In defence of untranslatability

Howard Sankey

IN DEFENCE OF UNTRANSLATABILITY

Howard Sankey

1. Introduction

This paper addresses criticisms of the concept of untranslatability which
Davidson [1984] and Putnam [1981] have raised against the
incommensurability thesis. The main themes of the criticism are present
in the following extract from Putnam [1981]:

> The incommensurability thesis is the thesis that terms used in another
> culture, say, the term 'temperature' as used by a seventeenth-century
> scientist, cannot be equated in meaning or reference with any terms or
> expressions we possess . . . [I]f this thesis were really true then we could
> not translate other languages—or even past stages of our own language—
> at all. And if we cannot interpret organisms' noises at all, then we have
> no grounds for regarding them as thinkers, speakers, or even persons.
> In short, if Feyerabend (and Kuhn at his most incommensurable) were
> right, then members of other cultures, including seventeenth-century
> scientists, would be conceptualizable by us only as animals producing
> responses to stimuli (including noises that curiously resemble English or
> Italian). To tell us that Galileo had 'incommensurable' notions and then
> to go on to describe them at length is totally incoherent.
>
> [1981: 114-5]

The central objection is that it is incoherent to talk about what is
untranslatable. Three lines of argument may be distinguished with regard
to this alleged incoherence. One argument is direct: it is incoherent to express
the content of an untranslatable language within the language into which
it is untranslatable. The other two are indirect arguments which assume
translation is necessary for understanding. First: ideas expressed in an
untranslatable language are incomprehensible, so claiming to understand
them is incoherent. Second: it is incoherent to conceive the speaker of an
untranslatable language as having a language at all. The direct argument
will be dealt with in Section 2, and the indirect arguments in Section 3.

In addition, Davidson [1984] argues that languagehood is inextricable
from translation. He claims that a 'dogma of a dualism of scheme and reality'
which fallaciously separates language from translation underlies the
incommensurability thesis. His attack on the dualism will be considered in
Sections 4 to 6.

2. The Direct Incoherence Argument

Putnam defines 'incommensurability' by saying that 'terms used in another culture . . . cannot be equated in meaning or reference with any terms or expressions we possess'. Given this definition, the direct incoherence argument is embodied in the last sentence of the quote: 'To tell us that Galileo had "incommensurable" notions and then to go on to describe them at length is totally incoherent'. For if Galileo's ideas are untranslatable into our language, then they cannot be expressed using our language, and it contradicts the claim of untranslatability to do so.

Davidson puts the point in the form of a paradox [1984: 183-4]. 'We are encouraged', he says, to 'imagine we understand massive conceptual change' by the use of examples, but 'the changes and the contrasts can be explained and described using the equipment of a single language' [183]. 'Kuhn', he adds, 'is brilliant at saying what things were like before the revolution using—what else?—our post-revolutionary idiom' [184]. The paradox is that the meaning expressed by the terms of an untranslatable language should be expressed in the very language into which translation allegedly fails.

Putnam and Davidson's remarks suggest the following argument. Suppose it is argued in language L that L* is untranslatable into L. Suppose as well that the argument in L employs examples from L* in the sense that it expresses the meaning of terms taken from L*. It follows from the latter that L* is translatable into L, for that is what expressing the meaning of terms from L* in L amounts to. But then the argument itself translates from L* into L in the course of arguing that L* is not translatable into L. If the argument is correct, then it is possible to translate from L* into L, so the conclusion is false. If the conclusion is correct, then it is impossible to translate from L* into L, so the argument is incorrect. Such an argument is incoherent.

This argument is sound but its scope is limited. Rather than being a general objection to untranslatability, it is a meta-argument to the effect that one form of untranslatability argument is self-refuting. Nothing follows from that about untranslatability itself. It is not even a general objection to all arguments for untranslatability. At most, it is a criticism of arguments in which untranslatability is argued for in the language into which translation fails. It does not apply if the language of argument and the untranslatable languages are distinct.[1] In fact, it only applies to arguments which employ examples. Arguments which do not express the meaning of untranslatable expressions are immune to such cricitism. As it is not fully general, the argument fails to show untranslatability to be incoherent. So it cannot be brought to bear on any particular untranslatability claim unless specifically shown to apply to it.

Of course, Putnam and Davidson employ the objection because they assume incommensurability falls within the ambit of the argument. They assume

---

[1] E.g., it might be argued in a metalanguage that a pair of object-languages is not intertranslatable without expressing untranslatable content in either object-language.

the language into which an untranslatable theory fails to be translatable is the language in which the argument for incommensurability is couched. Instead of translation failure between delimited theoretical terminologies, they identify the language into which translation fails with language as a whole.

This interpretation is explicit in Putnam's definition of 'the incommensurability thesis [as] the thesis that terms used in another culture . . . cannot be equated in meaning or reference with any terms or expressions we possess'. And it is evident in the inference he draws: 'if this thesis were really true then we could not translate other languages . . . at all'. When Davidson notes Kuhn's paradoxical use of 'our post-revolutionary idiom' to discuss pre-revolutionary science he assumes that the modern language into which out of date theory fails to translate is contemporary English. Davidson also takes the language into which translation fails to be a total language because he discusses incommensurability in the context of complete translation failure [1984: 190-1]. As Davidson and Putnam interpret incommensurability, the language of argument and the language into which translation fails are one and the same.

Given such an interpretation, it remains only to note that Kuhn and Feyerabend use examples extensively. Since they use examples in arguing for incommensurability, their argument for untranslatability is open to the charge of incoherence. For if the language into which translation fails is the very language in which they argue for untranslatability, then their .use of examples is indeed incoherent.[2]

This objection could be met by denying that Kuhn and Feyerabend express the meaning of the examples they discuss. One might claim that they give only approximate or partial translations. But this would be a misrepresentation. Kuhn and Feyerabend's exposition of the meaning of expressions is what shows them to be untranslatable in the first place.

