

# Radical Interpretation and Compositional Structure

Peter Pagin

In this paper I shall be concerned with the relation between a particular account of linguistic meaning and the property of compositionality in natural language.<sup>1</sup> The account, proposed by Donald Davidson, is that based on considerations about radical interpretation. I shall argue that there is a fundamental conflict between, on the one hand, the view that the meaning of expressions of natural languages is determined *purely* according to canons of radical interpretation, and, on the other hand, the view that natural languages exhibit compositional structure. I shall also argue that if there is such a conflict, this speaks against the proposed account.<sup>2</sup>

## 1. Radical interpretation versus compositionality

I think it is fair to say that over the years, from the early meaning theoretical papers and up to the attacks on the doctrines of linguistic conventions, the emphasis in Davidson's writings has shifted from linguistic structure to linguistic interpretation. The former is prominent in papers

---

1. By "compositionality" I understand a feature of a language, in virtue of which the compositionality principle holds:

The meaning of a complex expression is determined by the meanings of its parts and its mode of composition.

As stated there is much that must be further specified in order to make it reasonably precise, not least as regards the term "meaning". For now, by "meaning" I shall intend semantic properties which are relevant for the present discussion, not necessarily intensional ones.

I shall further assume that natural languages do exhibit syntactic structure (but not necessarily that there always is a uniquely correct syntactic analysis).

2. Note that I shall not claim that you cannot *combine* compositionality and radical interpretation. Clearly you can, and Davidson has impressively done so. I shall only claim that if radical interpretation is your *only* basic idea about linguistic meaning, then there is a conflict, for compositionality cannot be justified from that idea.

such as ‘Theories of meaning and learnable languages’, ‘Truth and meaning’ and ‘Semantics for natural languages’, and the latter particularly stressed in papers such as ‘Communication and convention’, ‘A nice derangement of epitaphs’ and ‘The social aspect of meaning’.<sup>3</sup> Roughly speaking, it seems that in the early papers, radical interpretation is regarded as a method for testing T-theories (i.e. semantic theories for individual languages), whereas later papers seem to assume a more independent and direct relation between linguistic meaning and radical interpretation.

My main aim in this paper is to explore the consequences of basing the *entire* account of meaning on considerations about radical interpretation. That Davidson seems to have moved in this direction is one reason for considering this option.<sup>4</sup> Another reason is that compositionality is an important property, perhaps the most important property, of human languages, and if we aim at a *fundamental* account of linguistic meaning, we cannot just take that property for granted. That natural language does have compositional structure is a tenet that needs justification, and if we aim at a fundamental account, that justification should flow from its assumptions. This is a reason for looking to radical interpretation for such a justification.

For this purpose I shall consider what I shall call the *RI account of meaning*. By the RI account of meaning I shall understand an account of meaning that involves the following two tenets:

1. A sentence, as uttered by a speaker *X* at a time *t*, means what it can be successfully interpreted as meaning.
2. Success of an interpretation is judged by the degree that the interpretation conforms to norms of *charitable* interpretation.

---

3. Because of space limitations I cannot go into the matter here. I did do that in an earlier version.

4. I don’t claim that Davidson has actually taken this position.

What the RI account of meaning is supposed to be depends, clearly, on what we include in “norms of charitable interpretations”. I do not want to be very specific about it. Clearly the ideas that Davidson himself, and some others, have advanced as making up “the principle of charity” must be included: that speakers are mostly (massively) right in their beliefs, that they are rational, coherent agents with reasonably stable beliefs, desires and preferences. But I would also want to include in these norms general features of thought and attitudes, which do not themselves concern linguistic structure. For instance, the idea of holistic individuation of thought content may well be counted among those norms, if it is a basic and general feature of thoughts, irrespective of particular properties of languages in which they are expressed. I only want to exclude principles which directly concern structural properties of languages or linguistic expressions.

I do not claim that Davidson has ever advanced the RI account of meaning. One reason to be cautious about it is precisely that Davidson in fact may have thought all along, as I do, not only that compositionality is an essential aspect of natural language meaning, but that this feature is not derivable from the RI account of meaning alone, and has to be added by separate considerations, such as our pre-theoretic knowledge of English as a language with significant semantic structure.

That compositionality is not derivable from the RI account of meaning alone is something I shall argue for. More precisely, I shall argue for the following three tenets:

1. The view that natural language speakers speak a compositional language, or the view that natural language interpreters have a compositional competence, cannot be justified from the RI account of meaning alone.
2. If natural languages have the semantic features which the most successful interpretational theories ascribe to them, then they are not compositional.
3. If 1 and 2 are correct, then the RI account of meaning is false.

