I. The Role of Syntax in a Theory of Meaning
A. Quine speaks of two projects for a theory of meaning. The first is to say when a phonetic or orthographic string is meaningful. This project is analogous to the construction of a syntax for a formal language, that is, a recursive definition of wff such as this.
   B. Sample definition of wff for predicate calculus.
      a. F, G, H, ... are predicate letters;   a, b, c, ... are names; x, y, z are variables.
      b. A term list is any string each item of which is a name or a variable.
      b. When P is a predicate letter and t is a term list, then Pt is a wff.
      c. If A and B are wffs then so are ~A, (A&B), (AvB), (A→B).
      d. If v is a variable, and A is a wff, then ∀vA is a wff.
      e. Nothing else is a wff.
   C. Why is this definition important to the creation of a formal semantics? It provides a **template** for the recursive definition of semantical value. We know that we will need a truth clause for each way b-d that a wff can be formed.

II. Davidson's Theory of Meaning
A. The Problem in Providing a Theory of Meaning.
   1. One might propose to define meaning of any sentence via the **referents** of the expressions that appear in it.
   2. The **Slingshot** is a widely cited argument that (in essence) goes back as far as Frege. It purports to show that whatever we take the referent of a sentence to be, true sentences must all have the same referent. So if meaning is to be defined by the way in which a sentence refers to something (whether it be truth values meanings or whatever), all sentences true sentences will mean the same thing.
   3. A formal account of this is given in Davidson's "Truth and Meaning" fifth paragraph. The notation: x^(x=x.R) = x^(x=x) reads: the set of things identical to themselves such that R holds is identical to the set of all things identical to themselves. Davidson claims that R is provably equivalent to x^(x=x.R)=x^(x=x), (and similarly for S). To see why, suppose R and S are true. Then the set of all things identical to themselves such that R holds, just amounts to everything, that is, the set of all things identical to themselves, and similarly for S. Therefore x^(x=x.R)=x^(x=x) is also true. Now suppose R is false. Then nothing qualifies as a member of x^(x=x.R), since the condition R is not met. But x^(x=x) on the other hand is the set containing everything, since everything is self identical. So when R is false, so is x^(x=x.R)=x^(x=x) since it claims x^(x=x.R) and x^(x=x) are identical, which they are not. The same argument works for S, and it also shows that x^(x=x.R) and x^(x=x.S) have the same referent. Now presuming that provably equivalent sentences have to have the same referent, and that expressions with the same referent may be substituted one for another, all these must have the same referent for arbitrary choices of sentences R and G having the same truth value: R, x^(x=x.R)=x^(x=x), x^(x=x.S)=x^(x=x), S. The upshot: a theory of meaning that appeals to reference must assign the same referent to sentences with the same truth-value. This is intolerable if reference is going to do the job to giving an account of meaning. (Note. There is an interesting literature on the cogency of the Slingshot. Not everyone buys it.)
   4. Could we do better bringing in meaning directly to a theory of meaning? Davidson says no. If we do so, we end up with a vacuous theory.
   B. **The Promise.**   By providing a formal semantics of a language, we may end up with all we need for a theory of meaning. Here we could avoid mentioning meanings directly and still avoid the slingshot by failing to treat expressions as having referents. As long as we can tell you the circumstance under which each sentence is true, we would have told you enough for you to know its meaning. So the project will be to say what a model of a language is, and to give a recursive account of the truth conditions for each expression in a systematic way following the formation patterns revealed in the

syntax of the language. There will be no need to ever bring in any conception of meaning, nor will there be a need to assign expressions a referent save in the sense that we tell you how to calculate whether a sentence is true given the same information about its parts.

    C. **The Project.**

        1. The idea is that if we can give a recursive account of the truth conditions for a language L as a theory in predicate logic, then this will allow us to prove sentences of the form:

    (T) S is true iff p

where p gives an account of the circumstances under which S is true. Then (T) would qualify as fixing the meaning of S.

        2. By 'theory' we mean a set of axioms to be added to the machinery of predicate logic. The theory of truth would allow us to deduce all true sentences of shape (T).

        3. Note this project does not explicitly trade in a concept of "meaning" nor does it assign referents to expressions. It merely provides information about how to understand S, in the course of proving sentences of shape (T).

    D. **Clarifications.**

        1. Why knowing the lexicon is not enough. It is not enough to know the meanings of the words recorded in the lexicon (dictionary) of L. This would be analogous to having the model part of a formal semantics. We also need to know how to calculate truth-values of complex expressions from the truth conditions of their parts.

