Notre Dame Journal of Formal Logic
Volume 23, Number 2, April 1982

Common Sense in Semantics

JERROLD J. KATZ*

Q: What is the principal difference between your conception of semantics and other conceptions?

A: On my conception of the subject, semantics is the study of meaning. This may seem an odd characterization–like a historian saying that history is a study of the past. But the conception of semantics as the study of meaning is far from an uninformative truism. In fact, the conception expresses a quite controversial view of semantics which is, moreover, unique among approaches to the subject in the philosophy of language, linguistics, and logic, in taking senses or meanings, *as they present themselves in our ordinary linguistic experience*, to be the proper objects of study in semantics. Other approaches are reductionistic. They seek to reduce the ordinary notions of sense and meaning away, replacing them with something else regarded from the metaphysical perspective of the reductionist as philosophically more respectable or scientifically more tractable. Ever since Russell's attempt to treat meaning as reference, we have had one attempt after another to treat meaning as something else. There have been attempts to reduce it to behavior-controlling stimuli, to images, methods of verification, stereotypes, truth conditions, extensions in possible worlds, use, illocutionary act potential, perlocutionary potential of various sorts, and even physical inscriptions. Indeed, the history of philosophical semantics in this century might well be written as a succession of metaphysically inspired attempts to eliminate the ordinary notion of meaning or sense.

---

*This essay was presented in the form of lectures to the students and faculty of the Philosophy Department, University College London. I wish to thank this audience and, in particular, Gerald Cohen, Colin McGinn, Herbert Heidelberger, Hidé Ishiguro, and John L. Watling.

Received July 3, 1981

Q: Can you explain what you mean in saying that, on your approach, senses or meanings are taken as they present themselves in our ordinary experience with natural language?

A: There are three aspects of this claim. First, I want to claim that there *is* a particular way in which senses or meanings present themselves to us. Second, knowledge of this way is a matter of common sense. Third, this way constitutes the phenomena that a semantic theory is obligated to save.

Sense and meaning, as ordinarily conceived, present themselves as the aspect of the grammar of expressions and sentences on which their semantic properties and relations depend. It is surely a matter of common sense that a sentence like "People sometimes procrastinate" is meaningful but a sentence like "Falsehoods sometimes procrastinate" is not, or that expressions like "bank", "ring", and "visiting relatives" are ambiguous, or that sentences like "Perhaps it will rain" and "Maybe there will be rain" are synonymous, or that expressions like "happy" and "sad" are antonymous. Moreover, it is also common sense that an expression or sentence of a natural language is meaningful when it has a sense, meaningless when it has none, ambiguous when it has more than one sense, synonymous with another when they have the same sense, and antonymous with another when they have opposite senses. Such facts are as certain as the propositions on G. E. Moore's list in "A Defense of Common Sense".

That they are also the phenomena that theories in semantics are obliged to save is shown by the fact that theories which fail to square with them are ipso facto rejected. Conformity with these semantic facts has traditionally served as a condition of adequacy for theories of meaning. The standard criticism of Russell's theory of meaning has been that its equation of meaning with reference gets synonymy wrong: it falsely claims that merely coreferential expressions like "creature with a heart" and "creature with a kidney" are synonymous. To take another familiar example, the standard criticism of the classical empiricist equation of meaning with mental images is that it falsely claims that expressions like Lewis Carroll's "slithy toves" are meaningful and ones like "nevertheless", "besides", "implies", "infinity", "insofar as", etc., are not. The Russellian equation sacrifices synonymy; the classical empiricist equation sacrifices meaningfulness and meaninglessness.

An adequate theory of meaning, then, has to save the phenomena that common sense recognizes to be semantic facts. It has to correctly represent these phenomena. It cannot ignore them or get them wrong. If a linguistic or philosophical theory of meaning were to say nothing about the facts that pretheoretic intuition recognizes as semantic, it could no more claim to be an account of *meaning* than, say, quantum mechanics or Keynesian economics. If a theory were to address itself to such phenomena, but get them wrong, it could not be a correct account of meaning. Thus, the proper criterion for judging the account of meaning that a theory offers is whether meaning is explained in a way such that, together with appropriate auxiliary assumptions, the explanation implies all and only true predictions about the semantic properties and relations of sentences over the full range of sentences in the language.

Q: I take it then that, as you see it, the primary facts in semantics are facts about the meaningfulness, meaninglessness, ambiguity, synonymy, and other such intuitively discernible properties and relations of sentences. If semantics is the subject that explains what meaning is, it will have to say what meaningfulness is, what meaninglessness is, what multiplicity of meaning is, what sameness of meaning is, what opposition of meaning is, and so on. But how is the "and so on" to be completed?

A: No theory can, at the outset, specify completely what it is about. This happens late in the course of developing a theory when principles initially adopted to account for the original limited set of phenomena are extended (usually with much revision) to account for a wide range of phenomena outside the original set.

No doubt in semantics the pretheoretically clear cases of semantic properties and relations, meaningfulness, meaninglessness, ambiguity, synonymy, and antonymy, are not exhaustive. Accordingly, the initial task in developing a theory of meaning is to develop a representation scheme for describing the meaning of sentences, and then on the basis of such descriptions to frame necessary and sufficient conditions for meaningfulness, meaninglessness, ambiguity, and the other properties and relations in the set of pretheoretically clear cases. Once substantial progress has been made in accomplishing this initial task, it becomes possible to tackle the question of how to handle properties and relations that pretheoretic intuition does not categorize as semantic or as nonsemantic (either because intuitions are not strong enough or because they conflict). To answer this question, one has to determine whether the principles of semantic structure so far developed for defining properties and relations like meaningfulness, meaninglessness, and ambiguity, and synonymy also enable us to define as yet uncategorized properties and relations. If the principles enable us to define some previously uncategorized property or relation—or do with some straightforward extension—then we can say that it is semantic, too.

Consider an example: let us assume that meaningfulness, ambiguity, and synonymy constitute the clear cases of semantic properties and relations. We can show that the further property of redundancy exhibited in (1) but not (2)

(1) naked nude
(2) naked nudist

is semantic as well. This is shown by the definition (D_1)

(D_1) A modifier-head construction is redundant if, and only if, it is synonymous with the expression that is its head.

which defines redundancy in terms of the acknowledged semantic relation of synonymy. Thus, redundancy belongs to the subject-matter of semantics because it is definable in terms of something that does.

The boundary question for semantics becomes more complicated in connection with the cases that fall outside of semantics. We cannot, of course, conclude that a property or relation is outside the subject-matter of a science

just from the fact that it cannot at one particular time be fit into the pattern of definition and explanation. Some crucial principle might not yet have been discovered or might not yet be stated with sufficient generality. But fallibility does not prevent us from being able to argue that a further property or relation is outside the domain of the science. We can argue from what is known about the science at the time that we are unable to define the new property or relation for reasons that have to do with essential aspects of the science. The only consequence of fallibility is that the strength of such negative arguments must depend on the strength of the considerations supporting the principles from which the generalization proceeds.

For example, even at this early stage in the development of semantics, we can argue that the property of rhyme is nonsemantic. Rhyme involves correspondence of terminal sounds in a pair of expressions, but we have every reason to think that the principles underlying the definitions of meaningfulness, meaninglessness, ambiguity, synonymy and similar properties and relations make no reference to sound patterns.