It must be denied instead that the language into which translation fails is the language of argument. So it must be denied that incommensurability entails untranslatability into a total language. This accords fully with the thesis of incommensurability, since the incommensurability of scientific theories is not a relation between total languages. It is a relation between the languages of theories, and the language specific to a theory is only a part of a language. Incommensurability is due to semantical differences in the terminology of theories: the terminology employed by a theory cannot be translated into the terminology of a theory with which it is incommensurable. Instead of untranslatability into a total language, it is a case of translation failure between sub-languages within language as a whole.

That the untranslatability is limited is evident from Kuhn and Feyerabend's discussions of the theories they take to be incommensurable. They do not

---

[2] Even their use of examples independently of such arguments would be incoherent. For if the examples cannot be expressed in our language, then they cannot be expressed in any discussion couched in our language.

claim translation failure into a total language, since they are concerned with semantic analysis of the vocabulary the theories employ.[3]

Kuhn makes the point explicitly in responding to the Davidson-Putnam argument [1983: 669-71]. He advocates 'local' incommensurability, which is untranslatability between sub-sets of the terms used by pairs of theories. This involves localised semantic difference within the context of shared everyday and scientific language:

> Most of the terms common to . . . two [incommensurable] theories function the same way in both; their meanings . . . are preserved; their translation is simply homophonic. Only for a small subgroup of (usually interdefined) terms and for sentences containing them do problems of translatability arise.
>
> [1983: 670-1]

So he does not even claim full translation failure between the special languages of theories: untranslatability is restricted to a central complex of interdefined terms.

The situation is similar with Feyerabend, who holds that the basic principles of a theory preclude the concepts of a theory with which it is incommensurable: 'the conditions of concept formation in one theory forbid the formation of the basic concepts of the other' [1978: 69]. The resultant untranslatability affects more than a central complex of terms. Yet it is still a relation between the languages of theories, rather than total languages.[4]

The picture of language which emerges is of natural language as a conglomerate of terminologies or local idioms with special areas of application. Untranslatability between theoretical languages constitutes a relation between sub-languages within a total language. Rather than untranslatability into a total language as assumed by Davidson and Putnam, what is at issue is localised translation failure between sub-languages contained in an encompassing language.

Thus the language into which the vocabulary of a theory fails to be translatable may be distinct from the language of argument. For the argument that a pair of sub-languages is not intertranslatable can be couched in a portion of language distinct from the language into which translation fails. Theoretical sub-languages may themselves be the topic of a discussion carried out within some other part of the language.

Thus consider two languages TL and TL* associated with two theories. Suppose that TL and TL* are sub-languages of a broader natural language L. It is possible to use L as a metalanguage to speak about the semantic

---

[3] See Kuhn on Newton and Einstein [1970: 101-2], phlogiston versus oxygen [1983: 675-6]; Feyerabend on impetus and momentum [1981c: 62-9], and classical physics versus general relativity and quantum mechanics [1981d].

[4] Feyerabend does not restrict incommensurability to scientific theories. Languages, world views, frameworks and forms of life may also be incommensurable (cf. [1987: 81], [1981b: 16], [1975: 269]). But such broader application of the concept of incommensurability does not imply that the language into which the vocabulary of incommensurable theories is untranslatable is a total natural language.

relations between TL and TL*. In particular, it may be argued in L that a term t* of TL* cannot be translated into TL. Such an argument need not be formulated in TL, for it can be formulated in L. Using L as metalanguage, t* can be referred to and shown to be indefinable in TL without being expressed in TL. Nor is there any need in the course of the argument to express the meaning of t* in TL. For t* may be defined in L used as a metalanguage for TL without expressing the meaning of t* in TL.[5]

In sum, the direct incoherence argument does not apply to the incommensurability thesis. Since the untranslatability in question is a relation between theoretical sub-languages, and since such sub-languages may be discussed within a metalanguage, no incoherence attaches to the untranslatability argument.

3. Translation and Interpretation

The Putnam passage suggests two arguments which do not proceed strictly in terms of translation. Both involve the assumption that translation is necessary for interpretation. This assumption is implicit in Putnam's inference from 'we could not translate . . . at all' to 'we cannot interpret organisms' noises at all'.

Discussion of incommensurability is supposed to be incoherent because of such inability to interpret. The first argument is that if a language is untranslatable, then the ideas expressed in the language cannot be understood, so 'to describe [incommensurable notions] at length is totally incoherent'. The second is that speakers of an untranslatable language cannot be known to have a language, and 'would be conceptualizable by us only as animals producing responses to stimuli'.

It must be asked what Putnam means by 'interpret'. To say that translation failure entails inability to 'interpret organisms' noises' suggests that the meaning of untranslatable expressions cannot be understood. So to interpret an expression is presumably to understand what the meaning of the expression is. But it is unclear how interpretation is related to translation.

Putnam may assume that interpretation of a speaker who shares one's own language constitutes homophonic translation from the speaker's idiolect into one's own idiolect. If such domestic interpretation is assumed to be a form of translation, then it is natural to take interpretation of a foreign language as a form of translation as well. Interpretation of a foreign language would then consist in translating foreign expressions into a home language

---

[5] For an example of a discussion of untranslatability which has this form, see Feyerabend's discussion of 'impetus' [1981c: 65-6]. He mentions 'impetus' and gives its meaning as defined in the impetus theory; then he shows that such a definition cannot be formulated on the basis of the principles of Newtonian mechanics. This is a point about the semantical limitations of the language of Newtonian mechanics, and the meaning of 'impetus' is not expressed in that language anywhere in the argument. The discussion is couched in English used as a metalanguage, so the language of argument and the language into which 'impetus' fails to be translatable are distinct.

and understanding their home language equivalents. Thus conceived, failure to translate immediately entails failure of interpretation.

This translational sense of interpretation can be construed in two ways. If translation must be exact, interpretation of a foreign expression would consist in understanding an exact equivalent in a home language. If translation may be loose, interpretation may consist in understanding a loose rendering of a foreign expression.

If 'interpret' were given the first reading, Putnam's inference would fail. It does not follow from failure of exact translation that the content of speakers' utterances cannot be understood. For that to follow, it would have to be the case that exact translation is a necessary condition for understanding such content. But there is no reason to assume failure of exact translation to entail that a language cannot be understood at all. Such failure neither precludes the production of a gloss or loose translation, nor does it prevent the language from being learned directly.

