Truth, Meaning and Decidability

Neil Tennant


Stable URL:
http://links.jstor.org/sici?sici=0026-4423%28197707%292%3A86%3A343%3C368%3ATMAD%3E2.0.CO%3B2-E

_Mind_ is currently published by Oxford University Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/oup.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.
Truth, Meaning and Decidability*

NEIL TENNANT

1. Compositional Theories of Meaning

Given the meanings of separate words, and the rules of syntax, the meaning of a proposition is determinate. This is the reason why we can understand a sentence we never heard before. You probably never heard before the proposition ‘that the inhabitants of the Andaman Islands habitually eat stewed hippopotamus for dinner’, but there is no difficulty in understanding the proposition. The question of the relation between the meaning of a sentence and the meanings of separate words is difficult ...

Russell, *The Analysis of Mind*

Russell’s own theory of judgment had failed to overcome just this difficulty:

... (the theory) does not show how it is that we should be able to understand the statement abbreviated as ‘James judges that p’ provided that we understand both the expressions ‘James judges (that)...’ and the sentence (whose oratio obliqua form is) abbreviated to ‘p’.

Geach, *Mental Acts*

More recently, Davidson has underlined the importance of this difficulty in a series of papers. The so-called ‘Davidsonian requirement’ on a theory of meaning for a natural language now enjoys wide currency; it may be stated as follows:

The theory must show how out of the meanings of parts the meanings of wholes are composed.

Davidson describes theories which meet this requirement as recursive. Since the latter term already enjoys technical senses to which I shall appeal in due course, I propose instead to talk of compositional theories of meaning.

What is it for a theory of meaning to be compositional? For

* I am indebted to Donald Davidson, Timothy Smiley, and Hans Kamp for helpful discussion. Earlier versions of this paper were read to the Cambridge Philosophy Research Seminar and the St. Andrews Philosophical Club.
this to be the case there must be, for each reading of any sentence of the language, a constituent structure which brings out clearly the configuration and relative scopes of expressions within the sentence on that reading. This constituent structure must be seen as perhaps partly manifest, perhaps partly underlying, but in some sense associated with or enjoyed by the sentence on the given reading. Thus with Chomsky's example:

They heard the shooting of the hunters

there may be associated one constituent structure which brings out clearly that the hunters were shooting, and another which brings out clearly that they were being shot. When intuiting as native speakers of English what the appropriate reading of the sentence is, we consider context. When conjecturing as theorists about English that the sentence has been taken on a particular reading, we postulate the associated constituent structure which, according to our theory, 'brings out' that reading. Part of the task of semantics, then, is to associate constituent structures with readings in a systematic way. In so doing we can then appeal to the constituent structures as theoretical entities upon which meanings are ravelled.

Where, in this account, lies the compositional aspect of the theory of meaning? The answer to this question is usually roughly as follows:

The constituent structure is seen to arise by a finite number of applications of operations to form new expressions out of old ones. Such operations are applied initially to expressions drawn from a vocabulary of primitives. The stock of operations (for the whole language) and the vocabulary are both finite. The meaning of any expression depends only on its constituent structure and the meanings of those primitive expressions which are terminal with respect to that structure. There are finitely many compoundings of meanings, beginning with the meanings of the primitive expressions, compounding them in accordance with the meanings of dominating expressions, to yield after finitely many steps the meaning of the whole expression. Despite the constraint of finitude the means of construction suffice for the presence of infinitely many sentences of the language; because of the constraint of finitude, however, a speaker with 'finite resources' is able to produce and understand any even hitherto unencountered sentence of the language.
This is, I think, a fair summary of the picture which emerges both from informal accounts of Frege's theory of sense and from some linguists' accounts of the workings of the semantical component of a transformational grammar. It is intuitively satisfying, because comprehensible; but it hardly constitutes more than a re-description of the problem to be solved by an adequate theory of meaning.

Davidson expresses this sentiment clearly in his paper 'Truth and Meaning'. He sees no advantage in talking of meanings and their compoundings. Instead, he suggests, the theory of meaning should be furnished holistically:

If sentences depend for their meaning on their structure, and we understand the meaning of each item in the structure only as an abstraction from the totality of sentences in which it features, then we can give the meaning of any sentence (or word) only by giving the meaning of every sentence (and word) in the language.