Q: Do you think that some things are nonsemantic from our common sense standpoint?

A: Yes, but reductionist theorists have tried to make us think common sense is wrong on these points—and unfortunately, with considerable success. For example, I think that we not only exclude phonological properties like rhyme but also syntactic properties like morphological form and, more importantly, properties of the referents of words (e.g., those that appear in stereotypes), psychological associations (e.g., images), and social constraints on use (e.g., emotional values of "four"-letter words). Later I will show how common sense reasserts itself against reductionist criticism.

Q: Supposing, then, for present purposes, that we have a reasonably clear idea of the subject-matter of semantics, how, on your view, do we obtain knowledge of the variety of particular facts about meaningfulness, meaninglessness, ambiguity, synonymy, etc., of sentences?

A: Suppose you are asked by a foreigner whom you are helping to learn English whether sentence (3) is meaningful:

(3) Pigs do not fly south in the winter.

or whether (4) is ambiguous

(4) Mary's will can't be broken.

or whether (5) and (6) are synonymous

(5) Give your money to the poor
(6) Give the poor your money

I think you will unhesitatingly say that they are. You will believe, moreover, that in such clear cases you *know* they are on the basis of your grammatical intuitions about them. In other cases, you might hesitate or be unsure of what the right answer is. Your intuition is unclear. But in the clear cases your acquaintance with semantic facts, as with phonological and syntactic facts, is through grammatical intuition.

It is both our initial source and final arbiter in questions of grammatical fact. Indirect evidence, operational tests, behavioral correlates, and the like have to be checked against intuition in order to determine that they are genuinely evidence about, tests for, and correlates of the grammatical structure in question. For example, if an operational test or behavioral correlate for clause boundaries in terms of breath pauses conflicts with our intuitions about clause boundaries, it will be rejected as not testing for or correlating with the proper grammatical structure. Thus, when nonintuitive sources of information about semantic, syntactic, or phonological fact are employed in grammatical argumentation, their employment must be legitimized by ultimately grounding the facts in clear intuitions of fluent speakers. In irreconcilable conflicts with clear intuition, it is always the nonintuitional source of evidence that goes.

Q: I think we have clarified the question of the subject matter of semantics sufficiently to return to the claim that your approach is distinctive in accepting the common-sense notion of meaning as the proper object of study in semantics. Can you now review some of the main reductive attempts and show how common sense "reasserts itself" against them?

A: If I am correct to think that attempts to replace the notion of meaning with something else are bound to fail because meaning is meaning, and nothing else, then one should be able to find examples, in the case of each such attempt, which exhibit the difference between meaning and what has been introduced to replace it. I shall now try to show that one always does find such counterexamples. My moral will be that, although the reducing theory in these attempts–the theory of reference, associationist psychology, first-order logic, speech act theory, or what have you–may be a truly marvelous theory in all sorts of ways, it is a marvelous theory of something other than meaning.

Let us begin with the attempted behaviorist reduction which claims that meaning is best understood in terms of the notion of behavior-controlling stimuli and with the attempted mentalist reduction which claims that meanings are images or "pictures in the head". Since meaning is tied far more tightly to linguistic constructions than their utterances are tied to eliciting stimuli belonging to any well-defined class, the semantic properties and relations of linguistic constructions can be expected to diverge significantly from the significant features of eliciting stimuli. One consequence of this is that, on the behaviorist reduction, the distinction between meaningful and meaningless expressions collapses completely. Nonsense words, as the behaviorist psychologists themselves have shown, are stably elicited by certain well-defined experimental conditions, but meaningful expressions do not have stable or well-defined stimulus conditions (see [4]). The same problem plagues the equation of meaning with images: meaningful words like "if", "nonetheless", "insofar as" come out meaningless because we do not associate a mental picture with them, but nonsense expressions like "slithy toves" come out meaningful because we do. Another problem is that the vast amount of imagery connected with a word like "hell" sharply contrasts with its degree of ambiguity.

Now, neither of these two reductionist attempts is taken very seriously in philosophy nowadays. Yet attempts that *are* taken seriously suffer from

the same sort of difficulty—failing to save the phenomena of semantics. One of the most widely accepted of the "serious" reductionist programs is the attempt to equate meaning with use. This program began with a simplistic notion of use which sacrificed a number of semantic phenomena, including synonymy. For obscene words, to take one example, and pejorative terms, to take another, are different in use from their nonobscene and nonpejorative synonyms. The use of words like "piss" and "Jew-down" is influenced by a large number of factors, such as culturally associated attitudes, desire to be accepted in society, etc., which do not influence the use of their synonyms "urine" and "bargain down" in at all the same ways. The equation of meaning with use also sacrifices the distinction between meaningful and meaningless sentences: sentences that are too long or too syntactically complicated to be used—ones with over two hundred words or ones having ten center-embeddings—will be predicted to be meaningless even though built up from meaningful components by operations (e.g., conjunction) known to preserve meaningfulness. Furthermore, the equation sacrifices the phenomena of ambiguity and antonymy, since almost every word of the language has ironic as well as literal uses. The reduction is thus forced to predict that "beautiful", "happy", "clever", etc., are ambiguous between their sense and the sense of their antonyms.

Alston, in a well-known paper on meaning and use [1], suggests that the equation of meaning with use is basically sound but that use theorists have failed to clarify the notion of use sufficiently and ought to clarify it along Austinian lines. The suggestion to so clarify it has been most systematically developed by Searle [37]. But, as I shall argue now, development of Alston's suggestion does not succeed in removing the difficulties we have found with the unreconstructed notion of use. On the contrary, Searle's speech act theory reveals all the more clearly that such difficulties are inherent in the equation of meaning with use by showing that they are also present in the reconstructed notion of meaning as illocutionary force potential.

Searle's attempt to formulate a systematic theory of meaning as use founders on the same differences between meaning and use that we have found with the simplistic notion of use: cultural, social, psychological, and other factors which contribute nothing to the meanings of words, nonetheless strongly influence the way they are used in speech. These differences show up in Searle's speech act theory as incongruities in that theory's pattern of explanation. Instead of the kind of coherent explanation we find in genuine theories, we find an assortment of observations about language and extralinguistic factors of a cultural, social, and psychological nature (see [28]).

In Searle's account of syntactic indicators of illocutionary force, one of the rules for using an illocutionary force indicator like "promise" says that its use involves the undertaking of an obligation. Another rule says that its use involves reference to a future act (relative to the speech point) in which the speaker is the agent. Now two other of Searle's rules, in no relevant way distinguished from these two in his account, say that it must not be obvious that, in the normal course of events, the speaker will perform the act that fulfills the promise, and that the promisee wants the speaker to fulfill the promise. The former pair of rules is about aspects of English verbs and hence about the language, and the latter pair is about extralinguistic factors, con-

cerning the conditions for something like typical cases of promising. The former rules express components of the grammatical meaning of "promise" which also play a role in other grammatical phenomena, like the contrast in meaning between "promise" and "advise" (no obligation but reference to future act), on the one hand, and between "promise" and "thank" (obligation but no reference to future act), on the other. In contrast, the latter rules merely express aspects of the context of use that a speaker normally or usually takes into consideration in using language. Thus, unlike the obligation and future act conditions, the conditions expressed in the latter rules need not be satisfied in straightforward, literal uses of sentences of the form 'I promise to do *X*'. Oaths and pledges are given even when it is obvious that the speaker will do the act(s) in the normal course of events, e.g., oaths like honest Abe Lincoln's oath to uphold the constitution, or pledges like Romeo and Juliet's or promises where a formal commitment is necessary like some promises to love, honor, and obey. And a promise can be made when the last thing in the world that the promisee wants is to see the promise fulfilled, e.g., a student's promise to turn in his or her three-hundred word paper in time for the instructor to read it during vacation.

