

# 4

## Davidson on Explicit Performatives and Mood-setters

ULF DALVAD BERTHELSEN

Scandinavian Institute, University of Aarhus, Denmark

I det følgende giver jeg en kort præsentation af Donald Davidsons semantiske analyse af eksplicitte performativer og modus. Jeg diskuterer hans strategi for at løse de semantiske problemer, der knytter sig til modus og argumenterer for, at hans analyse lider under det faktum, at han accepterer illokutionær kraft som et semantisk primitiv. Dette medfører, at han i stedet for at løse et semantisk problem blot flytter problemet fra den semantiske analyse af eksplicitte performativer til den semantiske analyse af det, han kalder sætningens 'mood-setter'.

### 1. INTRODUCTION

Donald Davidson's approach to natural language semantics is founded on two main ideas:

1. Natural language is a special case of rational behaviour.
2. Extensional first order predicate logic has all the power needed to describe and explain natural language meaning.

In brief, this means, firstly, that it does not make sense to detach the study of natural language meaning from the studies of action and rational behaviour and, secondly, that reference and predication are the basic means by which language users relate to the external world.

In his paper *Moods and Performances* (Davidson 1979) Davidson addresses a number of semantic problems related to grammatical mood along these lines. He suggests a Tarski-style semantics that adds to the truth conditions of the propositional content a set of truth conditions for the mood-setter, i.e. the

expression (e.g. grammatical mood or change of word order) that indicates the illocutionary force of the utterance in question.

In the following I will give a brief presentation of Donald Davidson's semantic analysis of explicit performatives and grammatical mood. On the basis of this presentation I will discuss his strategy for solving the semantical problems related to these matters and argue that his analysis suffers from the fact that he takes illocutionary force to be a primitive. Thus, instead of eliminating a semantically unanalysed element he simply removes the problems from the analysis of explicit performatives just to reinstate them in the semantic analysis of the mood-setter.

## 2. DAVIDSON'S ANALYSIS OF EXPLICIT PERFORMATIVES

As his starting point Davidson takes Austin's analysis of the so-called explicit performatives (Austin 1962), because they present what Davidson calls "an intolerable discrepancy between the semantics of certain first-person present-tense verbs and their other-person other-tense variants" (Davidson 1979: 117). According to Davidson this discrepancy is due to Austin's division between indication of illocutionary force and representation of propositional content and arises if we accept that the question of truth only concerns the propositional content and not the expression that indicates the illocutionary force of a given utterance.<sup>1</sup>

- (1) *Jones asserted that it is raining.*
- (2) *I assert that it is raining.*

The two examples illustrate Davidson's point. (1) is a report of the content of Jones' assertion and (2) is an explicit performative. The main verb in both sentences is indicative and the only differences are: 1. a shift from past tense to present tense and 2. the substitution of a pronoun for a proper name. Since Davidson aims at an extensional truth conditional semantics it is obviously intolerable if 'assert' in (2) does not contribute to the truth conditions of (2) in the same way that 'asserted' does in (1). (1) is true if and only if Jones asserted that it is raining no matter if it was actually raining at the time of the utterance. This, on the other hand, is not the case in (2). Here it seems reasonable to say that a speaker who has uttered (2) has said something true only if it was in fact

raining at the time of the utterance. To use Austin's terminology we could say that (2) is happy if it was raining at the time of the utterance and unhappy if it was not, whereas the happiness of the utterance of (1) is absolutely independent of the truth value of the embedded clause in (1). This, however, does not tell us whether or not the main clause can be said to represent a set of truth conditions, and it does not tell us anything either of how the main verbs in (1) and (2) are semantically related.

To overcome these problems Davidson suggests that we abandon what he calls the 'usual semantics' for explicit performatives. What Davidson refers to is a variety of theories based on Frege's approach to semantics. Frege (1892) defines the meaning of the embedded clauses in speech reports and attitude reports as their '*ungerade Bedeutung*', which means that they, as opposed to non-embedded sentences that have truth values as their extensions, are regarded as referring expressions that have their own intension (i.e. their meaning in the Fregean metaphysical sense) as their extension. There are at least two reasons for Davidson's rejection of this approach. Firstly, he wants to avoid at all costs a semantic theory that involves any kind of intensionality and, secondly, the Fregean approach seems to present the same difficulty as Austin's approach, namely, that we end up with two different meaning concepts (direct and indirect meaning) for the same word.