To take interpretation in the second way as loose translation is to implausibly exaggerate incommensurability. Untranslatability of theories in the sense relevant to incommensurability does not entail a total lack of common semantic features: expressions of untranslatable languages may share aspects of reference and even meaning. Though 'dephlogisticated air' cannot be translated into the oxygen theory, some of its tokens co-refer with 'oxygen'. If 'interpretation' is taken in a loose sense, Putnam's denial of interpretation is stronger than licensed by incommensurability.

In the context of the argument interpretation cannot be taken to consist in translation. Interpretation must be separable from translation. Though in some cases interpretation may depend on translation, it cannot have translation as a constitutive component. To interpret an expression must be, quite simply, to understand what it means. And to understand an expression is not to translate it, nor is understanding restricted to what is expressed in a home language. Rather, to understand consists simply in knowing the meaning of an expression, whatever language it belongs to.

Putnam's inference from failure to translate to failure to interpret does not require that interpretation consist in translation. That inference can be made if translation is assumed to be a necessary prerequisite of interpreting a foreign expression. To say that translation is necessary for interpretation is distinct from saying that it is a component of interpretation.

The assumption that translation is necessary for interpretation is a restrictive assumption about the nature of understanding. As distinct from taking interpretation to be itself a form of translation, it takes understanding to be limited to expressions couched in one's home language. If the assumption were true, we would be unable to come to know the meaning of an expression not translatable into our language.

Let us now consider the first of Putnam's two arguments. It derives immediately from the assumption that translation is necessary for interpretation. Suppose that there is a language which we are unable to translate. Then, by the assumption, we cannot understand what is expressed

in the language. And in that case we cannot know what ideas are expressed by the speakers of such a language. But advocates of the incommensurability thesis do claim to know what expressions of untranslatable languages mean. Thus they say both that the expressions cannot be translated and that they know what the expressions mean. But this is incoherent: for if the expressions cannot be translated, their meanings cannot be known; and if the meanings are known, then the expressions can be translated.

Given the assumption of the necessity of translation, the conclusion of incoherence no doubt follows. However, the assumption is itself implausible. For understanding a foreign expression need not consist in understanding its translational equivalent within one's home language. Bilingual speakers do not translate 'in their heads' while conversing in a foreign language, so a bilingual may understand a foreign expression not translatable into his home language. Moreover, if a foreign language must be translated before it may be understood, then no language could ever have been translated in the first place. Still worse, if understanding a new language really did require translation into a prior language, it would be impossible to learn one's own first language. In any case, it is as a matter of fact unnecessary to translate in order to learn a second language, since it is possible to learn a new language by the method of direct immersion. For these reasons we may conclude that understanding a foreign language is not contingent upon translation and that the first argument may be rejected.

This rebuttal is patterned on the responses of Kuhn [1983] and Feyerabend [1987], who claim that the language of a theory incommensurable with one's own can be understood. Feyerabend rebuts Putnam by pointing out that 'we can learn a language or a culture from scratch, as a child learns them, without detour through our native tongue' [1987: 76]. Kuhn distinguishes between translation of a language and interpretation of an initially unintelligible language. He characterises interpretation as follows:

> Unlike the translator, the interpreter may initially command only a single language. At the start, the text on which he or she works consists in whole or in part of unintelligible noises or inscriptions . . . If the interpreter succeeds, what he or she has in the first instance done is learn a new language . . . whether that language can be translated into the one with which the interpreter began is an open question. Acquiring a new language is not the same as translating from it into one's own. Success with the first does not imply success with the second.
>
> [1983: 672-3]

This distinction enables Kuhn to rebut Putnam as Feyerabend does: an untranslatable theory may be interpreted, so there is nothing incoherent about claiming to understand the meaning of untranslatable expressions.[6]

---

[6] This rebuttal may be further supported by noting that the untranslatable language to be interpreted in the case of an incommensurable theory is not a total language. For what is at issue is untranslatability within a single language, and what must be interpreted is an unknown area of that language. Interpretation of theoretical terminology untranslatable into one's own theory is therefore not the radical project of learning a completely unknown language without the benefit of any common language.

Let us turn to Putnam's second argument, which is that untranslatability prevents language attribution. It too assumes translation to be necessary for interpretation. From this it follows that if a speaker's utterances cannot be translated, then it cannot be known what the utterances mean. And if no meaning can be attributed to the utterances of a speaker, then there is no evidence that the speaker has a language. But advocates of incommensurability describe speakers as having untranslatable languages and they attribute meanings to the speakers of such languages. But that is incoherent: for if utterances cannot be translated there is no evidence the speaker has a language; and if meaning is attributed to the utterances, that presupposes the speaker does have a language.

Our discussion of the first argument disposes of this argument's initial premise. For if translation is unnecessary then the meaning of untranslatable utterances can be known. However, we may also question the inference from inability to interpret a speaker's utterances to lack of evidence for language attribution. This inference is apparent in Putnam's remark that:

> if we cannot interpret organisms' noises at all, then we have no grounds for regarding them as thinkers, speakers, or even persons ... members of other cultures ... would be conceptualizable by us only as animals producing responses to stimuli.
>
> [1981: 114-5]

That argument would succeed if it were true that an untranslatable language cannot be recognised as a language. For it is indeed incoherent to deny possession of a language to an organism while saying of that organism that it possesses concepts which it expresses in language. What should be questioned is whether a speaker whose language we are unable to translate cannot be known to possess a language.

Putnam apparently assumes that if the meaning of sounds or inscriptions cannot be interpreted then there is no reason to take the organism which produces them to have a language. Davidson takes a similar view when he asks us to reflect 'on the close relations between language and the attribution of attitudes':

> On the one hand, it is clear that speech requires a multitude of finely discriminated intentions and beliefs ... On the other hand, it seems unlikely that we can intelligibly attribute attitudes as complex as these to a speaker unless we can translate his words into ours. There can be no doubt that the relation between being able to translate someone's language and being able to describe his attitudes is very close.
>
> [1984: 186]

With both Putnam and Davidson, the suggestion appears to be that knowledge of semantic content or propositional attitude is required to justify language attribution.