'Compositional' theory seems to give immediate expression to the intuition that we understand the meaning of a sentence by understanding the meanings of its parts and grasping the way these are put together to get the meaning of the whole sentence, but in at least two passages in 'Truth and Meaning' Davidson seems unhappy with this:

recursive syntax with dictionary added is not necessarily recursive semantics

and earlier:

knowledge of the structural characteristics that make for meaningfulness in a sentence, plus knowledge of the meanings of the ultimate parts, does not add up to knowledge of what a sentence means.

He does not characterize a theory of meaning independently, as in the sketch above, and proceed from there to establish that a theory of meaning 'is' a theory of truth. Rather, he asks what requirements we would place on a theory of meaning, and argues that they are precisely those which we would place on a theory of truth:

What (is demanded) is a theory that has as consequences all sentences of the form 's means m' where 's' is replaced by a structural description of a sentence and 'm' is replaced by a singular term that refers to the meaning of that sentence,
from which he concludes that

What we require of a theory of meaning for a language L is that without any (further) semantical notions it places enough restrictions on the predicate 'is T' to entail all sentences got from the schema ('s is T iff p') when 's' is replaced by a structural description of a sentence and 'p' by that sentence.

This is the basis of his bold suggestion that a compositional theory of meaning is best furnished by a recursive theory of truth.

2. The Notion of Recursiveness in Truth Theory

What exactly is meant by recursiveness in the case of truth theory? It certainly does not mean that the truth theory is decidable. A set is decidable when there is an effective test for deciding, given any object, whether it is a member of that set. For example, the truth-table test is an effective test for membership in the set of propositional tautologies; but we know, by Church's theorem, that there is no effective test for membership in the set of valid formulae of first-order logic. Now to say that a truth theory is recursive does not mean that there is an effective method which enables use to decide, given any sentence of the language in which the truth theory is framed, whether that sentence is a theorem of the truth theory. As far as I know, no-one has ever considered or solved the problem of whether extant truth theories are recursive in this sense (the answer is probably negative).

Recursiveness of truth theory is a quite different matter from decidability. The recursiveness lies elsewhere—in the form of definition of an important predicate of the theory, namely the satisfaction predicate. To say that a truth theory is recursive is to say no more than that the satisfaction predicate is defined recursively; or, as is sometimes also said, inductively.

We are familiar, for example, with the clause-by-clause form of Tarski's definition of satisfaction for first-order extensional languages. There is a clause for each connective, for each quantifier and for atomic predicates. Since the operators distinguished with their own clauses in the definition of satisfaction in an obvious sense 'are' the means of constructing new formulae out of old ones, we can see that, given any formula, a finite number of applications of the satisfaction definition suffices to unravel the satisfaction conditions of the whole formula.
There is a formal criterion for the recursiveness of the definition of a predicate (due to Montague) but it is a laborious business applying it (just as it is a laborious business giving a strict logical proof of any theorem of arithmetic). In the cases which concern us we may just as well inspect the definition to make sure that it is of the right form. This is the case with Tarski’s theory; with Kripke’s theory for the language of modality; with Kaplan and Montague’s theory for higher-order languages; and with the author’s own theory for restricted quantification. In each of these cases we may say (using standard technical terminology) that there is, in the case of any formula of the language under discussion, a recursion over the subformula relation, which is well-founded. By saying that the subformula relation is well-founded we mean that the process of unravelling any formula into its constituent subformulae, and repeating the operation on the latter, must come to an end after finitely many steps with formulae which themselves have no subformulae. Because each formula can be unravelled in only one way, any function defined by recursion over the subformula relation is well defined. In a model-theoretic version of the truth definition, the function in question is that which assigns to each formula X and interpretation M the truth-value of X under M.

It is important to realize that the recursiveness of a theory of truth in no way guarantees that we can effectively determine, concerning any sentence and any interpretation, whether that sentence is true under that interpretation. Nor, to repeat, does it mean that there is an effective test for theoremhood in that theory.

3. The Metalanguage: Three Alternatives

In his original paper Tarski gave a broad characterization of the metalanguage. It had to contain structural descriptions and trans-

---

3 Not least because infinite interpretations can’t be assigned Godel numbers; and the existence of such a method in the case of ‘pure’ truth theory renders science superfluous.
lations of expressions in the object-language; and it had to obey 'the usual laws of logic'. For his treatment of the language of the calculus of classes, he could have taken the first-order language of set theory as his metalanguage. In his metatheory he would need such set-theoretic axioms as would guarantee the existence of sequences and justify the method of definition by recursion; in short, quite a powerful formal apparatus. But in developing a truth theory one implicitly assumes what is required. Moreover, for readability one avoids rigorous symbolization in the metalanguage and uses instead so-called 'logician's English'.