Instead, Davidson employs the strategy presented in *On Saying That* (Davidson 1968). The main features of this strategy are 1. that truth is ascribed to utterances, 2. that indirect speech reports are treated as two distinct utterances, and 3. that 'that' is treated as a referring expression used to refer to utterances. Applying this to (1) and (2) we get:

- (3) *Jones asserted that. It is raining.*
- (4) *I assert that. It is raining.*

When uttered, (3) and (4) each consists of two distinct utterances of which the former contains an expression that refers to the latter. Henceforward I will use 'explicitation' to refer to the first type of utterance and 'content' to refer to the second. We can now make the following generalization:

(5)  $x \alpha$  that.  $p$ .

It is now clear what Davidson has in mind. Speech reports and explicit performatives share the same logical form. Whenever  $x$  is replaced by a singular referring expression,  $\alpha$  is replaced by 'said' or any performative verb, and  $p$  is replaced by any first order sentence<sup>2</sup> we get a configuration of sentences that, when uttered, will constitute either a speech report (Jones said that.  $p$ .), a report of a performed speech act (Jones promised that.  $p$ .), or an explicit performative (I promise that.  $p$ .).

A clear advantage of this result is that we can now use the same semantic theory to treat utterances of ordinary descriptive sentences and utterances of explicit performatives. In either case the verb in the main clause will be treated as a relation that yields either T or F whenever the referring expressions that are taken as arguments have proper extensions. Thus, we need no other semantic concept than that of truth condition to explain the meaning of the explication.

In order to specify the relation between the explication and the content, Davidson introduces a somewhat unorthodox manoeuvre. Instead of invoking some kind of intensional concept Davidson simply conceives of the relation as referential. The two utterances are related by the referring expression 'that' that serves as the second argument in the explication. This means that (1) is true if and only if Jones asserted that it is raining and that (2) is true if and only if the person in question by uttering (2) did in fact assert that it is raining. In either case the truth value of the content does not affect the truth value of the explication. What could affect the truth value of the explication would be a situation where 'that' for some reason failed to refer, where the actual content of Jones assertion did not match the content of the speech report, or where the person uttering the explicit performative did not for some reason satisfy the proper requirements of assertion.

Granted that we accept the enrichment of our ontology with utterances, Davidson has now provided a purely extensional semantics for explicit performatives – or has he? There is still one important issue left to be dealt with, and it raises the question whether Davidson's analysis is at all satisfactory. The

problem is that the analysis, despite its virtues, so far has told us nothing about the fact that even though the relation between the explicitation and the content is not truth functional the content of an explicit performative might be used as one of the premises in a future deduction, whereas this could never be the case for the content of a speech report. To put it differently, the content of explicit performatives can sometimes be asserted, whereas the content of speech reports can never be asserted. This suggests that Davidson has not succeeded since there remains a tinge of intensionality to the content of the speech reports.

Davidson's solution is to say that the content of explicit performatives in certain situations can be uttered non-assertively (Davidson 1979: 119). Davidson does not spell out exactly what it means to utter something 'non-assertively', but it might be thought of as a special kind of illocutionary force that is used when the speaker does not want to be committed to the truth value of the utterance, which is indeed the case with the content of both speech reports and attitude reports.<sup>3</sup> In other words, Davidson uses this special non-assertive force as a defence against intensionality as he maintains that there is nothing more to the concept of meaning than what can be expressed by specifying truth conditions within the framework of extensional predicate logic: Words always have the same meaning, i.e. their contribution to the truth conditions is the same no matter where they appear. The only thing that differs is the force with which the sentence in which they appear is uttered.

### 3. DAVIDSON'S ANALYSIS OF MOOD

With the idea of non-assertion as a special kind of illocutionary force rather than as lack of illocutionary force, we have finally reached Davidson's analysis of mood. Having gone through his analysis of explicit performatives the analysis of mood seems rather straightforward since the semantic analysis in many respects is similar to the analysis of explicit performatives.