This suggestion is surely mistaken. Why should knowledge of semantic content be necessary for language recognition? Surely, formal and contextual features count for something. Codes may be recognised as codes without

being broken. Fragments of dead languages may be recognised as such prior to translation. Travellers recognise native speech as the local tongue without understanding it.

Why must psychological content be determined to identify behaviour as linguistic? In many social and physical settings the observed behaviour of humans is identifiable as linguistic without access to attitude or meaning. In any case, mental state need not be entirely inscrutable in the absence of knowledge of a language. The rough character of attitude or meaning can be known from observation of non-linguistic aspects of behaviour.

Even if there were no way to determine the presence of language without access to attitude or meaning, it would still not follow that an untranslatable language could not be identified as a language. The presumed necessity of content ascription by means of translation presupposes the necessity of translation for interpretation. To say that a speaker's meanings or attitudes can only be known if the speaker's language can be translated is to assume the only way to understand is via translation. But the possibility of direct understanding or acquisition of a language means that meaning and belief are interpretable without translation. A bilingual may determine psychological and semantical content for speakers of an untranslatable language without translating back into a home language.

In any case, the problem of recognising language is largely irrelevant to incommensurability. A rival theorist is not an organism whose possession of language is in question. Scientists with untranslatable theories may share a natural language. So the problem of whether a rival theorist possesses a language is resolved prior to discussion of theory. Nor is it as if the discovery of semantic variance between theories throws into question the status of a scientist as a speaker of language. For shared use of a background language is a precondition of narrowing a linguistic difference down to difference of theory.

This completes criticism of Putnam's two arguments. We will now briefly consider a related argument which derives from Davidson's discussion of interpretative charity [1984: 195-7].[7] Davidson applies the principle of charity to the problem of radical interpretation.

The problem is how to interpret meaning without independent access to belief: 'a man's speech cannot be interpreted except by someone who knows a good deal about what the speaker believes . . . and . . . fine distinctions between beliefs are impossible without understood speech' [1984: 195]. He assumes that 'the basic evidence for a theory of radical interpretation . . . [is] the attitude of accepting as true, directed to sentences'. But such evidence does not determine meaning: 'if we merely know that someone holds a certain sentence to be true, we know neither what he means by the sentence

---

[7] I say 'derives' advisedly. Davidson puts the principle of charity to a different use. However, the argument discussed in the text follows immediately from Davidson's analysis of interpretative charity. He uses the principle against partial translation failure and concludes that no sharp distinction between difference of language and of belief can be drawn [1984: 197].

nor what belief his holding it true represents' [196]. Charity is invoked to
extract meaning from the thin evidence of sentences held true.

To determine meaning, assumptions must be made about belief: 'if all
we know is what sentences a speaker holds true, and we cannot assume
that his language is our own, then we cannot take even a first step towards
interpretation without knowing or assuming a great deal about the speaker's
beliefs' [196]. Belief attribution should be governed by charity. The rough
idea is for the agent to come out on the whole as a believer of truths:

> We get a first approximation to a finished theory by assigning to sentences
> of a speaker conditions of truth that actually obtain (in our opinion) just
> when the speaker holds those sentences true. The guiding policy is to
> do this as far as possible, subject to considerations of simplicity, hunches
> about the effects of social conditioning, and of course our common-sense,
> or scientific, knowledge of explicable error.
>
> [1984: 196]

The principle of charity is justified because the agreement it provides is
a precondition of interpretation:

> Since charity is not an option, but a condition of having a workable
> theory, it is meaningless to suggest that we might fall into massive error
> by endorsing it. Until we have successfully established a systematic
> correlation of sentences held true with sentences held true, there are no
> mistakes to make. Charity is forced on us; whether we like it or not,
> if we want to understand others, we must count them right in most matters.
>
> [1984: 197]

The connections Davidson draws between interpretation, translation and
charity seem to license the following inferences. Since charity involves taking
sentences of our language which we hold true as the content of alien
utterances, charity implies translation. Since charity is necessary for
interpretation, successful interpretation entails translation. Therefore,
interpretation of an agent is inconsistent with translation failure. Thus to
interpret a scientist as having a theory untranslatable into one's own is
incoherent.

In effect, Davidson's use of the principle of charity combines Putnam's
two arguments. In accordance with the first argument, charity makes
translation·necessary for interpretation. In accordance with the second, it
makes translation necessary for interpreting an agent as a speaker. Two
objections may be raised against this use of the principle of charity.

In the first place, the link between charity and translation must be severed.
Just as interpretation does not require translation, interpretative charity does
not require translation. Davidson assumes that charitable interpretation of
an agent assigns truth conditions in a home language to sentences of an
alien language. But while such charity might be generally advisable, it is
not necessary. Charity may be applied directly within the alien language.
Charity may be incorporated into the direct method of language acquisition.
In learning a language directly without translating the interpreter can, and

perhaps should, assign maximum plausible truth conditions as well as reasonable belief. Interpretation seeks coherence and assigns plausible truth-values whether or not it results in a translation.

In the second place, charity is unsuitable for theoretical discourse. The principle of charity can be refined in various ways to allow for varying degrees of error. But the general principle of assigning maximal truth is unacceptable as a principle of interpretation when applied to theoretical languages. Maximal assignment of truth to the statements of a scientific theory overlooks the possibility of large-scale error. But the history of science abounds with theories that have been profoundly mistaken. Moreover, there are compelling epistemological reasons to take a fallibilist stance towards all theories, past and present. Surely, in the interpretation of scientific language no assumption about the truth of theoretical assertions should be made.[8]

Now, against this second objection, it might be argued that attribution of massive error makes behaviour unintelligible. That is, to deny of an agent that any of its beliefs are true is to make it inexplicable how it manages to engage in successful action. But to say that a theory is totally or mostly false is not to say that the entirety of an agent's beliefs are false. Moreover, a false theory can have true consequences and be put to practical use. And a theory which is strictly false but nearly or approximately true may serve as a guide for action.