        2. How do we determine "meanings" of terms and predicate letters? Since the theory is based on an account of 'is true for L', where do the contributions of terms as predicate letters come in? Well, in a way, Davidson treats them as syncategorematic, in the sense that they do not get assigned meanings directly, but their semantical contributions are determined holistically by how they affect the truth-values of all sentences of L in which they appear. Therefore the theory is, in this sense, holistic.

        3. How can Davidson overcome the issues raised by Quine? According to Quine the truth value of no atomic sentence can be determined alone. It is whole webs of belief that confront the evidence for their eventual acceptance or rejection. So how can Davidson make sense of truth conditions for atoms of L? Well in a way he doesn't. For atoms, one need not solve the **epistemological** problem of whether they are true or not. Instead, all we need is to know the circumstances that would guide the quest for a truth-value. So for atoms, we can simply assert such claims as:

    (TGG)   'Grass is green' is true in English iff grass is green

Presuming we get such claims on the table for the (finitely many) atomic sentence of L, the recursive truth conditions can do the rest of the work for complex sentences of L.

        4. Why is the theory not trivial? It is trivial in the extreme to provide a theory that generates all true claims of form (T). Just provide instructions for taking sentence S placing it once to the left of an 'iff' in quotes followed by 'is true' and once on the right unadorned to obtain: 'S' is true iff S. (Sample 'Grass is green' is true iff grass is green.) However, this does not provide the kind of theory Davidson is seeking. To qualify, the theory must generate all results of the kind (T) using a recursive definition of truth that shows how truth-values of larger expressions depend on those of their parts. This definition has to define truth conditions for each of the formation structures defined in the grammar of the language. The trivial theory just outlined does not do this.

III. **Problems with Davidson's Program**

    A. Is it possible to formalize a natural language?

        1. The success of Davidson's program is an "empirical" matter. It depends our ability to construct of a formal semantics for the language for which we hope to have a meaning theory. I have already mentioned that the project of defining the grammar for a natural language like English has stalled. Since this is a precondition for going on to

give a formal semantics, the prospects look bleak. Davidson admits that other smart people (eg. Tarski and Quine) are not optimistic, and at the end of "Truth and Meaning" he gives a laundry list of issues that must be dealt with, none of which I would count as solved in the 50 years since Davidson wrote: counterfactuals, subjunctive conditionals, probability, causality, adverbs, attributive adjectives ('good man' 'good book'), mass terms, expressions for belief, perception and intention, imperatives, interrogatives …

      2. A quick look at belief will help reveal some of the issues. Notoriously, the operator 'John believes that' has no truth table. It is not possible to tell from the truth-value of 'grass is green' what John will believe about it. So the project of giving truth conditions for this expression is foiled. Davidson mentions a possible solution: that 'John believes the Earth is round' is to be analyzed as 'John responds affirmatively to "the Earth is round"', treating it as an atomic sentence without the sentential operator 'John believes'. (I think Wittgenstein suggests this analysis in the <u>Tractatus</u>.) But Davidson himself shows that this is wrong. Variations on this strategy due to Davidson and others have not been widely accepted.

      3. There are also formal worries. We know from Tarski's work on paradoxes of truth that a truth predicate 'is true' cannot be defined for a language meeting some elementary conditions, notably that it have names for its own sentences (via some device such as quotation) Otherwise (presuming a standard logic) we can derive contradictions in the theory. Davidson's reply seems lame to me: Well, we can still work out meaning theories for fragments of our language suitable for conducting philosophical discussion (but not discussion about truth???).

B. Would a Truth Theory Provide a Theory of Meaning?

      1. Even if the project could succeed, there are serious objections to the thesis that it provides us with a theory of meaning. The issues are related to the fact that a theory providing truth conditions fails to yield information with fine enough grain to qualify as having fixed meanings in ways we would expect. The considerations recorded by Soames are intricate, but the following gets at the gist of it. Soames argues that knowledge of a correct truth theory for L of the kind described is neither necessary nor sufficient for understanding L. It is not necessary, since it is implausible to presume ahead of time that knowledge of the machinery of Tarski style theories is the only way to come to understand a language. It is not sufficient by the following reasoning.

      2.   Suppose we have a correct recursive truth theory for L that entails these two results:

        (TGG) 'Grass is green' is true iff Grass is green

        (Not) 'not A' is true iff A is not true

Then by the rules of propositional logic, we may obtain (TBAD).