Perhaps the clearest instance of a nonsemantic rule in Searle's discussion are his preparatory rules for asserting, viz., that the speaker have reasons for the truth of the assertion and that it not be obvious to both speaker and hearer that the latter already knows the fact(s) asserted (see [37], p. 66). If the former were semantic, there would be no baseless assertions and nobody just shooting off his mouth. If the latter were semantic, there would be no boring relatives who always state the same tired views. If both were, our examination system might be jeopardized.

These considerations, although only some of the cases that can be marshalled to show the basic incongruity in Searle's pattern of explanation, suffice for present purposes to support my claim that Searle's account is a concretion of two distinct things: on the one hand, observations concerning the grammar of syntactic constructions, and on the other, observations concerning extragrammatical (sociological, psychological, etc.) aspects of the context of utterance use. Given that my claim is correct, a coherent system of explanation is obtained by eliminating speech act theory in favor of two distinct theories, one dealing with the way in which meaning is determined in the language, and the other dealing with the way in which extragrammatical factors influence a speaker's use of language.

But before we give up on Alston's suggestion that perhaps the simple notion of use can be adequately specified, we have to consider the other principal attempt in contemporary philosophy to explicate it, the explication of Austin's idea of perlocution instead of his idea of illocution. This attempt was initiated by Grice who saw how to restrict the notion of perlocution so that its utilization as the basis for an account of meaning does not produce absurd consequences. Since on the unrestricted notion of a perlocutionary act (an act performed *by* uttering something) causal effects of any sort, including breaking the listener's ear drum, count as perlocutionary acts, they become part of meaning. Grice's notion of meaning restricts the totality of causal effects of an utterance to just those that a speaker intends to produce

in the listeners in virtue of their recognition of this intention to produce them (see [19]).

Grice sketched a line of argument to show that this restricted perlocutionary notion is the basic notion of meaning (notions like that of the meaning of a sentence are derivative) and that the principles that underlie the reasoning of speakers and listeners in the performance of these perlocutionary acts enable us to specify the notion of use satisfactorily. I say Grice *sketched* such a line of argument because he never actually argues that his account of utterer's meaning does not itself presuppose the notion of sentence meaning. Without such an argument, the possibility is left open that sentence meaning is definable in terms of utterer's meaning only because the former notion has been assumed in the statement of the principles of intentional reasoning underlying utterer's meaning.

The fact that this possibility is left open is a serious difficulty for an application of Grice's notion of utterer's meaning to suitably specify the concept of use. But worse, Grice's own account of this notion seems to make explicit appeal to the grammatical notion of sentence meaning. This appeal occurs when Grice says that the speaker has a repertoire of semantic procedures for expression types of the language that "equip" the speaker to use these expression types in any circumstances in which the conditions specified in the procedures are met (see [19], pp. 231-234). Grice does not explain his notion of a semantic procedure, but the semantic procedure for the expression type "bachelor" equips speakers to use its tokens to refer to anything meeting the specified conditions of being an adult human male who never married. Given the role of the notion in Grice's account, the notion can be nothing but that of the content of the speaker's knowledge of the grammatical meaning of sentences.

Schiffer has tried to supply arguments for the claim that perlocutionary meaning is the basic notion (in [36], pp. 6-7). He offers two. The first is that what one would "normally or ordinarily mean by uttering x" is a necessary condition for knowing the meaning of the whole utterance type x ([36], p. 5). But, even if this claim were true, it would not establish the priority of the perlocutionary notion of meaning. The problem is the same circular presupposition. What a speaker normally or ordinarily means by uttering an expression may depend on the prior notion of a compositionally fixed, grammatic meaning. Grammatical meaning might be necessary to provide the norm against which the utterance is compared to determine that it has its normal or ordinary meaning.

Schiffer's second argument is that the priority of utterer's meaning follows from the fact that

> . . . it is possible for a person to mean something by uttering x even though x has no meaning. ([36], p. 7)

Schiffer gives the example of someone uttering "grrr" to inform someone of his or her anger. But the possibility of meaning something by an antecedently meaningless sign-token is not incompatible with the claim that grammatical meaning is prior, anymore than the possibility of meaning by a sign-token

the opposite of what its sign-type means is incompatible with this claim. If I say "Snidely is a fine friend", referring to someone who recently betrayed my friendship, the meaning of my sign-token is the meaning of the sign-type "Snidely is a rotten friend". The fact that it is possible for an utterance of "fine" in the proper tone of voice or occurring in the proper circumstances (i.e., where everyone knows of the betrayal) to mean the opposite of what "fine" means in the language does not at all show that the meaning of "rotten" which the utterance of "fine" bears on such occasions is not an antecedent part of English grammar. Ironic usage shows only that speakers, in addition to their knowledge of grammar, must have extragrammatical conventions for hooking up sign-tokens uttered in an ironic tone or in the proper circumstances with the grammatical meaning of its antonym.

Similarly, the fact that, in the example Schiffer gives, it is possible for a sign-token of "grrr" to have the meaning that the sentence "I am angry" has in the language need show nothing more than that such conventions also hook up the grammatical meaning of "I am angry" with such utterances of "grrr" (by exploiting extragrammatical knowledge that canine grrrs are causally connected with canine anger, perception of onomatopoetic relations, and recognition of the speaker's intent to express such a hook up, in the way the ironic usage exploits knowledge of the betrayal). In order for Schiffer to have shown that meaning is not a contextually prior aspect of the grammar of the language, he would have to have shown—as he clearly has not—that a theory that maintains the priority of grammatical meaning has no way to introduce new meanings for expressions, but must suppose that the perlocutionary meaning of *every* sign-token is associated with its sign-type in the grammar.

Such theories suppose only that every meaning functioning as the perlocutionary meaning of a sign-token is associated with *some* sign-type in the grammar. The point may be made clear if more is said about the conception of language and use underlying these theories. On this conception ([28], pp. 13-22), there are (as suggested in the discussion of Searle's speech act theory) two distinct theories: grammar and pragmatics. Both describe sound/meaning correlations, but the correlations are different. Grammar describes the correlation of sound-types with meanings in the language: the correlation is given by the compositional function that determines the meaning of syntactically complex expressions from the meanings of their parts and syntactic relations. Pragmatics describes the correlation of sound-tokens (specifically, uses of sentences) with their meanings in the context of utterance: the correlation is given by conventions which determine the utterance meaning of a use of a sentence as a function of its grammatical meaning in the language plus extragrammatical information about the context. Such conventions may connect the use of a sentence with a meaning that the sentence type does not have in the language (as in the example of an ironic use of "Snidely is a fine friend"), but they also may connect a use of a sentence with a meaning that the sentence type does have in the language (as in the case of a mathematics teacher saying in class "The number two is the only even prime"). Pragmatics, on this view, is about how knowledge of context enables speakers to diverge from the compositional sound/meaning correlation in the language without loss of comprehension. The significant feature of this conception of the

relation between language and its use is that, being drawn from a grammatically specified sound/meaning correlation, utterance meanings themselves are not another kind of meaning, but simply grammatical meanings under a different correction.