When we consider truth theory for rich fragments of natural language, however, whether or not the metalanguage is formalized becomes important. There are three alternatives, to each of which I shall devote some discussion:

(I) The first is to take our metalanguage to be a fragment of a natural language—say English—embellished with the requisite set theory. Tarski's material adequacy condition would then be as follows:

Each instance of the schema 'X is true iff p' must be provable, where 'X' is replaced by the structural description of a constituent structure underlying a sentence of the object-language on one of its readings, and 'p' by a translation of that sentence on that reading into English.

Now p will be a translation of an object-sentence on one of the latter's readings only if p is unambiguous. For example, if X and Y are descriptions of the constituent structures underlying the English sentence

Polishing machines can be dangerous

on its two readings, we would have an instance of the adequacy schema both

X is true iff machines which polish things can be dangerous and

Y is true iff it can be dangerous to polish machines.

The problem would be to settle on a workable system of proof and definition in unformalized English which will allow the derivation of such instances of the schema. This problem is not to be underestimated. In a more difficult example, with many embeddings of quantifiers and verbs of propositional attitude, a
fully unambiguous circumlocution on the right hand side would be extremely complicated. We might manage a short one like
\( \forall x (Jx, \forall y(My, yBxKLyx)) \) is true iff John and Mary are such that she believes him to know that she loves him;

But as complexity increases the translation on the right hand side becomes more and more unwieldy, until our understanding of 'the former . . . the latter . . . such that' constructions is strained to the point of collapse. At this point we would be well-advised to embark upon some formalization of the metalanguage itself, so as to avail ourselves of the useful quantifier-variable notation in making the right hand side more perspicuous. This indeed is the second alternative:

(II) Take the metalanguage to have the structural and symbolic resources of the object-language, in addition to whatever set theory is necessary. For example, with restricted plurality quantifications, we might assume the metalanguage to have the quantificational resources of the object-language. Then, with the constituent structure

\[ N(Dx)\times M(By \& Rxy)ySy \]

representing the quantificational form of the English sentence

Nearly all the delegates had read mostly soft-covered books; and with the new clause

\[ s \text{ satisfies } Q(Fx)xGx \text{ iff } Q(s(x/a)) \text{ satisfies } Fx)a \text{ s(x/a) satisfies } Gx \]

in the definition of satisfaction (where underlining indicates translation into the metalanguage), we would have the following (informal) proof of the T-sentence

\[ N(Dx)\times M(By \& Rxy)ySy \text{ is true iff } \_N(Da)aM(Bb \& Rab)bSb: \]

\[ N(Dx)\times M(By \& Rxy)ySy \text{ is true } \]

\[ \_\exists s \text{ satisfies } N(Dx)\times M(By \& Rxy)ySy \]

\[ \_\exists N(s(a/x)) \text{ satisfies } Dx)a \text{ s(a/x) satisfies } M(By \& Rxy)ySy \]

\[ \_\exists N(D(s(a/x)(x)))a \_M(s(a/x)(b/y)) \text{ satisfies } Bx \& Rxy)b \text{ s(a/x)(b/y)} \]
satisfies \text{Sy}
\text{iff}
\exists s \, N(Da)a \, M(s(a/x)(b/y)) \text{ satisfies By} \land s(a/x)(b/y) \text{ satisfies Rxy} \land S(s(a/x)(b/y)(y))
\text{iff}
\exists s \, N(Da)a \, M(Bs(a/x)(b/y)(y)) \land R(s(a/x)(b/y)(x), s(a/x)(b/y)(y)) \land b \land Sb
\text{iff}
\exists s \, N(Da)a \, M(Bb \land Rab) \land b \land Sb
\text{iff}
N(Da)a \, M(Bb \land Rab) \land Sb.

Note that the translation is verbatim. The success of any truth theory providing such instances of the adequacy schema depends on our thorough preliminary understanding of the logical behaviour of restricted quantifiers and other operators in the metalanguage.\footnote{Note that the informal proof of the T-sentence in the case of restricted plurality quantification may be more difficult to transform into a strictly formal proof than in the classical case. In the classical case we depend heavily on the substitutivity of logical equivalents, and the derivability of that rule in a system of, say, natural deduction. I do not see how the same rule is to be derived in plurality logic; it may have to be taken as primitive. Moreover, the most interesting plurality logic, even for un-restricted quantification, is incomplete—see A. Slomson, ‘Remark on Altham’s “Logic of Plurality”’, \textit{Journal of Symbolic Logic}, 39 (1974), p. 418 (Abstract).} We may illustrate this point with another example: we may prove
\[ \tau x(Fx, Gx) \text{ is true iff } \tau a(Fa, Ga); \]
But unless we know exactly what we mean by ‘\( \tau \)’ in the metalanguage, we shall not be able to claim that we have characterized the truth conditions of \( \tau x(Fx, Gx) \). Whether or not \( \tau x(Fx, Gx) \) is true if and only if there is just one \( F \) and it \( G \)’s depends on whether or not we intend that reading of ‘\( \tau \)’ as used in the metalanguage.