The tenets can be alternatively rendered as follows:<sup>5</sup> Suppose we shall come up with a *theory* of

the language  $L$  of a speaker  $X$ , at a time  $t$ . The theory shall state the meaning of the sentences of  $L$ , and hence provide a long list of meaning specifications in some form.<sup>6</sup> The criterion of *truth* of such a theory is provided by the RI account of meaning. That is, the theorems are true if, and only if, utterances of  $X$  can be successfully interpreted as meaning precisely what the theorems specify. Then the question is whether there is any justification for the view that such a theory shall exhibit compositional structure, e.g. as Davidson's T-theories do. Tenets 1 and 2 can then be reformulated as concerning such theories of particular languages, or speakers at times.

## 2. RI and the justification of compositionality

I shall here consider a number of possible reasons for believing in the compositionality of natural language. I shall start with some classical arguments, and see how these fare under the RI account of meaning, and then I shall consider some further reasons derivable from this account itself. I shall try to show that in none of these cases do we get a justification of compositionality.

Two things should be kept in mind. First, I am *not* going to argue that *intuitively* successful radical interpretation, i.e. radical interpretation which is successful by ordinary, intuitive standards, does not require compositionality (indeed, the intuitive standards may themselves involve pre-theoretic conceptions of semantic structure). That question is considered in passing by Martin Davies.<sup>7</sup> As Davies argues, compositional structure “crucially facilitates radical interpretation”. But, as Davies also notes, that there is a practical requirement of structure, for the sake of interpretation, is “consistent with the theoretical possibility of a language that is partially or

5. This way of putting the matter was suggested to me by Barry Smith.

6. That form can be "s means that p" or "s is true iff p" or something else. Which it shall be can perhaps not be decided in advance. Davidson's original reason for preferring "is true iff" over "means that" was that the former is extensional, and hence allows that standard extensional logic may be applied in *deriving* the theorems. But the derivations are needed precisely in order for the theory to exhibit the compositional structure of the object language sentences. If there is no reason to believe in compositional structure, this reason for preferring "is true iff" over "means that" is not available.

7. Davies, 'Tacit knowledge, and the structure of thought and language', pp 149-152.

wholly lacking in semantic structure”. I agree with Davies on these points, but I am arguing something quite different. My point will be that if our basis is the RI account of meaning, i.e. if success of radical interpretation is judged by the principles of *charitable* interpretation (and not by intuitive standards), then there will be no justification of compositionality.

Secondly, I am not going to complain that the RI account cannot offer any *proof* of the compositionality principle, or that it cannot provide a *knock-down* argument. I don't believe that the compositionality principle (as applied to natural languages) *can* be proved, or that there are any knock-down arguments for it. I think there are good arguments for the principle, but these are not of the knock-down kind. The argument I favour is an inference to the best explanation, or, somewhat stronger, an inference to the only reasonable explanation. But in both forms, the argument is disputable.

What I shall complain is that the RI account does not offer any reasonable argument at all. Or, to be more precise, I shall try to show that all the candidates below, which are those I (and others, so far) can think of, are unsatisfactory, but not really that no better candidates can be had.

## 2.1 Learnability

One of the classical arguments for compositionality is based on the idea that natural languages must be learnable, since speakers do in fact learn them. This was Davidson's own argument.

When we can regard the meaning of each sentence as a function of a finite number of features of the sentence, we have an insight not only into what there is to be learned; we also understand how an infinite aptitude can be encompassed by finite accomplishments. For suppose that a language lacks this feature; then no matter how many sentences a would-be speaker learns to produce and understand, there will remain others whose meanings are not given by the rules already mastered. It is natural to say that such a language is *unlearnable*. This argument depends, of course, on a number of empirical assumptions: for example, that we do not at some point suddenly acquire an ability to intuit the meanings of sentences on no rule at all; that each new item of vocabulary, or new grammatical rule, takes some finite time to be learned; that man is mortal.<sup>8</sup>

---

8. 'Theories of meaning and learnable languages', pp 8-9.

This argument presupposes that natural languages do have an infinite (or extremely large finite) number of meaningful sentences (with different meanings), and a few rather uncontroversial things about human cognitive capacity. Let's say that such a language is *very rich*. If we are allowed to assume that very rich languages exist as entities to be learned, then we do have a good argument for compositionality. That a very rich language has a compositional structure is (part of) a good explanation of how a speaker can learn it. It is probably the best explanation, and possibly the only reasonable explanation. The question is whether we are allowed to make the assumption.

In the present context, that question comes down to the question whether the assumption can be justified on the basis of the RI account of meaning. Since this account proceeds by way of interpretation of individual speakers, that justification must involve at least three steps. In the first step we would have to find a justification for ascribing to a single speaker the property of being a speaker of a very rich language. In the second we would have to ascertain that two or more speakers speak the same very rich language. And in the third we would need a reason for believing that at least one of them has learned that language from the others.