We may even suppose that Grice's principles are the correct account of the conventions that correlate utterance-tokens with grammatical meanings: speakers use an expression to mean something by uttering it with the intention to produce appropriate comprehension by virtue of the audience's recognition of that intention. These principles, as we have seen above, require a repertory of semantic procedures that connect signs with meanings, that is, with grammatically determined conditions for, inter alia, the reference of the signs. The above conception of grammar and pragmatics satisfies this requirement by taking the speaker's knowledge of the grammatical sound/meaning correlation to be this repertoire of semantic procedures. The conception thereby makes it possible to plug a gap in Grice's pragmatic theory. In so doing, however, we adopt just the grammatical/extragrammatical distinction that prevented Searle's systematization of Austin's notion of illocution from saving the simplistic notion of use, thus also preventing Grice's notion of perlocution from saving it.

There is another important tradition of reductionist thinking. It attempts to sophisticate Russell's equation of meaning with reference in much the same way that Searle's and Grice's work attempts to sophisticate the equation of meaning with use. The principal contemporary figure in this tradition is Davidson. He presents a program for replacing the intensionalist paradigm for semantic analysis '*s* means *p*' by the extensionalist paradigm '"*s*" is true if, and only if, *p*'.[1] I will show that this program makes the same equation of meaning with reference as Russell's and, hence, sacrifices the same range of semantic phenomena.

As we saw at the outset, Russell's equation sacrifices synonymy: coreferential expressions like "creature with a heart" and "creature with a kidney" must be counted as synonymous even though they are not. Davidson's program sacrifices synonymy, too. Coreferential sentences like

(7) Snow is white
(8) Grass is green

are counted as the same in meaning since they are the same in truth value.[2] Davidson's program treats (10) as an equally good account of the meaning or logical structure of (7) as (9).

(9) "Snow is white" is true iff snow is white
(10) "Snow is white" is true iff grass is green

Moreover, since the truth conditions for every true sentence of English can be given on the basis of any true sentence, and the truth conditions for every false sentence of English can be given on the basis of any false sentence, on Davidson's program there is an optimal semantic analysis of English consisting of an infinite list of biconditionals in which each true sentence is paired with "Two plus two equals four" and another in which each false sentence is paired with "Two plus two equals five". Therefore, Davidson's program posits that English has only two meanings.

Davidson is himself aware of the fact that his program is at this point counterintuitive, but he claims that this "grotesqueness" (Davidson's own description) is not really a vice of his program. He says,

> . . . the grotesqueness of [(10)] is in itself nothing against a theory of which it is a consequence, provided the theory gives the correct results for every sentence . . . if [(10)] followed from a characterization of the predicate 'is true' that led to the invariable pairing of truths with truths and falsehoods with falsehoods–then there would not, I think, be anything essential to the idea of meaning that remained to be captured. ([9], p. 457)

There are three things to be said to this. First, Davidson needs to show that the pairing of truths with truths and falsehoods with falsehoods is all there is to meaning. But when such pairings lead to grotesque consequences, he says, well, so what, pairing truth with truths and falsehoods with falsehoods is all there is to meaning. Davidson is here simply arguing from his own theory. Second, what makes the consequence grotesque is that it flatly contradicts strong pretheoretic intuitions that constitute our firmest hold on the notion of meaning; hence, the theory implying these consequences ought to go, not the intuitions. Third, if in the case of less sophisticated equations of meaning with reference, like Russell's, we take such grotesque consequences to refute the attempted reduction, we cannot consistently turn our backs on the guidance of pretheoretic intuitions in essentially the same situation.

Davidson tries to soften the blow of such grotesqueness by saying:

> It may help to reflect that [(10)] is acceptable, if it is, because we are independently sure of the truth of [(7)] and [(8)] but in cases where we are unsure of the truth of a sentence, we can have confidence in a characterization of the truth predicate only if it pairs that sentence with one we have good reason to believe equivalent. It would be ill-advised for someone who had any doubts about the color of snow or grass to accept a theory that yielded [(10)], even if his doubts were of equal degree, unless he thought the color of the one was tied to the color of the other. ([9], p. 312)

But how can reflections about people's degree of confidence in sentences be relevant to the adequacy of Davidson's claims about meaning? The existence of such grotesque consequences depends only on whether biconditionals like (10) are true. Moreover, the particular example doesn't matter; indefinitely many such biconditionals *are* true–whether we are skeptical or not–and hence indefinitely many grotesque consequences follow–whether we recognize them or not.[3]

Davidson also tries to absolve his theory from blame by separating it from such alledgedly questionable examples. "It is not easy," he says, "to see how [(10)] could be party to [the enterprise of characterizing the truth predicate]" ([9], p. 457). Davidson's claim is that his theory proper ought not to be blamed for such grotesque consequences because it is not itself responsible for asserting that (7) and (8) are equivalent. But it is hard to see why it matters whether such sentences are involved in the enterprise of characterizing the truth predicate in any direct sense. We would not let another theory, say, a theory of color perception, off the hook if, in conjunction with simple truths like grass is green and snow is white, it implies grotesque conse-

quences about vision. Even though theories are not responsible for the auxiliary assumptions used to deduce consequences from them, *good* theories do not have grotesque consequences when conjoined with simple truths. Since the grotesqueness does not derive from the simple truths themselves, and since it has its origin somewhere, we take it that the grotesque consequences must originate in some hidden weakness of the theory. The simple truths merely help to reveal their source.

Although (10) and similar cases may not be directly involved in the enterprise of characterizing the truth predicate, other cases which raise the same problem for Davidson's theory are. Let the sentences $S_1, \ldots, S_n$ be a Davidsonian characterization of the truth predicate for English. Then we obtain essentially the same grotesque consequences from $S_1, \ldots, S_n$ by replacing "*s*" in the schema *T* by a name or structural description of a sentence S_i in $S_1, \ldots, S_n$ and replacing "*p*" with a different sentence S_j in $S_1, \ldots, S_n$. Since the sentences $S_1, \ldots, S_n$ are the axioms of Davidson's theory, they are presumably all true and nonsynonymous with one another. Hence, the biconditional resulting from these replacements will express a grotesqueness just as (10) does. But unlike (10) cannot be disavowable as not being party to the enterprise of characterizing the truth predicate.

Closely related to Davidson's reductionism is a family of positions in the Russellian tradition that may be collectively termed "epistemic role theories". These include Davidson's own theory of interpretation, Harman's conceptual role theory, Field's probabilistic semantics, and others.[4] Epistemic role theories assimilate language to belief systems in the same manner Davidson's move from the '*s* means *p*' paradigm to the '"*s*" is true if, and only if, *p*' paradigm, and these theories accord complexes of synthetic beliefs the functions that analytic beliefs have in classical theories of meaning, in order to eliminate meanings in favor of complexes of beliefs with the same causal role in the speaker's use of language (see [20] and [12]). Different epistemic role theories choose different complexes of beliefs or handle their causal role in language use in different ways, but all stand or fall on whether their common assimilation of language to belief systems sacrifices or saves the semantic phenomena.