It may be preferred not to rely on this purportedly trouble-free understanding of operators in the metalanguage, and to try for translations on the right hand side which spell out in greater detail the truth conditions intended. Thus we might state the satisfaction clause for the description operator \( \tau \) as follows:
\[ s \text{ satisfies } \tau x(Fx, Gx) \text{ iff } \text{card } \{ a | s(a/x) \text{ satisfies } Fx \} = \text{card } \{ a | s(a/x) \text{ satisfies } Fx \land Gx \} = 1. \]
In general, we might eschew ‘translations’ of quantifiers into the metalanguage and introduce in their place mappings of cardinal numbers to truth values, along lines suggested by Mostowsiki. Thus for each quantifier $Q$ of the object-language we assume a corresponding mapping $Q^*$ of ordered pairs of cardinals to the truth values $t$ and $f$. Then the satisfaction condition for restricted quantification would be as follows:

$$s \text{satisfies } Q(Fx) \forall x \exists x$$

iff

$$Q^*(\text{card } \{a|s(a/x) \text{satisfies } Fx & Gx\}, \text{card } \{b|s(b/x) \text{satisfies } Fx & \sim Gx\}) = t$$

The effect of these alterations is to keep the metalanguage classical, insofar as the existential and universal quantifiers are the only quantifiers used. But the preservation of classical reasoning in the metatheory is offset by the loss of perspicuous translations on the right hand side of Tarski’s schema. According to this more mathematical specification of truth conditions, our earlier example would become

$$N(Dx) \forall x M(By & Rxy) \exists y s y \text{ is true}$$

iff

$$N^*(\text{card } \{a|D a & M^*(\text{card } \{b|Bb & Rab & Sb\}, \text{card } \{c|Bc & Rac & \sim Sc\}) = t\}, \text{card } \{d|Dd & M^*(\text{card } \{b|Bb & Rab & Sb\}, \text{card } \{c|Bc & Rac & \sim Sc\}) = f\}) = t$$

It is only insofar as the complicated mathematical expression here can be regarded as a ‘translation’ of the object-sentence into the metalanguage that we may accord the biconditional any importance as an instance of the adequacy schema. The value of these ‘cardinality’ interpretations is that they permit precise formulation and solution of problems concerning axiomatizability and definability in languages containing the corresponding quantifiers; while the underlying logic of the metalanguage remains classical.

(III) Another departure from ‘pure’ truth theory—with which the original adequacy schema is properly concerned—is the introduction of explicit reference to models or interpretations of the object-language. In the case of classical first-order languages, an interpretation consists of a domain of discrete individuals standing in determinate relations to one another. Each $n$-place predicate

---

1 See A. Mostowski, ‘On a generalization of quantifiers’, *Fundamenta Mathematica*, 44 (1957). For more details on the adaptation of Mostowski’s method to restricted quantification see J. Altham and N. Tennant, *ibid.*
is assigned as extension a set of n-tuples of members of the domain. If M(R) is the extension of a two-place predicate R in M, and if D(M) is the domain of M, we would have as an instance of the adequacy schema
\[ \forall x \exists y Rxy \text{ is true in } M \text{ iff } \forall a \in D(M) \exists b \in D(M) \langle a, b \rangle \in M(R). \]
Explicit mention of M as a parameter in the truth predicate alters little of substance in the truth definition. Translation becomes relativized to interpretation in an innocuous way. Moreover, it is now possible to define logical consequence in terms of truth:
\[ \Delta \text{ logically implies } X \text{ iff for every } M, \text{ if every member of } \Delta \text{ is true in } M \text{ then } X \text{ is also.} \]
It is precisely in order to give analogous definitions of logical consequence in richer languages that model-theoretic versions of the truth definition are favoured. (Kripke, for example, defines models with a richer internal structure in the case of modal logic, and defines truth and logical consequence in a way parallel to that indicated above.) In giving these versions of the truth definition logicians do not hesitate to use 'mathematical' interpretations of expressions of the object-language in preference to what might ordinarily be regarded as their 'translations'. They do this in order that logical consequence be a mathematically well defined relation. But then the putative instances of the adequacy schema are not all simple, as has been pointed out above. We have to rely on our grasp of the adequacy of each clause in the truth definition, and of the way these clauses act in concert to give an overall representation of the truth conditions of any sentence, to see that the definition as a whole is materially adequate. So according to the third and final alternative, the adequacy condition will be understood as follows:

Each instance of 'X is true in M iff p_M' must be provable, where 'X' is replaced by a structural description of a constituent structure, and 'p_M' by a mathematical translation of what it says under the interpretation M.