Now clearly, if we could justify the assumption that speakers speak a very rich language *without* compositional structure, then we know in advance that the learnability argument is flawed, because then, if speakers learn such a language from each other we know that it cannot be correctly explained by means of compositionality, simply because we know that the language is not compositional. On the other hand, if we can justify directly the assumption that speakers do speak very rich *compositional* languages, then we already have an argument for compositionality, and need not add any extra considerations about learnability.

The only reason for turning to learnability as an argument would be that we can justify the assumption that speakers do speak very rich languages without being able, directly, to justify the belief that they are compositional. But I cannot see at all how we could end up in that situation, for I don't think there could be any good reason for ascribing to a speaker the property of

speaking a very rich language *unless* on the basis of believing that a compositional theory gives the best account of the finitely many utterances he has actually made.

If this is correct, then the learnability argument fails, in the sense that compositionality would already have to be justified if we had good reasons for the premisses of the learnability argument. In fact, I think that this holds generally, not just within the context of the radical interpretation approach to meaning.

In this context, the conclusion is even sharper. For the judgement that a speaker does speak a very rich language must be based on radical interpretation, issuing in a theory about the speaker's language. That theory is either compositional or non-compositional, definitely rendering the learnability argument either superfluous or flawed. And if it is added that radical interpretation must respect compositional structure because of the learnability argument, then the whole reasoning is viciously circular, since the learnability argument, if it isn't flawed, depends on the assumption of compositional outputs of radical interpretation.

## **2.2 Understanding new sentences**

Closely related to the argument from learnability is the argument from understanding new sentences. It is said that a language must be compositional in order for it to be possible to understand new sentences, sentences you have never used or encountered before. This argument, however, can be understood in two different ways.

One way to understand it is to think of a language as an independent entity, containing many sentences. Only a fragment of all of its sentences have actually been used by its speakers. Frequently, however, speakers are confronted with sentences never used before, and they manage to understand them. Compositionality is then invoked for explaining this.

Again, however, we must ask for a justification of the assumption that there are sentences which have never been used and which nonetheless have a well determined meaning. I think the situation here is exactly parallel to the situation we had with the learnability argument. The only reasonable justification of the view that sentences never used before do have a well determined meaning (differ-

ent from meanings of sentences that have been used), is compositionality. And so there will be no need to add extra considerations about the understanding of new sentences.

The other way to understand the argument is more effective. We shift to utterances, or sentences as used by speakers. Then we can ask how it is possible for an audience to interpret a new sentence, which he may never have used or encountered before, as the speaker *intends it* to be interpreted. And the natural answer is that it can be understood because of its compositional structure. The speaker employs the compositionality of his language in order to find an expression of his thought, and by grasping the meaning of the parts and the mode of composition the interpreter manages to understand the utterance in the right way.<sup>9</sup>

This does provide a good argument for compositionality on certain empirical assumptions. If we can take for granted that speakers often do understand each other, and often right away, even when they use sentences not used before, and express thoughts not expressed before, then we have a good reason for believing that we need compositionality for explaining this. But from the viewpoint of the RI account of meaning, this *cannot* simply be taken for granted. Rather, whether two speakers understand each other is something that would have to be *ascertained* by applying radical interpretation.<sup>10</sup> Either we interpret them both, and check whether they understand each other, or they apply the methods of charitable interpretation to each other. Either way, we will again find ourselves in this situation: if the two speaker are successfully interpreted as speaking a compositional language, then we don't need special considerations about

---

9. This situation is depicted in the opening passage of Frege's 'Gedankengefüge' ('Compound Thoughts'), and Frege also provides compositionality as the explanation, in the form of a structural correspondence between the sentence and the thought.

10. Davidson explicitly says that the assumption that two speakers speak the same language is in need of justification. "Speakers of the same language can go on the assumption that for them the same expressions are to be interpreted in the same way, but this does not indicate what justifies the assumption. All understanding of the speech of another involves radical interpretation." 'Radical Interpretation' pp 125-26.

Indeed, assuming that A and B interpret each other homophonically, the question whether they understand each other is equivalent to the question whether they speak the same language.

communication for justifying compositionality, and if the speakers are successfully interpreted as speaking non-compositionally, then an argument from communication will be flawed, from the standpoint of the RI account.

So the results about the compositionality, or non-compositionality, of the speakers' languages must be established *before* the question whether they understand each other can be answered. We must therefore do without assumptions about mutual understanding when it comes to justifying compositionality within the RI framework.