The general policy of overall interpretative charity towards speakers should not be enjoined upon the interpreter of theoretical discourse. For the purpose of interpreting theoretical discourse, we are not therefore obliged by the principle of charity to impose translational equivalences upon scientific theories. So the forcing move from charity to intertheoretic translatability may be rejected. The possibility of interpretation does not rule out translation failure between theories.

4. The Scheme-Content Dualism

There is another side to Putnam's claim that we have no reason to take uninterpretable organisms as 'thinkers, speakers, or even persons'. Namely, uninterpretable linguistic activity is evidentially indistinguishable from non-linguistic behaviour. The point is clearer with Davidson, whose arguments in its favour we will consider in the following two sections:

> nothing . . . could count as evidence that some form of activity could not be interpreted in our language that was not at the same time evidence that that form of activity was not speech behaviour.
>
> [1984: 185]

The point against untranslatability is this: for neither an untranslatable language nor for non-linguistic behaviour can semantic content be given in our language; so inability to translate is indeterminate between being

---

[8] Davidson may intend to exempt theory from maximal assignment of truth, for he does say we should assign truth 'subject to considerations of . . . common sense, or scientific, knowledge of explicable error' [1984: 196]. But he fails to elaborate the point.

evidence that a language is untranslatable and that it is not a language at all.

As Davidson himself notes, to conclude that no evidence could show an untranslatable language to be a language 'comes to little more than making translatability into a familiar tongue a criterion of languagehood' [1984: 186]. Rather than assuming translatability to be a necessary condition of languagehood, Davidson argues that language is neither conceivable nor recognisable as such independently of translation.

The central argument of Davidson's [1984] is directed against 'the dualism of conceptual scheme and empirical content' which underlies the conception of language as independent of translation. This 'dualism' posits an opposition between language, which embodies a conceptual system, and reality, upon which that system imposes order. The opposition of scheme versus content bypasses translation and characterises language as something bearing the scheme-content relation to reality. Because this severs language from translation, Davidson must dispose of the dualism in order to show that evidence for an untranslatable language is indeterminate.

According to Davidson, the scheme-content dualism disconnects languagehood from translation as follows:

> something is a language, and associated with a conceptual scheme, whether we can translate it or not, if it stands in a certain relation (predicting, organizing, facing, or fitting) to experience (nature, reality, sensory promptings) . . . The images and metaphors fall into two main groups: conceptual schemes (languages) either *organize* something, or they *fit* it (as in 'he warps his scientific heritage to fit his . . . sensory promptings'). The first group contains also *systematize, divide up* (the stream of experience); further examples of the second group are *predict, account for, face* (the tribunal of experience). As for the entities that get organized, or which the scheme must fit . . . either it is reality (the universe, the world, nature), or it is experience (the passing show, surface irritations, sensory promptings, sense-data, the given).

[1984: 191-2]

Such a relation gives substance to languagehood not contingent upon an interlinguistic relation of translatability. Identification of a language need not therefore involve translation, but may be based on evidence of the right sort of relation between putative linguistic behaviour and the world. Thus the dualism allows a language to be recognised as such without translation into a home language.

Davidson gives a dual analysis of the scheme-content relation: either schemes organise reality or experience, or they fit reality or experience. The accent with the first pair of relations is on the taxonomic function of language; with the second it is on its predictive or explanatory function. Davidson argues against the organising idea that it does not give translation-independent content to the idea of a language. He argues against the idea of language fitting reality that it separates truth from translation and leads illegitimately to the idea of a true but untranslatable language.

5. Schemes Organise the World

The organising idea is that something which is a language is recognisable as such because of its classificatory function. Translation fails because languages arrange things differently.

Against this version of the dualism, Davidson first notes that only pluralities can be organised:

> We cannot attach a clear meaning to the notion of organizing a single object (the world, nature etc.) unless that object is understood to contain or consist in other objects. Someone who sets out to organize a closet arranges the things in it.
>
> [1984: 192]

He then argues that it can only be determined that a language organises things differently if the language can on the whole be translated:

> A language may contain simple predicates whose extensions are matched by no simple predicates, or even by any predicates at all, in some other language. What enables us to make this point in particular cases is an ontology common to the two languages, with concepts that individuate the same objects. We can be clear about breakdowns in translation when they are local enough, for a background of generally successful translation provides what is needed to make the failures intelligible. But we were after larger game: we wanted to make sense of there being a language we could not translate at all. Or, to put the point differently, we were looking for a criterion of languagehood that did not depend on, or entail, translatability into a familiar idiom. I suggest that the image of organizing the closet of nature will not supply such a criterion.
>
> [1984: 192]

So, while admitting extensional variance between languages, Davidson denies that 'organizing the closet of nature' gives translation-independent content to languagehood.

Extensional variance raises the possibility of translation failure. Davidson's tactic is to play down its scope. Rather than argue against semantic differences between languages, he argues that translation failure must be limited if it is to be intelligible: 'we can be clear about breakdowns in translation when they are local enough'. But this point is no objection to incommensurability, which is at most local translation failure. Theoretical sub-languages are embedded in larger languages, so translation between theories may well fail against 'a background of generally successful [albeit homophonic] translation'.

Davidson's crucial assumption is that it is necessary to translate to determine that a language divides the world up differently. This supports his conclusion that translation failure is intelligible only if it is local and occurs in the context of broad translational success. Of course, if either the assumption or the conclusion were true, then the idea of organising reality would not offer a means independent of translation for recognising a language.

But the assumption begs just the question at issue. It may be true that

to find out that the classificatory systems of languages differ, the languages must be understood. But Davidson simply assumes that translation is necessary for understanding another language and the classificatory system it embodies.

To assume this is to lose sight of the purpose of the argument. Davidson is arguing that the idea of organising reality gives no translation-independent content to the notion of being a language. To show this he has to argue that there is no way to determine whether a language organises reality without translation. But he simply assumes this. Surely, in the context of arguing that translation is necessary for the determination of classificatory difference, it begs the question to assume that translation is the only way to find out about such difference.

Davidson concludes that the 'image of organizing the closet of nature' does not enable sense to be made of total translation failure. This conclusion bespeaks a certain verificationism.[9] For it assumes that failure to specify a test for the presence of an untranslatable language entails that no content has been given to the concept of such a language.