Earlier we clarified what it is for a theory of truth to be recursive. Now, in the light of the foregoing discussion of three different methods of framing one's truth theory, we can consider some of
Davidson's claims on behalf of truth theory. In this connection I must draw attention to a very strong requirement which he places on a theory of meaning. He says in 'Truth and Meaning'

What (is demanded) is a theory... that provides an effective method for arriving at the meaning of an arbitrary sentence structurally described

and

In this paper I have assumed that the speakers of a language can effectively determine the meaning or meanings of an arbitrary expression (if it has a meaning), and that it is the central task of a theory of meaning to show how this is possible. (My italics)

If by this Davidson means that the theory of meaning must provide effective methods which model the effective determination of meanings by speakers, then his assumptions have the status of philosophical theses of some importance.

The first assumption—that speakers effectively determine meanings—is a philosophical thesis about linguistic competence and mind going even further than the theses of innatism which emanated from Chomskyan linguistic theory.¹ A Chomskyan grammar provides only an enumeration of pairings of 'deep' structural descriptions with 'surface' sentences of the language concerned. No mention is made of effective methods for determining the deep structure underlying a surface sentence on any of the latter's readings.

The second assumption—that a theory of meaning should show how it is that speakers can effectively determine meanings—is a philosophical thesis about the power of theory and the limits of natural language itself. I do not think that Davidson has chosen the word 'show' deliberately in order to be on the safe side of a Tractarian dichotomy between showing and saying. If he were being cautious and this way he would be entitled to stop short at simply laying bare constituent structures. Instead he wants to say that this operator affects truth conditions this way, and so on, for each operator of the language concerned.

What could possibly be the effective method, within Tarskian truth theory, which could be taken as modelling speakers' effective determination of meaning? Let us consider an inadequate answer

¹ See the contributions of N. Chomsky and H. Putnam to a symposium on innate ideas, in Synthese, 17 (1967).
TRUTH, MEANING AND DECIDABILITY

first: there should be an effective method which, given any sentence structurally described, will simply determine a sentence of the metalanguage such that (without necessarily supplying the relevant proof) the biconditional ‘X is true iff p’ is a theorem of the truth theory (where ‘X’ is replaced by the structural description and ‘p’ by the sentence determined by the method). But in the case of homophonic truth theory (where the object-language is contained in the metalanguage) there is, trivially, an effective method which meets this requirement—namely, the mapping from structural description to sentence described. The effective method we are seeking must therefore be one whose effectiveness or methodicalness cuts deeper.

Let us therefore try a second answer, which seems to be the only plausible candidate in this setting: there should be an effective method for getting from any structural description to a proof of the appropriate biconditional. The existence of such a method is prima facie compatible with the undecidability of the truth theory as a whole, since the set of instances of the T-schema is only a fragment of that theory.¹ The proposal under consideration, then, is that this method could be regarded as the ‘effective method for arriving at the meaning of an arbitrary sentence structurally described’.

In the case of classical first-order languages an effective method does exist. An informal argument to this effect is as follows:²

Every T-sentence is provable. There is an effective method to determine, given any sentence of the meta-language, whether it is a T-sentence (that is, of the form ‘X is true iff p’). So, given a structural description X, apply the following method:

Enumerate proofs in the truth theory. Check each to see whether it is a proof of a T-sentence involving X on the left hand side. After finitely many steps we will come across a proof of such a T-sentence, since every T-sentence is provable.

The proof of the fact that every T-sentence is provable is by induction on the length of formulae. (In his paper ‘The Concept of

¹ This may be clearer with an arithmetic parallel: arithmetic with addition and multiplication is undecidable; while the proper fragment of arithmetic with addition but not multiplication is decidable (Godel, Presburger respectively).