That such theories sacrifice semantic phenomena is clear from the case of common and proper nouns. The use of common nouns, on these positions, "must depend on a background of largely unmentioned and unquestioned true beliefs" ([10], p. 21). For example, the reference of "the earth" is fixed by such beliefs as: "this earth of ours is part of the solar system, a system partly identified by the fact that it is a gaggle of large, cool, solid bodies circling around a very large, hot star" ([10], p. 21). But, then, the use of proper nouns is fixed, and according to these positions is fixed, in the same way. For example, the reference of a proper noun like "Hitler" will be fixed by such beliefs as that he was dictator of Nazi Germany, propounded racial doctrines about the superiority of Aryans, carried out the destruction of millions of European Jews, and so on. Given the parallel treatment of common and proper nouns, epistemic role theories have to say that *both* have meaning, just as Russell's equation of meaning with reference has to say that "creature with a heart" and "creature with a kidney" have the same meaning.

However, facts about the phenomena of antonymy, synonymy, ambiguity, and redundancy in English show that common nouns have meaning but proper nouns do not. For example, common nouns have antonyms, or expressions with incompatible meaning, but proper nouns do not. The antonym of "bachelor" is "spinster" but what is the antonym of "Socrates" or an expression incompatible with "Bertrand Russell"? Mill made virtually this point. Putting his point in the formal mode, we would say that the proper noun "Dartmouth" has no meaning since it is not incompatible with "not located at the mouth of the river Dart". The same semantic asymmetry between common and proper nouns is found in synonymy phenomena: common nouns have synonyms, but proper nouns do not. "Marijuana" and "pot" refer to the same substance *because* they have the same meaning; "Mark Twain" and "Samuel Clemens" refer to the same literary figure but the explanation cannot be sameness of meaning. Note that it is absurd to say "Marijuana might not be pot" but not absurd to say "Mark Twain might not be Samuel Clemens". Saying the former is like saying that genuine coins might be counterfeit, but one can easily imagine a revisionist literary historian saying the latter.

Nor is the multiple reference of a proper noun a matter of multiple sense. Although it is true to say that the multiple reference of "bank" is due to its ambiguity, it is absurd to explain the multiple reference of a proper noun like "Mary Jane" (to a girl, a boat, a pet turtle, etc.) in the same way. Finally, nonrestrictive relative clauses on proper nouns do not exhibit semantic redundancy, whereas on common nouns they do. For example, in "The nightmare, which is a dream, . . . " or "The king, who is a monarch, . . . " the relative clause is redundant. But in "Gödel, who discovered the incompleteness of arithmetic, . . . " or "Aristotle, the most famous student of Plato and teacher of Alexander, . . . " the clause is not redundant. It is amplitive, telling us which Gödel (the mathematician, not the Bauhaus architect) and which Aristotle (the philosopher, not the ship owner).

Thus, given the commitment of epistemic role theories to proper nouns and common nouns *both* having meaning, the theories sacrifice not only the distinction between the meaningful and the meaningless but also the semantic phenomena of antonymy, synonymy, ambiguity, and redundancy. Epistemic role thories may be quite marvelous theories of something, but it isn't meaning.

There is another important objection to taking such theories to be theories of meaning. The principal thing about the notion of epistemic role is that it has to do essentially with mental states and processes: Beliefs are themselves mental states and the epistemic role of a complex of beliefs is the causal role it has in thought processes. Thus, suppose that we had a reconstruction of meaning in terms of the notion of epistemic role and suppose even further that, as far as we know, the reconstruction sacrifices no semantic phenomena. Ought we adopt it?

Meaning is, on all sides, regarded as referring to the grammatical features of sentences on which their logical implications depend. Adopting the reconstruction will, accordingly, commit us to characterizing logical implications in terms of cognitive structures assigned to sentences in the reconstruction. Since, furthermore, it is an *empirical* question what such cognitive structures

are, it will be an empirical question—presumably in psychology—what sentences the reconstruction will say that a given sentence implies. It may turn out that the reconstruction contains no surprises for the logician. "Henry is a bachelor" implies "Henry is unmarried", "Henry is a bachelor" implies "Henry is a bachelor or Henry is American", and so on. But it may also turn out that the reconstruction contains some surprises: "Henry is English" implies "Henry is stuffy". Presumably, the surprises won't really appear in connection with such simple cases. But it *is* ex hypothesi an empirical question, and therefore, in adopting the reconstruction, we commit ourselves to the possibility of arriving at logical principles that conflict with laws of logic. About such a commitment, Frege wrote:

> . . . what if beings were found whose laws of thought flatly contradicted ours and therefore frequently led to contrary results even in practice? The psychological logician could only acknowledge the fact and say simply: those laws hold for them, these laws hold for us. I should say: we have here a hitherto unknown form of madness. Anyone who understands laws of logic to be laws that prescribe the way in which one ought to think— . . . and not natural laws of human beings' taking a thing to be true—will ask, who is right? . . . The psychological logician cannot ask this question; if he did he would be recognizing laws of truth that were not laws of psychology. ([17], p. 14; see also pp. 15-16)

Since epistemic role theories allow the possibility for arriving at logical principles that conflict with the laws of logic, epistemic role theories may be quite marvelous theories of something like the psychology of human reasoning but they cannot be theories of meaning.

Q: Let me interrupt a moment. Doesn't this criticism apply as much to your own view of semantics as it does to the epistemic role view? In "Mentalism in Linguistics" [23] you also claim that grammatical structure is psychological structure.

A: Just such considerations as Frege's were responsible for changing my view of the ontology of grammatical structure from the psychological position I took in that paper to something like Frege's. The difference between my theory of semantics and epistemic role theories, which allows me but not epistemic role theorists to consider more than one ontological option, is that my theory is a nonreductive explication of the formal grammatical structure underlying semantic properties and relations while their theories are reductive programs proposing to replace meaning with a specifically psychological notion. Thus, they are wedded to psychologism, whereas nothing forces me to interpret the semantic representations as representations of mental structure. I am free to interpret them as representations of abstract objects. Very little changes in the formal theory of semantic structure, say, as set out in earlier works like *Semantic Theory* [25]. Epistemic role theories do not have a nonpsychological option because the notion of epistemic role is by definition a psychological notion.

Before returning to reductionism in semantics, I want to say a bit about my present view concerning what a theory of semantic structure is a theory of.

During most of the time I was developing my conception of intensional semantics, I assumed that a theory of semantic structure is part of a theory of linguistic competence in Chomsky's sense, namely, part of a theory of the ideal speaker-hearer's knowledge of the language. I thought that a theory of semantic structure is properly viewed as a theory of semantic competence, that is, a theory of the ideal speaker-hearer's knowledge of meaning. My present view is that semantics, and the rest of grammar as well, is better thought of as being about a class of abstract objects, sentences, rather than being about human knowledge, however idealized. I now think linguistics is a branch of mathematics, not a branch of psychology (see [32]).