What we need to ask is whether it is a *necessary* feature of interpretation that linguistic communication, normally and typically, proceeds by employing compositional structure. And it seems that, from the standpoint of RI, it is not. If we are to get an argument for compositionality from the understanding of new utterances, we must get it on the assumption that the interpreter has the knowledge needed for understanding the utterance *before* the utterance is made. His linguistic competence, including knowledge of compositional structure of the speaker's language, should allow him to know *in advance* what that particular sentence, as used by that speaker, means, or would mean if uttered, because that is precisely what compositionality is good for. If we put it in terms of having a theory, as Davidson often does, we can say that the interpreter should be able to *predict* what the new sentence will mean in the mouth of the speaker.<sup>11</sup>

But there is *no* general requirement in the RI account of meaning that the interpreter be able to predict meaning. All that is required is that the interpreter can arrive at the right interpretation, even if he can do so only *after* the utterance has been made, exploiting some extra knowledge he has gained in the circumstances of the utterance, or later (even much later).

Davidson himself stresses this when criticizing the idea that linguistic meaning as such is determined by conventions. There is no general requirement of successful prediction (and that is why, according to Davidson, conventions aren't necessary for linguistic communication).<sup>12</sup> Correctness of

---

11. It is not, of course, a matter of predicting what a speaker will say. It is a matter of knowing what a sentence would mean if it were uttered by the speaker (barring, of course, slips of tongue and other aberrations).

an interpretation is judged solely by the norms of charitable interpretation. And charity can only be applied to attitudes already revealed, i.e. with respect to utterances *already made*. Unless compositionality is assumed, interpretations of *unuttered* sentences are irrelevant to the evaluation of the interpretation scheme as a whole.<sup>13</sup> If you add compositionality, certain interpretations of sentences uttered commit you to certain interpretations of sentences not uttered, and so induce predictions. But if you only have charity to begin with, predictions are not induced and not required.<sup>14</sup>

Hence, within the framework of the RI account, we cannot appeal to compositionality for explaining the understanding of new sentences. So within this framework, this version of the classical argument for compositionality fails. Rather, for justifying compositionality within RI, we have to look for specific features of the account itself.

### 2.3 Regularity

It is often stressed that for a speaker to be interpretable, he must use language in a regular way. If we could not find any correlation between his assent to sentences and features of his surrounding, then we would never be able to interpret even his observation sentences. And if he often changed his linguistic habits, suddenly starting to express different thoughts with the same

---

12. 'Communication and convention' p 278. and 'A nice derangement of epitaphs' pp 440-44. It seems that in these papers Davidson has been thinking about prediction in terms of possible deviation from previous use. Understanding is possible in spite of deviation, which is why you don't need conventions. But the possibility of deviation is irrelevant when compositionality is at stake.

13. You may abstract away from the idea of a real interpreter needing real observations, and think of the principle of charity as applied directly to the speaker's attitudes, whether revealed or not. Still, the point against the RI account remains, since there would be no requirement of interpreting any sentence to which the speaker does not already have any attitude (like holding true). The point would be lost only on the assumption that a speaker in general has an attitude to every sentence of his own language.

14. In 'The structure and content of truth', p 313, Davidson says: "Thus, a theory of truth is a theory for describing, explaining, understanding, and predicting a basic aspect of verbal behavior." There is no doubt that Davidson wants predictive capacity, as there is no doubt that he wants compositionality. But the question is how to get the requirement of predictive capacity out of the criteria for correct interpretation, when judged by charity. As far as I can see, you plainly don't.

old sentences, then again we could never get evidence enough for a correct interpretation.

I am not sure to what extent this is true, but be that as it may. The point I am making does *not* depend on the assumption that the speaker does, or even may, deviate in any way from her previous language or speech habits. The speaker may be perfectly regular. My point is just that you don't get compositionality from the demand that a speaker speaks regularly in the sense of never changing the meaning, or use, of any single sentence. That demand does not even suggest compositionality. To motivate compositionality there would have to be a more complicated kind of regularity, in the use of sentence *parts*, since what is needed is a pattern for relating sentences which the speaker has *not* so far used to those sentences which she *has* used. You don't get this kind of regularity by requiring only that the speaker be consistent in her use of each single sentence. This places no requirement at all on sentences not yet used.

The question is whether anything belonging with the RI account can motivate the requirement of semantic regularity in the use of sentence parts. Again, my challenge to the RI account does not depend on any assumption on my part that the speaker *can* vary the use of sentence parts over time. We may well assume that the speaker doesn't, and cannot, but the question is how to learn about it from the principles of the RI account.

## 2.4 Holism

That the interpreter need not be able to predict the meanings of new utterances does not, however, imply that the interpreter need not have a system of interpretation. There can be other considerations, based on the norms of charitable interpretation, that forces the interpreter to be systematic, and perhaps compositional, too.