² This was communicated to me by Davidson, who attributes it to Scott Weinstein.
Truth in Formalized Languages' Tarski confined himself to the remark that

The proof would require the setting up of an entirely new apparatus: in fact it involves the transition . . . to the metametatheory, which would have to be preceded by the formalization of the metatheory . . . ;

he contented himself with giving derivations of sample T-sentences to make plausible the claim that every T-sentence is provable.) Notice, however, that the argument above is indirect; it establishes only the existence of some effective method for getting from structural descriptions to proofs of T-sentences. It does not guarantee that the 'p' in a T-sentence obtained in this way is the most economical or direct translation of the sentence structurally described. Even if this were remedied\(^1\) it would still be the case that the proof given above does not yield a method bearing any intimate relationship to the process of unravelling a sentence's meaning according to its structure. Finding a proof has been divorced from the process of proving. It is therefore doubtful whether we have succeeded in finding a satisfactory model of the process of effectively determining meanings.

It should be plausible that we can improve upon this result. There is indeed an effective method yielding proofs of T-sentences from structural descriptions, which closely parallels the structure of the sentence concerned. It is best illustrated with one of Tarski's own examples:

\[
\forall x \exists y Rxy \text{ is true} \\
\text{iff} \\
\exists s \text{ s satisfies } \forall x \exists y Rxy \\
\text{iff} \\
\exists a \forall x (a/x) \text{ satisfies } \exists y Rxy \\
\text{iff} \\
\exists a \forall b (a/x)(b/y) \text{ satisfies } Rxy \\
\text{iff} \\
\exists a \forall b R(s(a/x)(b/y)(x), s(a/x)(b/y)(y)) \\
\text{iff}
\]

\(^1\) By, for example, being equipped with an effective mapping \(t\) which when applied to structural descriptions yields verbatim translations (should they be available), and then searching for a proof of 'X is true iff \(t(X)\').

I owe this point to Hans Kamp.
\[ \exists s \forall a \exists b \, R(a, b) \]

iff

\[ \forall a \exists b \, R(a, b). \]

This is an (informal) proof of the T-sentence

\[ \forall x \exists y \, Rxy \text{ is true iff } \forall a \exists b \, R(a, b). \]

This proof, which is readily transformable into a strictly formal one in a system (say) of natural deduction, displays those steps which are involved in the proof of any T-sentence, and which were involved also in our earlier example concerning restricted plurality quantifications. These steps are:

(i) straightforward application of the definition of truth
(ii) straightforward applications of the appropriate clauses in the definition of satisfaction, where the order of application was straightforwardly determinable from the structure of the formula;
(iii) straightforward simplification of terms of the form \( s(a_1/x_1) \ldots (a_n/x_n)(x_i) \); and finally
(iv) discarding the irrelevant initial existential quantifier.

Each step in the proof is mechanical (that is, there is no need for ingenuity in deciding how to proceed after any stage of the proof); so it would not be difficult to formulate an algorithm for obtaining a similar proof in the case of any T-sentence structurally described. Moreover, if the metalanguage contains the structural resources of the object-language, the translation \( p \) in the proven instance of the schema '\( X \) is true iff \( p' \) will be a direct or verbatim translation of the sentence \( X \).

Thus, so far as quantification is concerned, we do seem to have an effective method for getting from any structural description to a proof of the appropriate biconditional; a method, moreover, which because of its meticulous unravelling of sentential structure may perhaps be put forward as a model of the allegedly effective process by which speakers determine meanings.

But to the best of my knowledge there is no general argument which guarantees that such a method exists in the case of truth theories for more richly expressive fragments of natural language.\(^1\)

\(^1\) On the contrary, if with Davidson we rule out model theoretic versions of the truth definition, the likelihood of this is decreased. See his paper 'In Defense of Convention T' in H. Leblanc (ed.): Truth, Syntax and Modality (North-Holland, 1973).
Just as the effective test for validity disappears when we pass from propositional to predicate logic, so also it may happen that the effective method for finding proofs of T-sentences is provincial to extant truth theories, and will not be available for truth theories concerning languages of greater expressive power. For such languages it may be the case that, on the basis of our understanding of the structure of a sentence and of the way the primitive expressions translate into the metalanguage, we make an ingenious choice of a sentence in the metalanguage; and then display our ingenuity once more by establishing as a theorem the T-sentence with that sentence on its right hand side. And what, now, would have become of the theoretical account of speakers' effective determination of meaning?