Davidson's verificationism is evident in his inference from the intelligibility of only local translation failure to the unintelligibility of total failure. He allows that 'we can be clear about breakdowns in translation when they are local enough'. And he claims that general success in translation is what makes such breakdowns 'intelligible'. Davidson's point is that untranslatable linguistic material can only be known to be language if translation failure occurs in the context of overall translation of the language. Otherwise there would be no semantic evidence that the untranslatable material is linguistic. From this Davidson infers that sense has failed to be made of total failure: 'But we were after larger game: we wanted to make sense of there being a language we could not translate at all.'

The inference appears to be based on the following reasoning. The idea that language organises reality can only be applied to the local case given background success in translating the language. Therefore the organising idea does not yield a test for determining the presence of a totally untranslatable language. So that idea does not give meaning to the concept of such a language. This final inference assumes that meaning is only bestowed upon concepts if a means of verification is specified.

The problem with this can best be seen from Davidson's alternative formulation of his conclusion:

> Or, to put the point differently, we were looking for a criterion of languagehood that did not depend on, or entail, translatability into a familiar idiom. I suggest that the image of organizing the closet of nature will not supply such a criterion.
>
> [1984: 192]

By 'criterion' Davidson seems to mean a test for language, not an account of what being a language consists in. He concludes that because the criterion

---

[9] A number of authors have noted Davidson's implicit verificationism here: among them Rorty [1982: 5-6] and Blackburn [1984: 61].

is inapplicable without translation no sense has been made of full untranslatability.

But the 'image of organizing the closet of nature' specifies a function which a language may perform. This is a criterion of being a language, as opposed to a criterion for recognising one.[10] It gives content to the notion of being an untranslatable language: viz. such a language organises the world differently. Such a criterion of languagehood gives content to the notion of being an untranslatable language whether or not it can verifiably be fulfilled.

Such verificationism imposes a fallacious constraint on meaning. To impart meaning to a concept is not contingent upon coming up with a test for applying it. For it is possible to specify mistaken tests for applying concepts. What enables this point to be made with respect to a given concept is a grasp of its content which is independent of such tests.

Even if this were not the case, Davidson's attack would still be beside the point. That we do not understand the notion of total untranslatability is no objection to the idea of an untranslatable language. Even if we have no conception of what total untranslatability involves, no existence claim follows from that about such languages. Neither from inability to verify the existence of a totally untranslatable language, nor from inability to give content to the concept of such a language, does it follow that no such language exists.

6. Schemes Fit Experience or Reality

On the second version of the dualism, a conceptual scheme or a language enables us to deal with the world by explaining and predicting facts. Schemes are a way of 'coping with (or fitting or facing) experience' [193]. Such metaphors emphasise prediction rather than classification, and take us 'from the referential apparatus of language . . . to whole sentences':

> It is sentences that predict (or are used to predict), sentences that cope or deal with things, that fit our sensory promptings, that can be compared or confronted with the evidence.
>
> [1984: 193]

The relation between the two versions of the dualism appears to be this: schemes which organise the world differently provide alternative ways of coping with experience. Since it is 'sentences that cope' and the 'referential apparatus' from which sentences are built varies with scheme, sentences from alternative schemes may be untranslatable and yet deal adequately with the world.

Davidson first argues that the idea of fitting experience reduces to that of being true. Schemes account for all the evidence:

> a theory may be borne out by the available evidence and yet be false. But what is in view here is not just actually available evidence; it is the totality of possible sensory evidence past, present, and future.
>
> [1984: 193]

[10] Clearly, it cannot be a sufficient condition of being a language, but it is perhaps necessary.

To deal with all such evidence is just to be true: 'for a theory to fit or face up to the totality of possible sensory evidence is for that theory to be true'. There is no need to maintain a dichotomy between fitting all the evidence and being true:

> the notion of fitting the totality of experience, like the notion of fitting the facts, or of being true to the facts, adds nothing intelligible to the simple concept of being true.

[1984: 193-4]

Instead of two versions of what schemes fit we have this: 'something is an acceptable conceptual scheme or theory if it is true' [194]. Since fitting experience or reality thus reduces to being true, 'the criterion of a conceptual scheme different from our own now becomes: largely true but not translatable'.

This raises the question of whether 'we understand the notion of truth, as applied to language, independent of the notion of translation'. Davidson takes Tarski's theory of truth as constitutive of our understanding of truth. Convention T requires translation from an object-language into the metalanguage in which the truth-predicate is defined, so our understanding of truth depends crucially on translation. It is worth quoting Davidson's remarks in full:

> We recognize sentences like ' "Snow is white" is true if and only if snow is white' to be trivially true. Yet the totality of such English sentences uniquely determines the extension of the concept of truth for English. Tarski generalized this observation and made it a test of theories of truth: according to Tarski's Convention T, a satisfactory theory of truth for a language L must entail, for every sentence s of L, a theorem of the form 's is true if and only if p' where 's' is replaced by a description of s and 'p' by s itself if L is English, and by a translation of s into English if L is not English. This isn't, of course, a definition of truth, and it doesn't hint that there is a single definition or theory that applies to languages generally. Nevertheless, Convention T suggests, though it cannot state, an important feature common to all the specialized concepts of truth. It succeeds in doing this by making essential use of the notion of translation into a language we know. 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.

[1984: 194-5]

Davidson's attack on the idea that schemes fit experience or reality has two steps. The first step is the reduction of the idea to that of being true. The second is the argument that truth is inextricable from translation. The two steps are linked in that the idea of an untranslatable scheme being true divorces truth from translation.

The problem with the first step is that fitting experience does not reduce to being true as far as scientific theories are concerned. Scientific theories

may be, and often are, mistaken. More to the point, theories which 'fit the evidence' in the sense of being empirically adequate may be false; for a false theory may entail true predictions.

Davidson does, it is true, restrict attention to theories which fit 'the totality of possible sensory evidence past, present, and future'. But this simply removes actual science from the ambit of the argument. What he says can neither be about actual science nor is his argument relevant to examples that have been put forward of untranslatable theories. For rarely, if ever, do actual theories fit all the evidence, much less all the future evidence.