This, of course, is not to say that there is no such thing as a study of human *knowledge* of grammatical structure. There surely is. Such a study is as legitimate, on my view, as psychological investigations into our knowledge of number. But it is as *il*legitimate, on my view, to treat the study of semantic structure as psychology as it is to treat the study of numerical structure as psychology. A psychological study of the ideal speaker's knowledge of meanings is as distinct from the study of meanings as the study of the ideal arithmetician's knowledge of number is from the study of number. And distinct in the same way: the study of a competence is a study of human *knowledge* while the study of meaning or number is the study of the *object* of such knowledge. Hence, on my present view, semantic theories, and theories of grammar, too, are not required to meet any empirical, "psychological reality" constraints, reflecting the way in which human knowledge is mentally represented and processed. Rather, a semantic theory of a natural language is required to do no more or less than account for the semantic facts about its sentences—which are meaningful and which meaningless, which ambiguous and how many ways, and so on—in the simplest, most revealing way. There may be more than one such theory, but then, as equally simple, equally revealing, comprehensive accounts of the semantic facts of the language, they will be equivalent and equally correct.[5]

Now, I'd like to return to reductionism. There is one more form of reductionism in the Russellian tradition that we need to consider. This form, which I will call "possible world semantics", equates meaning with extensions in possible worlds. Possible world semantics is a reaction to narrowly truth-theoretic approaches like Davidson's. It prides itself on a more powerful theory of language than is possible within the Davidsonian framework but one that is not so powerful as to embrace meanings in the classical sense. It claims to be able to explain more than Davidsonians can without being subject to Quine's criticisms of classical semantics. But, in fact, it neither avoids Quine's criticisms nor succeeds in explaining enough semantic phenomena.

Possible world semantics is just as vulnerable to Quine's criticisms as classical semantics. The reason, as Quine himself pointed out ([34], p. 47), is that the notion of possibility is bound up with that of meaning. The possibility of a world is a matter of the consistency of its description, but consistency is partly a matter of meaning: there are no possible worlds corresponding to the descriptions "married bachelor", "mortal who will live forever", and "genuine coin of the realm which is counterfeit" because

the descriptions are semantically inconsistent. Since embracing the notion of possibility is also embracing meaning, possible world semanticists like Hintikka can hardly promote their theory as superseding the classical theory of meaning.[6]

In the course of this critical examination of reductionism in semantics, we have seen a pattern emerge: there is an initial suggestion that meaning be reduced to some notion that, on examination, turns out to sacrifice semantic phenomena; one or another revision of the notion is developed, but soon these, too, are shown to sacrifice the same semantic phenomena. We saw this pattern exemplified in the suggestion that meaning is use and Searle's revision of the simplistic notion of use; we saw it exemplified again in the suggestion that meaning is reference and the revisions of Davidson and epistemic role theorists. Now we see it exemplified once more in the revision we are calling possible world semantics. This should not be surprising insofar as all of the reductionist approaches in the Russellian tradition use purely extensional apparatus to handle the phenomena for which classical semantics invokes the concept of meaning. Possible world semantics extends the notion of reference in the Russellian and Davidsonian approaches to reference in all possible worlds, but this extension, since it leaves the basically referential character of semantic analyses unchanged, must sacrifice the same semantic phenomena as the Russellian and Davidsonian approaches.

The simple equation of meaning with reference leads to the definitions (D_2) and (D_3).

(D_2) An expression or sentence e is meaningful (has a meaning) just in case e refers to something in the actual world, i.e., has a nonnull extension.

(D_3) An expression or sentence e_1 is synonymous with another e_2 (has the same meaning) just in case e_1 and e_2 refer to the same thing(s) in the actual world, i.e., have the same extension.

These, as we have seen, face counterexamples: "witch", "unicorn", "golden mountain", etc., refer to nothing in the actual world but are meaningful; "creature with a heart" and "creature with a kidney" co-refer in the actual world but are different in meaning. Possible world semanticists replace definitions (D_2) and (D_3) with (D_4) and (D_5)[7]:

(D_4) An expression or sentence e is meaningful (has a meaning) just in case e refers to something in at least one possible world, i.e., there is a possible world in which e has a nonnull extension.

(D_5) An expression or sentence e_1 is synonymous with another e_2 (has the same meaning) just in case e_1 and e_2 refer to the same thing(s) in all possible worlds, i.e., have the same extension in every possible world.

With (D_4) and (D_5), the foregoing counterexamples disappear: there are hypothetical circumstances containing unicorns, witches, and golden mountains, and hypothetical circumstances where creatures with hearts have no kidney.

But, although these counterexamples disappear, others just like them remain. To find such counterexamples, we do what critics of simpler referential

approaches did, namely, find nonsemantic factors that constrain the extensions of expressions and sentences so that their meanings can vary independently of the extensional structure recognized in referential theories. Critics of the simple referential theory found evolutionary and biological constraints that precluded meaningful expressions like "unicorn" from having a referent in the actual world and ensured that nonsynonymous expressions like "creature with a heart" and "creature with a kidney" are coreferential. What constraints will serve the same critical purpose in the case of the more sophisticated possible world approach?

Part of the notion of possibility, as we observed above, depends on the notion of meaning. But only part. Part of the notion of possibility and necessity depends on extralinguistic matters of logical and mathematical fact. Such fact constrains extension in all possible worlds independent of a wide range of semantic variation. Thus, logical and mathematical fact gives rise to counterexamples that are exactly parallel to those that have been brought against the simpler referential theories. For example, corresponding to the counterexamples to (D_2), we have expressions like "the largest natural number", "a consistent division by zero", and "a consistent and complete formalization of arithmetic" which are meaningful but have a nonnull extension in every possible world. And, corresponding to the counterexamples to (D_3), we have sentences like "One plus one is two" and "Every even number is the sum of two primes" which are nonsynonymous but have the same extension in every possible world (for further discussion, see [22], pp. 86-96).

Possible world semantics may be thought to fare better than, say, Davidsonian semantics because the former allows more than two meanings in English. But, just as Davidsonian semantics counterintuitively claims that all true sentences have the same meaning and all false sentences have the same meaning, possible world semantics counterintuitively claims that all necessarily true sentences have the same meaning and all necessarily false sentences have the same meaning.

Q: Your mention of Quine raises the question of whether it is legitimate of you to assume a domain of meaning. I can grant you that your arguments go through on this assumption: *if* there is such a thing as meaning, then it is irreducible. But if Quine has shown that there is no such thing as meaning, these various nonintensional approaches are saved. The various equations that you have been criticizing thus no longer need be viewed as reductionist programs, but can now be viewed as proposals concerning how best to handle the logical structure of natural language in the absence of any hope to handle it on the basis of the classical notion of meaning.

A: Your point is exactly right. But it works the other way as well. What makes Wittgensteinian, Davidsonian, and other reductive approaches seem viable in spite of their failure to save semantic phenomena is the belief that Quine has proved that the sacrificed phenomena were not worth saving in the first place. Let's accept the biconditional that these approaches are off the hook if, and only if, it has been shown that there is no such thing as meaning in the classical sense. But the fact is that Quine's arguments against meaning show nothing of the sort.

These arguments, set out in "Two Dogmas of Empiricism" [34], have a structure that has eluded Quine's friends and foes alike. This structure has to be fully appreciated in order to understand why Quine's criticisms fail.

Grice and Strawson picture Quine's overall argument as a non sequitur, generalizing from cases of failure to draw the analytic-synthetic distinction to the sweeping conclusion that no distinction exists [18]. Thus, on their view, even if each of Quine's specific criticisms were cogent, there would be no grounds to conclude

> That there is such a distinction to be drawn at all is an unempirical dogma of empiricists, a metaphysical article of faith. ([34], p. 37)

But Grice and Strawson fail to appreciate that the cases Quine examines are, in an appropriate sense, exhaustive. Quine's overall argument considers *all* the places where it would be reasonable to expect clarification of the notion of meaning. Hence, if Quine's specific criticisms in each place were cogent, he would have shown that there can be no clarification of this notion in any place where one might reasonably expect it.