5. The Relationship Between Ordinary Sentences and Their Canonical Representations

Although Davidson's concern in 'Truth and Meaning' was to sketch the overall form of a theory in such a way as to give a strong indication of its promise and adequacy, there remains an important part of the theory of whose form he gives no clear account. This is the part which deals with the correlation between sentences of the natural language and sentences of the canonical language which do duty for them when Tarskian methods are brought to bear. Davidson does not want to be accused of improving or reforming a language, but wants to describe and understand it. '... the point', he says, 'is not that canonical notation is better than the rough original idiom, but rather that if we know what idiom the canonical notation is canonical for, we have as good a theory for that idiom as for its kept companion'.

Again, this passes over the problem which Davidson has explicitly set himself: the problem of accounting for speakers' effective determination of meaning. The theory to which he aspires should be of such a form that, given any sentence of the language, there is an effective assignment of a canonical sentence to it on each of its readings; and, as already discussed, there should be an effective method of culling from the canonical the meaning which the theory is to assign to the original. It is one thing to observe that

Philosophers have long been at the hard work of applying theory to ordinary language by the device of matching
sentences in the vernacular with sentences for which they have a theory

but quite another to indicate how the theory could be extended (in a manner which Davidson's general position obliges him to make clear) to bring the 'device of matching' within theoretical purview.

Davidson alludes to the work of Chomsky; but fails to grasp the full import of the problem whose solution might well avail itself of some adaptation of Chomskyan method. He remarks that getting the transformational relationships right between active and passive sentences would allow us to extend to passive sentences the treatment which succeeds in giving the truth-conditions of the active sentences. He does not see, however, that the transformational relationships can be exploited to this end, and the account achieve overall the explanatoriness he requires, only if we have a *decidable* grammar which will have told us, effectively, what the deep structures of the active sentences were in the first place.

In outline, Davidson hopes to give the truth-conditions of any sentence S of the natural object-language in two steps:

\[
\begin{align*}
(1) \quad \text{canonical sentence } X \text{ (or constituent structure, deep structure, logical formula) structurally described} \\
(2) \quad \text{proof of 'X is true iff p', where p is a translation of S on the reading in question into the metalanguage}
\end{align*}
\]

(1) and (2) must be effective steps. Nothing Davidson has said guarantees that (2) will be so in the case of more richly expressive fragments of natural language; and the effectiveness of (1) is almost beyond the dreams of contemporary linguistics.

It may seem that what I have said about Davidson's account of a theory of truth and its importance for a theory of meaning leads to a pretty gloomy prognosis for future work along these lines. If so, it matters little; for hardly any theory of significance has
been developed beyond the 'massive contribution' of Frege and Tarski. Davidson has written at considerable length only on the overall form of a theory. His work on adverbial modification admittedly brings a contribution to theory-building itself, but does so in a way which makes all the more acute the need for a systematic account of the relationship between the vernacular and the canonical—because quantification over events certainly does not advertise itself in the surface forms of sentences with adverbial constructions.

6. The Status of Davidson's Thesis

Davidson's basic claim, considerations of effectiveness aside, was that

a theory of meaning for a language L shows 'how the meanings of sentences depends upon the meanings of words' if it contains a (recursive) definition of truth-in-L.

We can restate this claim as follows:

A theory of meaning is compositional if it contains a recursive definition of truth.

The closing passages of 'Truth and Meaning' show that Davidson holds the converse also, as a working assumption. Let us therefore call the claim above, with 'if' strengthened to 'if and only if', Davidson's Thesis. What status may we now accord this thesis?

Since compositionality is an informal notion Davidson is deprived of any strict or formal argument for his thesis. A great deal must therefore turn on:

(i) the possibility of an effective method for arriving at proofs of Tarskian biconditionals (i.e. steps (1) and (2) above); and
(ii) the fact that the recursion in the definition of satisfaction is based upon the same features of sentential structure as is the intuitive compounding of meanings.

---

1 T. Burge has made some progress with proper names, S. Weinstein with demonstratives. See their papers in Nous (1974).

2 This is the case with all model theoretic accounts of richer fragments. See N. Tennant, Recursive Semantics for Knowledge and Belief (Ph.D. Dissertation deposited at Cambridge University Library); E. B. Richards, 'A Point of Reference', Synthese, 28 (1974), and 'A Theory of Adverbs' (forthcoming).
Even if the effective method of (1) does exist it would still be impossible strictly to establish that that shows how speakers effectively determine meanings. In support of this claim I advance the following analogy.