Perhaps the holistic nature of thought provides one such consideration. So let us assume that thought is holistic in roughly the sense of Davidson's presentations of this idea:

There are good reasons for not insisting on any particular list of beliefs that are needed if a creature is to wonder whether a gun is loaded. Nevertheless, it is necessary that that there be endless interlocked beliefs. The system of such beliefs identifies a thought by locating it in a logical

and epistemic space.<sup>15</sup>

In order to entertain the thought that the gun is loaded, e.g. in the form of a desire that it be loaded, the speaker must have a number of true beliefs about guns, about being loaded, and these other beliefs will interlock with yet further beliefs. And the same goes for the interpreter. So the interpreter cannot ascertain that the speaker is expressing a belief that the gun is loaded unless he can also ascertain that the speaker has a great number of other beliefs. And this in turn requires the interpreter to provide interpretations of a number of other utterances of the speaker. The interpreter must therefore interpret not only a great number of utterances of the speaker, but the beliefs and other attitudes he thereby ascribes to the speaker must form a coherent pattern. The interpretations must support one another.

This much does, I think, follow from the holistic nature of thought, together with principles of coherence and rationality as norms of interpretation. But the next step, to the conclusion that there must be some compositional structure in the expressions of the attitudes, is not taken. To be sure, if there are semantic relations between attitudes expressed by sentences, then there are also corresponding semantic relations between the sentences expressing them. But there is still no requirement, and it is not even suggested by these considerations, that there be *syntactic* relations, between those sentences, that mirror the semantic relations (e.g. that, for at least some sentence parts, sentences containing that part are *about* the same thing). It is quite sufficient that the thoughts themselves form a system of the required kind, and that the utterances and sentences are interpreted as expressing them.<sup>16</sup> Syntax isn't touched by holism.

Note that I do *not* claim that it is possible for a speaker to *have* a coherent system of beliefs without speaking compositionally. In fact I believe that to be impossible. I only claim that as far

---

15. Davidson, 'Thought and talk' p 157.

16. As far as I understand, Davidson slurs over this distinction when he writes "If we suppose, as the principle of charity says we unavoidably must, that the pattern of sentences to which a speaker assents reflects the semantics of the logical constants, it is possible to detect and interpret those constants.", in 'The structure and content of truth', p 319.

as the standards of charitable interpretation go, there is no reason to doubt that possibility.

A possible further point about holism is this. Since holism requires the existence of a whole system of beliefs as a condition of having thoughts, and since every thought must be expressible by a sentence, it might be that the sheer required *number* of meaningful sentences could only be realized by a compositional language. That would be very plausible, if infinitely many beliefs, and therefore infinitely many sentences, were required. But that does not seem to be required by holism, on any conception of holism that I know of.<sup>17</sup>

Even so, it can be claimed that it is inconsistent with holism that a person has a fixed finite number of beliefs.<sup>18</sup> Rather, a person frequently forms new beliefs. Therefore, there cannot be a fixed finite number of sentences in a person's language, for the new beliefs must be expressible also, and this might be held to imply that the language expressing the beliefs must be compositional.

But this is not right either. As long as a person is free to fit the words to his thoughts as he pleases, he can freely make up the new sentences, one by one, which he needs to express his new thoughts. There is nothing for compositionality to explain there, and therefore no justification of it from holism. Of course, it can be objected that we are not free to mean whatever we want by our sentences, and I would say that this claim is correct, but the reasons for this go far beyond holism. So they give no support whatsoever of the idea that holism can justify compositionality.

## 2.5 Underdetermination

You might, however, want to protest here: it should not be enough to assign meanings to expressions to the effect of ascribing a coherent system of attitudes to the speaker. For clearly, unless you impose some further constraint on how to match sentences with thoughts, you can come up with many different attributions, since many different sets of attitudes form a coherent system. In order

---

17. The relation between holism and compositionality is studied in my 'Is compositionality compatible with holism?'

18. I see no reason myself why this should be inconsistent with holism, even though, of course, it is quite implausible in itself.

eliminate this massive underdetermination, you must assume a pattern of correspondence between the thought and the expression of the thought.

Indeed, underdetermination is one of the basic themes in Davidson's account of meaning. He has reminded us many times that an assertive utterance depends on two factors: what the speaker believes and what his words mean. We can interpret the words one way, provided we are prepared to attribute to the speaker the corresponding belief, but for each utterance there are many pairs of belief and meaning that can explain why the speaker holds the sentence true. Moreover, Davidson has given as a reason for not attributing propositional attitudes to animals that the attributions will be heavily underdetermined by the evidence for them.<sup>19</sup>

And clearly, by requiring compositional structure, we can reduce the number of possible interpretations, since then we require that syntactic relations between sentences correspond to semantic relations between them, and many otherwise possible interpretations will fail to satisfy this requirement.