        (TBAD) 'Grass is green' is true iff (Not) and Grass is green

 (The principle is that from ⊢ T ↔ G and ⊢ N it follows that ⊢ T ↔ (N&G).)
But (TBAD) is wrong about what 'Grass is green' means. The problem is that any truth theory based on standard logic is going to deliver many many sentences of the shape (T) that add superfluous items unrelated to the meaning our intuitions accept as correct. Note that all those unacceptable sentences of shape (T) give correct truth conditions for 'grass is green' ((TBAD) is in fact true), but only <u>one</u> gives the meaning. Therefore knowing the truth theory is not sufficient for knowing which of all the (T)-shaped sentences captures knowledge related to our understanding.

      3. Perhaps further constraints can be placed on the theory so that (TGG) is singled out from among rivals like (TBAD). Intuitively, we don't need the conjunct (Not) to know the meaning of 'Grass is green', but it is far from clear how to define such a constraint without illicitly helping ourselves to our intuitions about meaning. One suggestion along these lines by Larson and Segal p. 36 (see syllabus for the citation) is to weaken the logic use to deduce sentences like (TGG) and block additional inferences to sentences like (TBAD). One idea is to disallow substitution of any sentences that were not used in the deduction. Since (Not) presumably was not used in the deduction of

(TGGZ) then (TBAD) is not obtained. But this does not work. Suppose we are trying to get the T sentence for 'Grass is not green'. Then (Not) would be used in the deduction, and so could be substituted into the (T) sentence for 'Grass is not green' to get another bad result. So a lot more research is needed to find the right alternative logic.

IV. Where Do We Go From Here?
   A. Could we define a formal semantics that takes meanings seriously?
      1. Soames suggests that the Davidson Program is a failure. The diagnosis is that a theory of truth is too weak to get a grip on language understanding. He considers looking at possible worlds semantics as one solution.
      2. Here one might identify propositions (meanings of sentences) with sets of possible worlds where they are true. Alternatively, a proposition expressed by sentence S could be identified with a function that assigns S a truth-value for each of the possible worlds. Something like this innovation seems necessary in any case to handle demonstratives and tense. The sentence 'I am (now) happy' is doubly indexical. To obtain its truth-value one needs to know the speaker, and time of utterance to fix referents for 'I' and 'now'. Perhaps it is also indexical with respect to the possible world at which it is to be evaluated. So an adequate truth theory would have as theorems sentences of the shape (Tint).
      (Tint)   'S' is true in L spoken by speaker p at time t in possible world w iff ...
Then the recursive theory of truth would define 'is true in L at i' where i would be a whole sequence of indices such as these. The attraction of this is that we have a way around the conclusion of the slingshot without resorting to Davidson's solution. Although R = 'grass is green' and S = 'snow is white' refer to the same thing (the True) in the actual world, it is reasonable to think their values differ at different possible worlds, hence the propositions assigned to R and S would be different. So the theory promises to have the right fineness of grain to handle meaning.
      3. However, Soames rejects the possible worlds answer, for reasons that have been well explored in the literature. While R and G are assigned different propositions, all necessary sentences will be assigned the same proposition, for necessary sentences are true in all possible worlds. This means that (say) '2+2=4' and '1+3=4' end up with the same meaning, which is counterintuitive.
      4. One response to the problem is to introduce "impossible" worlds - worlds that allow necessary truths to be false (cf. Quine: no sentence is unassailable), but Soames claims to have good arguments that this tactic must fail. There is a nice literature here for a project.
      5. Instead Soames suggests a solution that is deployed at a number of places in the literature, most notably by M. Cresswell in Structured Meanings. Take Frege's idea that a meaning is partly defined by a 'mode of presentation', that is, by the syntactic shape of the expression in question. So since '2+2=4' and '1+3=4' have different symbolic structures, their meanings will differ in concert. The idea is that the meaning of the first asserts that two added to itself amounts to four, but that the second says something about addition of one and three.
      6. Soames admits that a theory of meaning of this kind has what I take to be a serious failing. It is not plausible to think that the theorems concerning propositions generated in such a way describe what a speaker needs to know to be a master of a language. (Soames p. 30-31.) Therefore, the theory of meaning is not a theory of competence of the speaker. It cannot hope to describe processes people use to actually understand language. So it will have no applications to cognitive science. But I took it that one of the main attractions of Davidson's program or its successors would be to locate principles a hearer employs to understand what the speaker is saying.