There are three such places: definition, logic, and linguistics. In the area of definition, there are three forms of definition to consider, lexical definition (paraphrase), explication, and abbreviational stipulation. Quine quite easily shows that a clarification of meaning cannot come in any of these forms. In the area of logic, he shows, convincingly, that the application of postulational or rule methods to the problem of clarifying meaning—as exemplified in Carnap's use of meaning postulates or semantical rules—can only specify which sentences of a language are analytic and which pairs synonymous, but cannot explain what analyticity or synonymy is ([34], pp. 32-37). Finally, in linguistics Quine shows, again convincingly, that attempts to clarify semantic notions by providing substitution criteria for identifying their extensions leads to vicious circularity.

If Quine could in this way establish that neither definition, nor logic, nor linguistics offers hope of clarifying meaning, he could conclude that there is no place from which such clarification will be forthcoming. The problem with Quine's argument, then, is not that it is a bad induction. It is a deductive argument from an exhaustive enumeration of cases. The problem is rather that what is shown in connection with one of the cases does not rule out clarification of meaning in that area. Quine shows that no clarification is possible in the areas of definition and logic, but what he says about linguistics does not rule out clarification there. Failure to provide substitution criteria for analyticity and synonymy does not rule out their clarification.

It is easy to see why Quine thought that it does. He says:

> So-called substitution criteria, or conditions of interchangeability, have in one form or another played central roles in modern grammar. For the synonymy problem of semantics such an approach seems more obvious still. ([34], p. 56)

Modern grammar meant, at the time this was written, the taxonomic theory expounded in the works of Bloomfield, Bloch, and Trager, etc. (these are the "modern grammarians" to whom Quine himself refers; see [34], pp. 50 and

52). Quine thus thought that failure to provide substitution criteria rules out the possibility of clarifying semantic notions in linguistics because the scientific orthodoxy of the time, taxonomic theory, said that such criteria are the only way to clarify linguistic notions. However, since the Chomskian revolution, generative theory has introduced a new way of clarifying linguistic notions. Quine's grounds for taking substitution criteria as the proper standard for clarifying linguistic notions are thus undercut and his argument to rule out clarification of semantic notions collapses.

Philosophical faith in the power of Quine's arguments against meaning rests on a failure to appreciate this consequence of the Chomskian revolution. Taxonomic theorists insisted that substitution criteria are the only legitimate means of clarifying linguistic notions because they held a physicalist conception of language which required all such notions to be built up from distributional regularities in a corpus of utterances. The Chomskian revolution replaced this conception with a psychological conception from which no such requirement follows. Instead of viewing grammars as data-cataloguing devices, as the taxonomists had, Chomskian linguists viewed grammars as scientific theories of the ideal speaker's knowledge of the language. This view allows the clarification of a linguistic notion to be a matter of its connections with other notions in a predictively powerful theory of sentence structure. Hence, instead of requiring that a linguistic notion be built up from the empirical base, it can, as it were, be dropped down from above (see [6], pp. 31-33).

There is, then, the option of explaining meaning, synonymy, and analyticity on the model of Chomsky's explanation of syntactic notions like 'well-formed'. We can construct an abstract system of semantic representations that formally describes the meaning of sentences, characterize semantic notions like meaningfulness, synonymy, and analyticity in terms of such formal representations, and then justify both the representational system and the definitions indirectly on the basis of how well they predict and explain judgments of fluent speakers about such semantic properties and relations of sentences. Quine's criticisms do not apply to this way of clarifying semantic notions: such theoretical clarification of semantic notions can be no more circular than the clarification of syntactic notions like 'well-formedness' on which it is modeled. Indeed, if there were an a priori argument establishing the vicious circularity of semantics, there would also be an argument establishing the vicious circularity of syntax, and further one establishing the vicious circularity of logic and mathematics (the notions of 'is logically equivalent to' and 'is numerically identical with' would fare no better than 'is the same in meaning as' judged by the standard of substitution criteria). Quine's demonstration of circularity is best viewed as a *reductio ad absurdum* of his claim that substitution criteria are a proper standard of clarification.

Q: Couldn't Quine bring in his thesis of the indeterminacy of translation to replace this earlier criticism of meaning?

A: No. Quine's defense of this thesis depends on the criticism of meaning in "Two Dogmas of Empiricism". Recall the point in *Word and Object* where Quine tries to respond to the objection that indeterminacy may be nothing over and above ordinary inductive risk ([35], p. 75). Quine's response is that

to suppose that truth in translation is no worse off than truth in physics "misjudges the parallel". We suppose that, in semantics, there is a language-independent semantic reality that constitutes the meanings that sentences and their translations express, just as in physics there is a theory-independent physical reality. We misjudge the parallel because there is no semantic reality for hypotheses of translation to be true or false of, no "free-floating, linguistically neutral meaning", as Quine puts it ([35], p. 76). There are, according to Quine, no meanings over and above the sentences of a language or languages for translation hypotheses to be right or wrong about:

> . . . radical translation tries our meanings: really sets them over against their verbal embodiments, or, more typically, finds nothing there. ([35], p. 76)

Quine *says* that there is no "linguistically neutral meaning" but what reasons does he have for saying this? The only reasons he has given is, first, those in "Two Dogmas of Empiricism" with which we have already dealt, and second, ones that depend on the same taxonomic picture of language as the argument concerning the circularity of attempts to clarify meaning on the basis of substitution tests. Quine claims that translation hypotheses like:

(i) "gavagai" translates as "rabbit"
(ii) "gavagai" translates as "rabbit stage"
(iii) "gavagai" translates as "undetached rabbit part"

can accommodate equally well any evidence from the speech dispositions of native informants ([35], pp. 72-86). But this is so only as long as what Quine allows as evidence are just judgments about extensional relations, expressed, typically, in assenting and dissenting behavior to queries about the application of an English term to a present gavagai. It seems clear that no equi-accommodation claim strong enough to support indeterminacy can be made (even granting compensatory adjustments in the translation hypotheses for other expressions) if we are allowed to consider evidence in the form of judgments about intensional relations. If we can ask a bilingual informant questions such as 'Is "gavagi" *synonymous with* "rabbit"?' or 'Does "gavagai" *bear the same meaning relation to* "rabbit" that "finger" bears to "hand" or "branch" bears to "tree"?' or 'Is the expression "kicking a gavagai" *closer in meaning to* "kicking John in the head" or "kicking John's head",' then there is no reason to think we will be unable to rule out translation hypotheses because they accommodate the available evidence less well than others. Extension of semantic skepticism to intersubject agreement, the treatment of other expressions, or any combination of such factors can be met with compensatory refinement of the eliciting conditions, augmentation of the informant pool, increase in the variety of the questions, construction of a theory which reveals underlying principles (explains deviant judgments, etc.), and so on. Nothing in this interplay between criticism and construction, which is familiar in all sciences, suggests more than mere evidential underdetermination of semantic hypotheses.

Quine would of course take the introduction of evidence from judgments about intensional relations to beg the question. He would claim that such evidence illegitimately presupposes that field linguist and informant under-

stand notions like 'synonymy', 'bears the same meaning relation', and 'is closer in meaning than' in the same way. Quine would also claim that introducing such evidence illegitimately presupposes that these notions can be made sufficiently clear.