But there are three problems to solve before you have the right to conclude that radical interpretation requires imposing compositional structure. First, you must explain why the plurality of possible interpretations is unacceptable. That is, you must explain why the method of charitable interpretation fails to distinguish between correct and incorrect interpretations when compositional structure is not required. The alternative view is that

Indeterminacy of meaning or translation does not represent a failure to capture significant distinctions; it marks the fact that certain apparent distinctions are not significant.<sup>20</sup>

That is, if you trust your method of interpretation you will say that it does capture all that is objective about meaning, and that intuitions to the contrary are misleading. Davidson did think so about his preferred method of radical interpretation (which, of course, included compositionality).

---

19. 'Thought and talk', p 164.

20. 'Belief and the basis of meaning', p 154.

The general question is why some underdetermination is acceptable and some not. Until there is a basis for drawing the line that can be derived from the RI account of meaning itself, it has not been shown that this account will force such further restrictions on interpretation as we asked for.<sup>21</sup>

The second problem is this. Even if it can be established that the RI account of meaning must involve further restrictions in order to reduce underdetermination, there is so far no reason why that restriction should be precisely compositional structure. Any odd restriction, like trying to match numbers of letters in words with number of words in sentences, can serve to reduce underdetermination. Too great underdetermination is itself just a problem. It does not direct you to the solution.

The third problem is an extra twist to the second. There may be some reason, derived from the norms of charitable interpretation, for preferring some extra restrictions over others. You may want the restrictions you choose to preserve those interpretations which fulfil the norms to the highest degree, and which rule out others. You will then get compositionality as your preferred restriction, if the compositional interpretations do get the highest marks. There is, however, reason to doubt this. I return to this question in the next section.

---

21. In several passages, e.g. in 'Reality without reference' pp 224-225, Davidson compares indeterminacy of meaning with measurement. His idea is that the difference between two acceptable T-theories for the same speaker is like the difference between the Fahrenheit and Centigrade temperature systems. The numerical values are different, but the assignments are still equivalent. As the pattern of assignment of numbers to physical states is the same in these two systems, so the pattern of assignment of truth conditions is the same across acceptable T-theories.

This suggests that acceptable T-theories can be mechanically transformed into each other by changing the semantic values of some parameters. If this is taken seriously, then we might say that the underdetermination of T-theories by evidence is too great if such a mechanical transformation is not possible between all accepted T-theories.

Unfortunately this suggestion remains a suggestion. But suppose that a non-trivial notion of mechanical transformation between T-theories is worked out. Then, however, as far as I can see, it is an unfounded hope that the set of the best T-theories contain only mechanically equivalent elements. But should it then turn out that *adding* the requirement of compositionality has this effect, i.e. the effect that all the best T-theories are mechanically equivalent, then there would be reason to ask whether this is a good argument for compositionality within the RI-account. Fredrik Stjernberg directed me to this thought.

## 2.6 The interpreter's theory

I want to conclude this section by considering Davidson's claim that we still do need a recursive representation of the interpreters competence at any moment of speech transaction.<sup>22</sup> Davidson claims that this is necessary, since there are infinitely many utterances which the interpreter can interpret, and must be able to interpret, at any given time.

However, as far as I can see, Davidson has himself undercut the reasons for taking this view. Even if, from purely syntactic considerations, we take seriously the idea that the competent interpreter must, at any given time, be able to interpret infinitely many possible utterances by the speaker, there is still no requirement that the interpreter be able to give correct interpretations of any such possible utterance until after it has been made, with the possibility of gathering the required evidence after the event.

And even if the interpreter does have a recursive prior theory, and the speaker does make an utterance correctly interpretable by that theory, there is no requirement, in the RI account of meaning, or, to my knowledge, in Davidson's considerations on interpretation, that the interpreter rely only on that prior theory, and indeed, no requirement that he rely on it at all. He might as well start from scratch. Of course, he cannot start from scratch in respect of evidence. If he treats each new utterance as the first piece of empirical evidence, then he will not get anywhere. But he can always start from scratch in respect of *theory*. What matters is how well a theory measures up to the norms of interpretation, not how old it is.

I agree with Davidson that actual interpreters do have an infinite competence, one which does require a recursive representation, or at least a recursive ingredient in its representation, and I think it is clear that they do use it, but I also think that the reasons for this view are quite external to the RI account of meaning.

---

22. 'A nice derangement of epitaphs', p 441, 'The social aspect of meaning', p 5.

### 3. Which are the best theories?

One natural way, and obviously the best way, if it can be had, to justify compositionality from the RI account of meaning, is to point out that, when we compare the different possible interpretations of a speaker, in respect of their satisfying the norms of charitable interpretation, it will be the compositional interpretations that come out with the best results. If I believed this to be the case, then I would not think there is any problem. But I don't.