However, there is one important deviation that has to do with the semantic representation of illocutionary force. When moving on to the semantic analysis of grammatical mood, an obvious approach seems to be the one presented by David Lewis in *General Semantics* (1972). Lewis suggests that we simply reduce all other moods to indicative by paraphrasing imperatives, interrogatives<sup>4</sup> etc. into explicit performatives constituted by two distinct indicative sentences. Yet,

Davidson rejects this approach because it does not satisfy a number of requirements that he poses for a proper treatment of mood:

1. It must show or preserve the relation between indicatives and corresponding sentences in other moods; it must, for example articulate the sense in which 'You will take off your shoes', 'Take off your shoes', and 'Will you take off your shoes?' have a common element.
2. It must assign an element of meaning to utterances in a given mood that is not present in utterances in other moods. And this element should connect with the difference in force between assertions, questions, and commands in such a way as to explain our intuition of a conventional relation between mood and use.
3. Finally, the theory should be semantically tractable. If the theory conforms to the standards of a theory of truth, then I would say all is well. And on the other hand if [...] a standard theory of truth can be shown to be incapable of explaining mood, then truth theory is inadequate as a general theory of language. (Davidson 1979: 115-16)

In relation to Lewis' suggestion the most important of these requirements is the second. Apparently Lewis' theory does not satisfy this requirement since he reduces all other moods to the indicative. This reduction seems to be the cause of Davidson's reluctance: Not only because it reduces the different moods to indicative, but also because this reduction apparently has as a consequence an elimination of all other types of illocutionary force than the assertive force.

Davidson's answer to this problem is to approach mood in the same way as he approached the explicit performatives. The only exception is that the mood of a given sentence cannot be paraphrased into an explicitation since that would yield precisely the corresponding explicit performative. The reason for Davidson's rejection of the paraphrase strategy is not entirely clear, but he seems to believe that there are some special irreducible qualities related to the different moods and that the semantic analysis of mood must specify exactly what that quality is. In order to do so Davidson starts out by saying that all sentence types, indicatives as well as non-indicatives, consists of two distinct parts: 1. a mood-setter and 2. an indicative core. The idea behind this move is the same as before. By splitting the utterance in two distinct parts it becomes possible to specify two distinct sets of truth conditions. Thus, the utterance of the mood-setter irrespective of its manifestation (inflection, change of word order, tone of voice

etc.) becomes an independent part of the speech act representing its own set of truth conditions. Unfortunately Davidson is not very specific on this point, and instead of suggesting a formal procedure of specifying truth conditions for mood-setters he gives us a somewhat intuitive formulation:

The mood-setter of an utterance of 'Put on your hat' is true if and only if the utterance of the indicative core is imperatival in force. (Davidson 1979: 120)

In this case it is the truth conditions for the imperative that has been specified, but similarly we could specify the truth conditions for the other moods. Bearing in mind that the mood-setter is related to the utterance of the indicative core by a referring expression (exactly as we saw it with the explicit performatives), we can now generalize Davidson's formulation by saying that the mood-setter is true if and only if the accompanying indicative core is uttered with exactly the illocutionary force indicated by the mood by which the mood-setter is represented on the surface level (inflection, change of word order, tone of voice).

This is in short Davidson's semantic analysis of mood, and we can now return to the three requirements to see how they are actually met by Davidson's proposal. As regards the first requirement Davidson shows how the split of the semantic interpretation allows the same indicative core to be related to a variety of different mood-setters without any changes in its semantics. Thus, the requirement is met in the sense that whenever we have different grammatical transformations of the same indicative core, we have a systematic way of showing that it is in fact the same indicative core but related to a different mood-setter.

As to the second requirement Davidson points out that the relation between mood and mood-setter is systematic since a particular mood always expresses the same meaning. To be more specific, the utterance of say imperatives always counts as an utterance of the imperative mood-setter that in turn always represents the same set of truth conditions. This, however, does not mean that the relation between mood and mood-setter is conventional in the sense that a particular mood can only be used to perform the type speech act to which it is conventionally related. Think for instance of Searle's famous example 'Can you pass the salt?' or of the utterance of 'Are you coming' followed by the utterance

of 'That is an order!' Instead, it means that the relation between a particular mood and its mood-setter is conventional in the sense that the concept of illocutionary force is always part of the meaning expressed by a particular mood since the concept of illocutionary force is part of the truth conditions represented by the mood-setter of that particular mood. This ensures 1. that the use of a particular mood (by convention) always expresses the same meaning (i.e. represents the same set of truth conditions) regardless of which type of speech act it is actually being used to perform, and 2. it ensures that each mood expresses something unique since each mood-setter is uniquely related to a particular illocutionary force.