Consider the problem of accounting for a child’s effective determination of solutions of arithmetical computations. Suppose we have a Turing machine which will simplify any correctly bracketed numerical term built up from numerals by the function symbols for addition, multiplication, subtraction and division. Now an intelligent child would give two answers to the verbal command ‘Two plus three times five’: ‘seventeen’ and ‘twenty-five’. Suppose we had an effective method (1) for assigning numerical terms to verbal commands, fed the former into our Turing machine (2) and obtained printouts:

\[2 + (3 \cdot 5) \rightarrow 17\]

\[2 + 3 \cdot 5 \rightarrow 25\]

Two plus three times five

Consider the structural analogy here with the process of assigning constituent structures X and Y to the English sentence ‘Polishing machines can be dangerous’, and then employing the effective method discussed earlier to get proofs of appropriate T-sentences:

\[X \rightarrow \text{proof of ‘X is true iff machines which polish things can be dangerous’}\]

\[Y \rightarrow \text{proof of ‘Y is true iff it can be dangerous to polish machines’}\]

It is quite possible that in the numerical example nothing in our
effective two-step process simulates or models the child’s computational activity (mental or pencilled), but only achieves the same end result as it. Analogously, it would appear to me, it would be extremely difficult to place enough constraints on the steps (1) and (2) in the semantical example to ensure that (1) simulates the process of parsing a sentence and that (2) simulates the process of determining its meaning (assuming that these are different processes).

It would appear, then, that Davidson’s Thesis should be accorded the same sort of status as Church’s Thesis. For, no matter how successful analyses in terms of constituent structure may be in vindicating the intuitive view that we grasp meanings of parts and compound them appropriately to obtain meanings of wholes, there remains the problem of the determination—or, better, the genesis—of the meanings of primitive expressions. It seems that there is no way to account for an expression’s having a certain meaning other than by appealing to the truth conditions of whole sentences in which it occurs. For the order of explanation, as Dummett has put it, the senses of sentences are primary. To this extent Davidson’s holism would be justified, and to this extent the intuitively appealing compositional theories of meaning remain as embellished descriptions of our original problem. They are to be replaced by precisely formulated theories of truth, which are not demonstrably better, but are felt to be better explanations of phenomena for which we require explanation.

Compositionality is an informal notion; the recursiveness of a truth definition is a formal notion. Analogously, computability of functions is an informal notion while general recursiveness is a formal notion. Church’s Thesis is that the latter should be regarded as the precise explication of the former. Davidson’s Thesis, analogously, is that recursiveness of the truth definition is a precise explication of compositionality of meaning. A justification of his thesis would run as follows:

Any language whose speakers have a compositional grasp of meaning admits of a recursive truth definition; and any language which admits of a recursive truth definition is such that its speakers have a compositional grasp of meaning.\(^{2}\)

---

1 In Frege (Duckworth, 1973), p. 4.
2 One point at which our analogy does not fare too well is that while Church’s thesis gains plausibility from the equivalence of several different formalizations of the notion of computable function, there is no support from a similar direction for Davidson’s thesis.
This justification is reinforced by (ii) above: the recursion in the truth definition is based upon the same features of sentential structure as is the intuitive compounding of meanings—at least for those languages for which truth definitions have so far been given.

The absence from truth theory of the effective methods discussed above does not render any the less important the provision of a truth theory for any language undergoing semantical investigation. Those ingredients of meaning which cannot be characterised in terms of truth conditions are best relegated to intentional theories of illocutionary force, of conversational implicature and so forth. The most important part of the sense of a sentence (that ingredient of its meaning to be captured by a characterization of truth conditions) remains its logical force. Since we grasp sense without being logically omniscient we can conclude that the most important part of our grasp of sense depends on an awareness of some of the most important sentences which would either entail or be entailed by it. To that extent at least a theory of meaning must ‘read a logic’ into the language; and the path to logical consequence is, as we know, properly via the notions of truth and interpretation. Even incomplete systems of natural deduction may then find application in showing how the most important entailment relations into which a sentence enters may be effectively determinable. In this way, perhaps, we may come closer than we were before to accounting for effective determination of meaning (sense), if such a process exists.

In the event of a general failure of effectiveness in truth theory we may have to challenge the original assumption that speakers effectively determine meaning in any sense which could be modelled (in a suitably strict sense) by semantical theories which provide effective methods and procedures. We could instead regard the circumstance as yet another indication of human creativity and ingenuity, shown in our ability to grasp the meanings of sequences of sounds. We may succeed all the time in working them out; we may be very effective in doing so; but that would not be to say that those meanings are effectively determinable in any way which can be simulated by effective methods in semantical theory. There may be processes underlying our understanding of natural language which are both highly effective and highly non-effective.