The reason is simple. The norms of charity are supposed to be met by way of maximizing, or “optimizing”, certain properties of the speaker's attitudes, e.g. maximizing the frequency of truths among his beliefs. And it is quite clear that it will be much easier to reach such a maximum if your interpretation isn't hampered by the restrictions of compositionality. You can, if you want to, decide not to interpret any utterance as expressing a belief you disagree with, or a desire you find absurd. Or, if you prefer to think that the speaker is wrong about certain things, interpret his utterances so that he does reveal false beliefs about those things. Whatever the norms, whatever the properties you wish to maximize, you will find it easier to maximize those properties without compositionality, and whatever your background anthropological theory about the speaker, you will find it easier to interpret him as confirming that theory if you need not impose compositionality on your interpretation.<sup>23</sup>

So with all likelihood, *ad hoc* non-compositional interpretations will get higher marks than compositional alternatives. At this point you may object that the compositional theories stand out precisely by assigning meanings to the infinitely many sentences in the speaker's language, and that the non-compositional alternatives don't.

I think that objection is twice mistaken. First of all, no strong reason has so far emerged for thinking that the speaker does have a language with infinitely many sentences, each with a well determined meaning. Until you can provide such a reason, you should not count it an advantage of a

---

23. You can, if you like, also impose a vacuous compositionality, i.e. one that satisfies a formal definition but could not serve any explanatory purpose. For examples of this, see Janssen, 'Compositionality', and Westerstahl, 'Purported mathematical arguments for the vacuity of compositionality'.

theory that it does assign meanings to infinitely many sentences.

But secondly, the objection is technically wrong. You can, quite trivially, assign meanings to the infinitely many sentences which, syntactically, belong to the speaker's language. To take the most trivial option: assign some arbitrary meaning, the same meaning, e.g. the meaning *that the moon is made of cheese*, to each of the infinitely many sentences that the speaker still has not revealed any attitude to. That will most clearly be wrong, but as long as the speaker has not revealed any attitude to any of them, there is no empirical evidence against it, and so no problem with charity.

There will be evidence against it, as soon as the speaker e.g. uses one of those sentences to make an assertion. But, when this happens, you simply revise your theory to accommodate the new data, by assigning a different meaning to the uttered sentence. There could not be anything wrong with that, from the point of view of the RI account of meaning, since that account does not require predictive success of interpretational theories. All that is asked for is successful accommodation, i.e. correct interpretation after the event. If predictive success had been required, then clearly *ad hoc* non-compositional theories would be quickly eliminated. But it is not, and your non-compositional theories still come out as the best.<sup>24</sup>

And if you say that you need to impose restrictions, in order to reduce underdetermination, the sound, conservative line to take is to impose restrictions which allow you to keep your best theories, i.e. some of the non-compositional ones, and eliminate others. So the non-compositional ones still come out as the best.

If this conclusion stands, there is an even stronger result. It is not just that you cannot justify compositionality from the RI account of meaning. If it will always hold that your best interpre-

---

24. It is natural to protest here that the compositional theory will be better than the proposed *ad hoc* rivals simply because it will be more true, i.e. contain a greater number of true meaning assignments. But recall that the criterion of truth of an interpretational theory, as far as this discussion goes, is given by the clauses of the RI account. A theory is true only to the extent that it conforms, as good as any theory, to the norms of charitable interpretation in its account of utterances made.

tational theories are non-compositional, then the proper conclusion is that natural languages *are* non-compositional.

#### 4. Why compositionality is preferable to radical interpretation

If you think that natural languages are compositional in nature, and you agree with what I have said before, you will think that what should be rejected is the RI account of meaning. In fact, you will think that, since those interpretations that come out as the best according to the norms of charitable interpretation, are non-compositional, and thus deviate from the correct compositional ones, the interpretations favoured by the RI account of meaning are simply wrong. They assign the wrong meaning to many sentences and other expressions.

But is there any fundamental reason for believing that natural languages are compositional, or, if you prefer, that natural speakers are to be compositionally interpreted? I think there is, and that the reason, in a way, is the Davidsonian version of the old reason. I think that the best explanation, and the only reasonable explanation, of why natural language speakers understand each other so often, so easily, in so diverse and complicated matters, and often with so little background knowledge about the other speaker, is that they understand each other through the grasp of the composition of what is spoken.<sup>25</sup> For an unbounded number of sentences ordinary speakers are in a position to predict what a sentence would mean in the mouth of a fellow speaker, since in many cases the clues provided in the speech situation explain very little, and in many cases there are no such clues worth mentioning. Yet understanding is immediate, and thus cannot reasonably depend on anything else than knowledge that the interpreter had before being exposed to the utterance.