Certainly, there is no need to assume purportedly untranslatable theories to be true. To take but one example, the phlogiston theory and Lavoisier's oxygen theory were both to varying degrees false. To say that a pair of theories is incommensurable carries no commitment to their truth: it is not to say that they are both untranslatable and true.

In any event, a theory which fits all 'possible sensory evidence' is not *ipso facto* true. Perhaps a theory which fits all the facts, observable and otherwise, is true; but if it fits only the 'sensory evidence', it does not follow that it is true. So even if a pair of untranslatable theories were to fit all the evidence, there would be no reason to suppose both were true: to describe such a pair as incommensurable is not therefore to say that they are true and untranslatable.

Part of the trouble is the choice of metaphor. 'Fitting the evidence' suggests empirical adequacy, which amounts to truth at an empirical level. But in any sense in which theories 'cope with experience' they need not strictly 'fit the evidence'. Even successful theories in actual science only fit the evidence imperfectly. Theories have empirical difficulties from the start and never fit all the evidence. Yet they may still 'cope with experience' in the sense of explaining and predicting phenomena, solving problems, and guiding research. To say that such theories 'fit the evidence' in any but a loose sense is mistaken. There is even less reason to say that they are true.

Since incommensurability need not be a relation between true theories, this breaks the link between the two steps of Davidson's argument. However, Davidson's Tarskian argument cannot be evaded so easily. For untranslatability implies the possibility of true but untranslatable sentences. If a sentence can be formulated in a language, then ordinarily either it or its negation is true. If a sentence cannot be translated from one language into another, then neither can its negation be so translated. Since either the sentence or its negation is true, untranslatability raises the possibility of a true but untranslatable sentence. So Davidson's attack on the separation of truth from translation must be confronted.

Davidson argues that our concept of truth is defined for English and languages translatable into English, so our grasp of the concept does not extend beyond languages intertranslatable with English. Convention T does not define a general concept of truth for unspecified languages. Rather, it defines a truth-predicate for a specific language and for sentences of languages intertranslatable with it.

A theory of truth for a language which conforms with Convention T entails a set of T-sentences for the sentences of the language and their translational equivalents. Recurring to the previous quotation:

> according to Tarski's Convention T, a satisfactory theory of truth for a language L must entail, for every sentence s of L, a theorem of the form 's is true if and only if p' where 's' is replaced by a description of s and 'p' by s itself if L is English, and by a translation of s into English if L is not English.
>
> [1984: 194]

The set of English T-sentences defines the English truth-predicate for the sentences of English and translational equivalents:

> sentences like ' "Snow is white" is true if and only if snow is white' [are] trivially true . . . the totality of such English sentences uniquely determines the extension of the concept of truth for English.
>
> [1984: 194]

Since no T-sentences can be formed in English for sentences not translatable into English, the truth-predicate of English is not defined for such sentences, which fall outside its extension.

Thus our concept of truth is given by the definition of the English truth-predicate which is defined exclusively for the set of English sentences and translational equivalents. Such a concept of truth cannot be understood independently of translation. For it would not be constitutive of understanding that concept to understand it as applied to untranslatable sentences: it would not be that concept if so applied.

On the face of it, this seriously undermines any notion of translation failure which depends on a translation-independent concept of truth. The argument does not, however, have any implication about translation failure between parts of a single language. Translation failure between theories does not require a translation-independent concept of truth. Since the languages of theories are sub-languages of a background natural language, they may be discussed within the inclusive natural language employed as metalanguage. English may function as metalanguage and the English truth-predicate may be defined over its embedded sub-languages.[11] Since English contains both sub-languages there is no need to characterise truth in English for sentences not translatable into English. Hence Davidson's argument poses no threat to the thesis of untranslatability of theoretical sub-languages.

Beyond this, however, it can also be shown that Davidson's argument against the translation-independent concept of truth is problematic in its own right. In the first place, the argument does not achieve its aim. It is meant to show, as against the scheme-content dualism, that something crucial to being a language (true assertion) has no content divorced from translation. But in order to show that one could not discover a language which turned out

---

[11] For example, an English T-sentence for a sentence of the impetus theory may be formulated as follows: 'Projectile bodies have impetus' is true if and only if projectile bodies have impetus.

not to be translatable, it needs to be shown that a language could not be recognized as such without translation. What it purports to show instead is that truth is indefinable for untranslatable sentences. But that does not show that a language could not be identified as such from non-semantic evidence. If a language which proved resistant to translation were to be so identified, that would present *a posteriori* the existence of untranslatable truth. In denying that truth can be disjoined from translation, Davidson rules out untranslatable truth *a priori*. But no argument is offered from the connection between truth and translation to the conclusion that language is unrecognisable as such in the absence of translation. So far from showing the impossibility of such language recognition, the argument merely assumes it.

In the second place, there is an underlying tension between the purported truth-translation nexus and Davidson's concession of local translation failure. As we saw, Davidson allows that 'we can be clear about breakdowns in translation when they are local enough' [192]. But if a sentence of a language which is on the whole translatable into English should turn out not to be so translatable, what is to be made of the possibility of its truth?

According to Davidson, the English truth-predicate is undefined for any sentence untranslatable into English. So on Davidson's own account our concept of truth is inapplicable to such a sentence. Yet either such a sentence or its denial is true. Whatever sense Davidson thinks can be made of the idea of an untranslatable sentence, he seems not to allow sense to be made of its truth.

Now such isolated translation failure might be dismissed as unproblematic, since linguistic modifications may remove local untranslatability. The fact that truth-conditions cannot be given for isolated sentences need not preclude sense being made of their truth. For, suitably modified, the language may translate recalcitrant sentences and subsume them under its truth-definition.

But when does it become intelligible to apply the concept of truth to such a sentence? If the sentence must await actual translation, problems arise with translating the truth-predicate. For until such an untranslatable sentence can be translated, the truth-predicate defined in its language does not have the same extension as ours. If our concept of truth can be applied to such a sentence prior to the requisite alteration of our language, then our truth-predicate can be applied to sentences for which no T-sentence in our language can be formed.