Thus, according to Davidson, we have a way of explaining how a sentence in a certain mood can be used to perform a speech act of another type than that expressed by its conventional meaning (cf. 'can you pass the salt?'). Unfortunately, Davidson's presentation lacks transparency. Therefore, it is not absolutely clear how he envisages the relation between the conventional mood-setter and the one changing the force of the utterance, but one way looking at it is this: Each mood always contain as part of its meaning the notion of a particular illocutionary force. However, in some cases the utterance of a sentence in a particular mood contains an element that indicates that the force expressed by the mood is overruled. This element might be an additional utterance ('This is an order'), a harsh tone of voice or something quite different, but the point is that the meaning of the original mood, despite the overruling, is still present.

Finally, if we accept that the first requirement is met, we must also accept that the third requirement is met. This is due to the fact that Davidson's split of the semantic interpretation that allows him to isolate the indicative core is exactly what also allows him to apply the semantic framework sketched in *Truth and Meaning* (Davidson 1967b). Thus, the theory is semantically tractable and it seems as if Davidson has succeeded in constructing a theory that satisfies all of the three requirements.

#### 4. THE MEANING OF 'ILLOCUTIONARY FORCE'

Opponents of pure extensional semantics might reject Davidson's theory all together, and even if we accept his foundational ideas mentioned in the introduction there are still problems left that need clarification. The most



interesting but also the most peculiar of these problems is Davidson's use of the notion of illocutionary force. Since Davidson is a notorious anti-metaphysicist, it seems strange that he apparently takes the notion to be primitive in the sense that it turns up undefined in the truth condition for the mood-setter. A possible explanation of his employment of the notion of illocutionary force is that it is supposed to be a defence against proposals like Lewis' which Davidson believes to be a token of inappropriate reductionism. The fact that Davidson puts a lot of effort into constructing his theory so that it satisfies the second requirement suggests that this is in fact the case. This leaves us with two obvious possibilities: either Davidson is right in his belief, and we might have to accept the dubious status of the notion of illocutionary force, or he is wrong and his analysis collapses into a version similar to Lewis'.

As announced I will argue that he is wrong, but I will also argue that this does not lead to a semantic reduction since the meaning that Davidson tries to capture by introducing the notion of force into the specification of the truth conditions of the mood-setter is captured perfectly well by paraphrasing the moods into explicit performatives.

The first step is to try to specify exactly what it means for an utterance to be uttered with a certain illocutionary force. If we compare it with the celebrated Tarski example "'snow is white' iff snow is white" we remember that the interpreter is supposed to have a clear intuition of the meaning of the meta-language sentence following the bi-conditional, i.e. he is supposed to know that 'snow' refers to snow and 'white' to the set of white entities etc. Since this is the semantic approach Davidson has chosen for the interpretation of the mood-setter, it seems fair to ask what clear intuition the interpreter is supposed to have of the meaning of the meta-language expression 'illocutionary force'.

One way of specifying this is to answer the following question: What conditions must an utterance satisfy in order to belong to the extension of predicates such as 'imperative in force', 'interrogative in force' etc. (from now on referred to as illocutionary predicates)? The most elaborate attempt to formalize the notion of illocutionary acts is made by Searle and Vanderveken (Searle & Vanderveken 1985; Vanderveken 1990), who, drawing on Searle's earlier work (Searle 1969), define illocutionary force by notions such as direction of fit, preparatory

conditions, and sincerity conditions. For the present purpose we can leave out the details and just assume that there is a set of necessary and sufficient condition for each of the different types of speech acts. Now, if this assumption is correct it gives us what we asked for, namely, a clear intuition of what it means for an utterance to be e.g. imperative in force. It simply means that all members of the set of necessary and sufficient conditions defining the imperative force are satisfied. For the imperative force this means among other things that it has world-to-word direction of fit, that the speaker prefers the outcome of being obeyed, and that the speaker has the proper authority.