25. There is a further question about why to think that speakers do understand each other, why to think that speakers have thoughts and also manage to get knowledge about each other's thoughts through language. I think there are good things to say about that, too. That must be done, I think, for answering Stephen Schiffer's attacks, in *Remnants of Meaning* and many other places, on the idea that compositionality is required in an account of linguistic communication.

In general there is a difficult question about the extent to which speakers must be ascribed some *tacit knowledge* of compositional principles as part of the compositionality explanation of understanding. This question has been discussed e.g. by Gareth Evans, Crispin Wright and Martin Davies.

I would guess that Davidson agrees with this. What he does not agree with, is my view of the philosophical significance of everyday communication:

I do not think I have ever conflated the (empirical) question how we actually go about understanding a speaker with the (philosophical) question what is necessary and sufficient for such an understanding. I have focused on the latter question, not because I think it brings us close to the psychology of language learning or use, but because I think it brings out the philosophically important aspects of communication while the former tempts us to speculate about arcane empirical matters that neither philosophers nor psychologists know much about. So let me say (not for the first time): I do not think that we normally understand what others say by consciously reflecting on the question what they mean, by appealing to some theory of interpretation, or by summoning up what we take to be the relevant evidence. We do it, much of the time, effortlessly, even automatically. We can do this because we have learned to talk pretty much as others do, and this explains why we generally understand without effort much that they say.<sup>26</sup>

I pointed to the fact that if we disallow ourselves what we *can* get from an explanation of actual interpretation, it may be difficult to get it at all, since it may not be available from an account of radical interpretation. In the case of linguistic conventions this is not the case: you surely do not get them from an account of radical interpretation, but neither do you get them, as far as I can see, from an explanation of actual interpretation. I agree with Davidson that the notion of a convention does not do any explanatory work. I further agree with Davidson that sharing a language with anyone else is not essential to having a language, speaking meaningfully, or thinking. I am convinced, however, that in the case of compositionality, the opposite holds: compositionality does do essential explanatory work for actual interpretation, and compositionality is an essential feature of the meaning of sentences that have it. But I have argued that compositionality cannot be derived purely from consideration about radical interpretation. If this is right, then the radical interpretation approach to linguistic meaning is methodologically misguided.<sup>27</sup>

Stockholm University

---

26. 'The social aspect of language', p 3.

## References

- Davidson, D, 1984, *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford
- ‘Theories of meaning and learnable languages’, in Davidson 1984
  - ‘Truth and meaning’, in Davidson 1984
  - ‘Semantics for natural languages’, in Davidson 1984
  - ‘Radical interpretation’, in Davidson 1984
  - ‘Belief and the basis of meaning’, in Davidson 1984
  - ‘Thought and talk’, in Davidson 1984
  - ‘Communication and convention’, in Davidson 1984
  - ‘Reality without reference’, in Davidson 1984
  - ‘A nice derangement of epitaphs’, in LePore (ed), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Basil Blackwell, Oxford 1986
  - ‘The structure and content of truth’, *Journal of Philosophy* LXXXVII, 1990
  - ‘The social aspect of meaning’, in McGuinness and Oliveri (eds), *The Philosophy of Michael Dummett*, Kluwer, Dordrecht 1994
- Davies, M, *Meaning, Quantification, Necessity*, London, Routledge and Kegan Paul, London 1981
- Davies, M, ‘Tacit knowledge, and the structure of thought and language’, in C Travis (ed), *Meaning and Interpretation*, Blackwell, Oxford 1986.
- Evans, G, ‘Semantic theory and tacit knowledge’, in Holtzman and Leich (eds), *Wittgenstein: to Follow a Rule*, Routledge and Kegan Paul, London 1981.
- Frege, G, ‘Gedankengefüge’, in *Logische Untersuchungen*, Vandenhoeck and Ruprecht, Göttingen 1966.

---

27. I am indebted to Hermann Cappelen, Donald Davidson, Kathrin Glüer, Sören Häggqvist, Mattias Högeström, Mikael Janvid, Ernie LePore, Dag Prawitz, Gabriel Segal, Barry Smith, Fredrik Stjernberg, Folke Tersman, Dag Westerståhl and Åsa Wikfors.

Grandy, R E, 'Understanding and the principle of compositionality', *Philosophical Perspectives* 4, 1990

Janssen, T, 'Compositionality', *Handbook of Logic and Language*, Elsevier, Amsterdam 1997

Pagin, P, 'Is compositionality compatible with holism?', *Mind & Language* 12, 1997, pp 11-33.

Schiffer, S, *Remnants of Meaning*, MIT Press, Cambridge, Mass. 1987

Westerståhl, D, 'Purported mathematical arguments for the vacuity of compositionality', forthcoming.

Wright, C, 'Theories of meaning and speaker's knowledge', in Wright, *Realism, Meaning and Truth*, Basil Blackwell, Oxford 1987.