In any case, to translate by altering a language is not strictly translation at all. If a sentence may only be translated by changing a language, then it cannot be translated into the unchanged language. But linguistic boundaries are fluid and arbitrary. No rules dictate when a fragment of a language becomes part of another or how large such a fragment may be. In principle, nothing prevents one language being appended in its entirety onto another. To permit application of the truth-predicate to sentences translatable by linguistic modification amounts to making the possession of truth-value depend on whether a sentence belongs to our language. But to have a truth-

value is not merely contingent upon belonging to our language. Nor does a sentence acquire truth-conditions only upon entry into our language.

In the third place, at least a *prima facie* case can be made that truth is separable from translation. Suppose one were to protest against Davidson that the concept of truth does not depend on translation. The Tarskian schema ' "s" is true if and only if p' is true if and only if p supplies a structural feature of truth which does not merely consist in a specification of the extension of 'true' for English. It is a constraint on the concept of truth such that nothing counts as a truth-predicate unless the sentence of which truth is predicated and the statement of truth conditions are equivalent. As against Davidson, the suggestion is that there is a general concept of truth of which the truth-predicates of particular languages are special cases.

To give some content to this claim, let us consider how one might come to recognise a truth-predicate for an untranslatable language. Consider a field linguist whom we may imagine to have encountered and mastered an alien language, call it 'Alien', which fails to translate into the linguist's home language, say English. What is to prevent such a linguist from recognising an Alien predicate whose use in Alien corresponds to the behaviour in English of the predicate 'is true'? Suppose the linguist identifies a predicate 'T' of Alien such that appending 'T' to a named Alien sentence 's' yeilds a sentence ' "s" is T' which is assertible when and only when 's' is assertible. Provided the linguist understands what 's' means and understands that ' "s" is T' is materially equivalent to 's', what reason could there be not to take 'T' as the truth-predicate for Alien?

Davidson's argument suggests the following objection to this proposal. Suppose the linguist reports in English, as regards the Alien sentence 's', that 's' is true. What does the linguist's report ' "s" is true' mean? Since truth-conditions cannot be given for 's' in English, the English truth-predicate cannot be used to say that 's' is true. So to say in English that 's' is true must mean that 's' is true-in-Alien, not true-in-English. But what does ' "s" is true-in-Alien' mean in English? 'True-in-Alien' is indefinable in English because no Alien truth-conditions are specifiable in English.

To give sense to saying ' "s" is true-in-Alien' in English one might say that 'true-in-Alien' is English for the Alien truth-predicate. The Alien truth-predicate and the English truth-predicate have similar functions in their respective languages. Each predicate behaves disquotationally: the result of appending either predicate to a sentence is a sentence assertible in identical circumstances to the original. In virtue of this formal resemblance both predicates instantiate a general truth-concept for particular languages, and 'true-in-Alien' can be used in English to translate the Alien truth-predicate.

It may be objected that the notion of 'true-in-Alien' is inconsistent with a semantic conception of truth. Since no truth-condition can be given for 's' in English, what it is to say ' "s" is true-in-Alien' in English cannot be defined in English.

Now, we may grant that the extension of the truth-predicate for a language is defined within the language by its T-sentences. No extensional specification

of 'true-in-Alien' can be given in English using English T-sentences since Alien is untranslatable into English. But it does not follow that no content can be given to 'true-in-Alien' in English. For the fact that the function of the Alien predicate is analogous to that of English 'true' enables 'true-in-Alien' to be defined as an English word for the Alien predicate which performs the same function in Alien as 'true' does in English.

It might be further objected that the Alien truth-predicate is not recognisable as such if it differs extensionally from English 'true'. It is not in virtue of disquotation that a truth-predicate is identifiable as such. In order to identify a truth-predicate, its extension must be determined. To identify such an extension as the extension of a truth-predicate, it must be the same extension as the extension of the English truth-predicate.

As against this, the way the imagined linguist recognises the Alien truth-predicate is precisely the same way in which the truth-predicate for English is identified. Given that the linguist understands Alien and recognises a predicate of Alien whose behaviour conforms to Tarski's schema, nothing further is required for recognising a truth-predicate. The objection reduces, in effect, to the previously criticised assumption that translation is necessary for understanding a language.[12]

University of Melbourne   Received May, 1988
and
Saint David's University College, Lampeter

BIBLIOGRAPHY

Blackburn, Simon [1984], *Spreading the Word*, Oxford: Oxford University Press

Davidson, Donald [1984], 'On the Very Idea of a Conceptual Scheme', in *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press, 183-198

Feyerabend, Paul K. [1975], *Against Method*, London: New Left Books

Feyerabend, P.K. [1978], *Science in a Free Society*, London: New Left Books

Feyerabend, P.K. [1981a], *Realism, Rationalism and Scientific Method: Philosophical Papers, Vol. 1*, Cambridge: Cambridge University Press

Feyerabend, P.K. [1981b], 'Introduction: Scientific Realism and Philosophical Realism', in [1981a] 3-16

Feyerabend, P.K., [1981c], 'Explanation, Reduction and Empiricism', in [1981a], 44-96

Feyerabend, P.K. [1981d], 'On the 'Meaning' of Scientific Terms', in [1981a], 97-103

Feyerabend, P.K. [1987], 'Putnam on Incommensurability', *British Journal for the Philosophy of Science*, 38, 75-81

Kuhn, Thomas S. [1970], *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press, 2nd ed.

Kuhn, T.S. [1976], 'Theory-Change as Structure-Change: Comments on the Sneed Formalism', *Erkenntnis*, 10, 179-199

Kuhn, T.S. [1983], 'Commensurability, Comparability, Communicability', *PSA 1982*, Vol. 2, East Lansing, Michigan: Philosophy of Science Association, 669-688

Putnam, Hilary [1981], *Reason, Truth and History*, Cambridge: Cambridge University Press

Rorty, Richard [1982], 'The World Well Lost', in *Consequences of Pragmatism*, Brighton: Harvester, 3-18.

---

[12] I am indebted to Henry Krips, Oscar Manhal and Graeme Marshall for discussion.