With this semantic interpretation of Davidson's mood-setter we now turn to the explicit performatives in order to compare the semantics of the mood-setter with that of the explicitation of the explicit performative. If we accept Davidson's proposal that 'that' should be treated as a referring expression, the explicitation turns out to be a ternary relation between a subject (the speaker), a direct object (the receiver), and an indirect object (the utterance immediately following the explicitation). The standard semantic interpretation of di-transitive verbs says that a sentence (or, to follow Davidson, the utterance of a sentence) containing a di-transitive verb is true if and only if the ordered triple consisting of the extension of the subject, the extension of the direct object, and the extension of indirect object does in fact belong to the extension of the di-transitive verb. Now, what does this mean in relation to explicit performatives? Let us take an example:

- (6) *'I order you that you read 'The Concept of Truth in formalized Languages' for tomorrow'.*

(6) is true if and only if the speaker (denoted by 'I') orders the hearer (denoted by 'you') the content of the utterance of 'you read 'The Concept of Truth in Formalized Languages' for tomorrow' (denoted by 'that'). Now, in order to evaluate the truth value of (6) we need to be a bit more specific as to what criteria we might apply in order to decide whether or not the ordered triple actually belongs to the extension of 'order'. An obvious candidate for such a set of criteria is, of course, the set of conditions defining the Searle/Vanderveken notion of the illocutionary force for orders. If we accept this, it follows that the

utterance of (6) is true if and only if the three members of the ordered triple together satisfy these conditions.

Thus, it turns out, provided that we accept Davidson's interpretation of 'that' and the Searle/Vanderveken definition of illocutionary force, that there is in fact no semantic difference between the interpretation of Davidson's mood-setter and the explicitation of the explicit performative. Consequently semantic reductionism does not follow from paraphrasing non-indicatives into explicit performatives. On the contrary the use of the different non-indicatives appears to be equivalent to the use of their corresponding explicit performatives.

So, it seems that Lewis was right after all. It does seem to be a plausible assumption that the non-indicatives can be reduced to explicit performatives consisting of two indicative sentences. More interestingly, it follows from the analysis that this does not lead to any kind of semantic reductionism, since Lewis' proposal not only meets the first and the third of the requirement posed by Davidson on a proper treatment of mood, but, as it turns out, also the second. This is so, because we now have a way of showing 1. that the uniqueness of the different mood is captured entirely by the illocutionary predicates and 1. that since the concept of illocutionary force is part of the meaning of the illocutionary predicates, the relation between explicitations and use is conventional in exactly the same way as Davidson believes the relation between mood and use to be.

With this, it has also been shown how Davidson's employment of the notion of illocutionary force fails to give the result he had hoped for. The way he states his theory leaves only two possibilities. Either the notion of illocutionary force is undefined, in which case Davidson has done nothing but replacing Austin's non-truth conditional indication of force with his own similarly dubious notion of force, or the notion of force is defined, in which case the meaning of the mood-setter has to be equivalent to the explicitation of the explicit performatives.

## 5. DAVIDSON TODAY

Despite the fact that Davidson is considered one of the most influential philosophers of the late twentieth century, his semantic analysis of intensional

contexts and illocutionary acts has not received much attention. Neither this nor his dubious employment of the notion of illocutionary force does, however, mean that his semantic approach to these matters is irrelevant for contemporary linguistics and philosophy of language.

Indeed, it turned out that Davidson's rejection of Lewis's paraphrase strategy was not convincing, but accepting Lewis's paraphrase strategy does not mean, that we are compelled to accept his semantic framework as well. Thus, the above discussion is not meant as an attempt to reject Davidson's approach to intensionality. Rather, it should be taken as an attempt to eliminate an obscure metaphysical notion from the Davidsonian framework in order to make it more plausible as an alternative to formalized approaches to natural language semantics such as Lewis's.

Since Davidson and Lewis wrote their landmark papers in the late sixties and seventies, formal semantics has developed tremendously, and the different scientific communities with an interest in formal semantics has almost entirely chosen in favour of the semantic tradition pioneered by Lewis and perhaps most prominently Richard Montague. This is mainly due to the fact that a lot of people from linguistics as well as from philosophy and computer science have shown a great deal of interest in the computational aspects of natural language semantics. Consequently, there seems to be a broad consensus that Davidson's semantic project is a dead end and that his philosophy, except for his event semantics (Davidson 1967b), has little to offer contemporary semantics. This, however, seems to be a somewhat unfair conclusion. It might be that the semantic framework outlined in Davidson (1967b) faces some formal difficulties, but that is no argument against the foundational views that underlies Davidson's semantic program. Thus, the insight that action and rational behaviour is an important key to understanding natural language meaning poses a great challenge to some of the most important contemporary semantic theories (see e.g. Asher & Lascarides 2003 and Blackburn & Bos 2005). The eagerness to compute meaning has led to theories that focus predominantly on interpretation and, therefore, have little to say about the role of the speaker and of the nature of reference and predication construed as acts performed by rational agents.

REFERENCES

- Asher, Nicholas, Alex Lascarides (2003), *Logics of Conversation*, Cambridge: Cambridge University Press.
- Austin, John L. (1962), *How to Do Things With Words*, Oxford: Oxford University Press.
- Blackburn, Patrick, Johan Bos (2005), *Representation and Inference for Natural Language*, Stanford/Cal.: CSLI Publications (CSLI Studies in computational Linguistics).
- Davidson, Donald (1979), "Moods and Performances", in: Davidson, Donald (2001), *Inquiries into Truth and Interpretation*, Second edition, Oxford: Clarendon Press.
- Davidson, Donald (1968), "On Saying That", in: Davidson, Donald (2001), *Inquiries into Truth and Interpretation*, Second edition, Oxford: Clarendon Press.
- Davidson, Donald (1967a), "The Logical Form of Action Sentences", in: Davidson, Donald (2001), *Essays on Actions and Events*, Oxford: Clarendon Press.
- Davidson, David (1967b), "Truth and Meaning", in: Davidson, Donald (2001), *Inquiries into Truth and Interpretation*, Second edition, Oxford: Clarendon Press.
- Frege, Gottlob (1892), "Über Sinn und Bedeutung", in: Frege, Gottlob (1994), *Funktion, Begriff, Bedeutung – fünf logische Studien*, Göttingen: Vandenhoeck & Ruprecht.
- Lewis, David (1972), "General Semantics", in: Davidson, Donald, Gilbert Harman (eds.) (1972), *Semantics of Natural Language*, Dordrecht: D. Reidel Publishing Company.
- Searle, John R., Daniel Vanderveken (1985), *Foundations of Illocutionary Logic*, Cambridge: Cambridge University Press.
- Searle, John R. (1983), *Intentionality – An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press.
- Searle, John R. (1969), *Speech Acts – An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.
- Vanderveken, Daniel (1990-91), *Meaning and Speech Acts*, Vol. I & II, Cambridge: Cambridge University Press.

## NOTES

- 1 It should be noted that it is not entirely clear that this is actually Austin's view. It is true that there are passages where Austin talks about indication of illocutionary force as non-descriptive (e.g. Austin 1962: 70-71), but these are found before he arrives at the final distinction between locutionary, illocutionary, and perlocutionary acts. By making this distinction Austin suggests that truth conditions and illocutionary force belongs to two distinct strata of the speech act, i.e. the locutionary act and the illocutionary act respectively. Davidson does not comment directly on this part of Austin's theory and it is, therefore, not obvious if Austin deserves Davidson's criticism.
- 2 Davidson is not being explicit about these matters, but it seems inevitable that his analysis will face considerable difficulties if we allow that sentences substituting *p* contain second order expressions, meta-language, embedded performatives etc.
- 3 The difference between asserting *p* and uttering *p* non-assertively might also be seen as an analogy to Searle's distinction between presentation and representation (Searle 1983: 23). Searle uses the distinction to explain intensional contexts as it allows him to say that a speaker who utters e.g. 'John believes that King Arthur slew Sir Lancelot' presents rather than represents the content of the embedded sentence. The difference is that the speaker by performing a presentation does not commit himself to the truth of the content even though the utterance has truth conditions, direction of fit, etc.
- 4 Even though interrogatives are marked by a change of word order rather than by inflection Lewis treats them similar to the non-indicatives.