Such presuppositions are indeed made but there is nothing illegitimate in making them. These presuppositions seem circular to Quine because the taxonomic theory of grammar demands that all linguistic notions be built up from the data without appealing to notions that have not already received physicalistic certification. But, independently of a taxonomic bias, there is nothing illegitimate about presupposing common understanding of semantic notions or the possibility of explicating them. One pursues theory construction in science on the basis of such presuppositions. One takes the attitude that such a course is the best way to determine whether or not they are true. If the presuppositions are false, attempts at such theory construction will come to naught; if they are true, the attempts will eventually be successful and we will learn in this way that the presuppositions are true. Since in sciences generally, we permit theory construction to test the faith we put in the existence of a subject-matter, why not in semantics, too?

Given that reductive approaches get off the hook in connection with their failure to save semantic phenomena if, and only if, Quine's skeptical arguments show that there is no such thing as meaning in the classical sense,[8] we may conclude, having removed these skeptical arguments, that such approaches are inadequate. Without skepticism about meaning, the failure of purported reductions of meaning to account for meaningfulness, ambiguity, synonymy, etc., has to be viewed as showing that these reductive approaches to meaning offer the wrong explanation of what meaning is.

Q: This may be a good point to turn to your conception of a semantic theory. Could you start by saying why you think your approach will offer the right explanation of meaning?

A: The principal reason for thinking so is that my approach is nonreductive. It does not assume that meaning is something else and then characterize the task for semantics as that of showing how meaning can be reduced to it. Instead, my approach assumes the viability of the common sense notion of meaning and characterizes the task for semantics as that of constructing a theory of this notion that saves *all* the semantic phenomena with which common sense acquaints us. Thus, my approach starts with no preconceptions about the nature of meaning: it goes on the more modest assumption that the nature of meaning may not be known until we have proceeded quite far in developing such a semantic theory and that our knowledge of its nature will be in terms of what a sophisticated semantic theory of this kind says it is. Another reason for confidence in this approach is that the semantic theory thus far developed confirms our faith in the common sense notion of meaning by shedding light on a number of heretofore recalcitrant problems in the philosophy of logic and language.

A semantic theory is a metatheory. It is a metatheory for theories about the semantic side of the correlation of sentences and senses in natural lan-

guages, that is, for the *semantic components* of grammars. A semantic theory has two parts, first, *a notation scheme* for formally representing the senses of sentences in natural languages, and second, *a complete set of definitions* for the semantic properties and relations of natural language. The notation scheme provides a set of *semantic representations* which are employed in semantic components to describe the senses of sentences. The definitions explain semantic properties and relations of sentences in terms of the senses of sentences; hence, the definitions are stated in the form of configurations of symbols in semantic representations. Such configurations constitute generalizations to the effect that the semantic structure they define *is* the grammatical basis for a sentence having the semantic property or relation. Accordingly, the assignment of semantic representations to sentences in a grammar automatically makes a set of predictions about the semantic properties and relations of the sentences. In this way, the success and failure of these predictions reflect on the metatheory as well as on the particular grammars that make them. Success confirms the representation of the correlation of sentences and senses in the grammar, the statement of the definition(s) used, and the notation scheme; failure disconfirms at least one of them. Hence, we can say that a semantic theory saves the semantic phenomena of language just in case its definitions characterize the condition for meaningfulness, meaninglessness, ambiguity, synonymy, etc., so that, semantic representations (in the notation of the definitions) can be assigned to the sentences of each natural language in a way that, under the definitions, makes all and only true predictions about their semantic properties and relations.

To make these ideas clearer, let us begin with a first approximation to a semantic theory that saves the phenomena in this sense. The minimal condition for a theory to be an *intensional* theory of the logical structure of natural language is that its notation scheme be a notation for senses. In this sense, Frege proposed an intensionalist theory of logical structure in natural language [14]. But the condition can be satisfied by a very restricted theory. Let the vocabulary of the notation scheme be the infinite list of numericals "1", "2", ... Each numerical is taken as the designation of a distinct sense. Let the semantic representations in this vocabulary be finite sets of such numericals including the null set. Let the definitions be (D_6)-(D_8):

(D_6) A sentence or constituent is meaningful if, and only if, the grammar assigns it a semantic representation; otherwise, it is meaningless, that is, semantically deviant.

(D_7) A sentence or constituent is ambiguous if, and only if, the grammar assigns it two or more semantic representations.

(D_8) Two sentences or constituents are weakly synonymous if, and only if, the semantic representations assigned to them have a member in common.

(D_8') Two sentences or constituents are strongly synonymous if, and only if, they are assigned the same semantic representations.

Let us imagine a grammar of English which assigns "creature with a heart" and "creature with a kidney" the semantic representations {17} and {27}, "One plus one equals two" and "Every even number is the sum of two primes" the semantic representations {875} and {61}, "piss" and "urine" {74} and {74}, and "Richard M. Nixon" the null set. Then, the first two expressions are predicted to be meaningful, unambiguous, and nonsynonymous; the first two sentences likewise; the second two expressions are predicted to be meaningful, unambiguous, and synonymous, and the last is predicted to be meaningless.

Since we have already seen that our intuitions about these expressions and sentences are in accord with these predictions, this first approximation to a semantic theory is confirmed. It is to be noted also that, simple as this first approximation is, it is completely successful in the cases where *all* reductionist approaches considered here fail.

It is natural to take this first approximation as a reconstruction within contemporary linguistic theory of Frege's conception of sense. The first approximation formalizes senses and adds explicit definitions of what are plausibly Frege's notions of meaningfulness, ambiguity, and synonymy. Of course, to do full justice to Frege's conception, this first approximation would have to be fleshed out in various ways. The most important would be to add functions from sets of numerals representing the senses of words to sets of numerals representing the senses of syntactically complex expressions and sentences as a reconstruction of Frege's idea that the sense of a sentence is built up from the senses of its syntactic parts.[9] But my interest in Frege's conception of sense is not with it for its own sake but as an illustration of a position that embodies the limitations of the first approximation in a very direct way. Frege's conception of sense is almost universally taken as the very model of a sense theory. I wish to show not only that his conception is only one among other models but that, in embodying the limitations of this first approximation, it is a very poor model.

The easiest way to see these limitations with Frege's conception is to contrast Frege's characterization of sense with my characterization. Frege characterizes sense nontheoretically and derivatively in terms of reference: sense, for Frege, is simply the mode of presentation of reference ([14], p. 57). Frege's technical interest in sense is not with it for its own sake, that is, with the construction of a theory of sense, but with sense as a means to attaining the end of a theory of reference. Aside from remarks here and there about compositionality and the like, the only role that sense plays in Frege's theorizing is to save his theory of reference from problems it would otherwise have to face unaided. Thus, Frege's concern with the technical use of *Sinn* is confined to resolving the paradox of identity and explaining the reference of expressions in opaque contexts.

In contrast, my characterization of sense is theoretical and nonderivative. I take senses to be whatever it will be necessary to take them to be to explain the meaningfulness, ambiguity, synonymy, and *all* the other semantic properties and relations. In short, a sense is whatever the semantic theory that best saves the semantic phenomena of language says it is. One clear difference

between this characterization and Frege's is that Frege's leads him to claim that proper names are meaningful. Since they present their referent(s) in some one or another specific way—for example, "Aristotle" might present Aristotle as Plato's most famous pupil or as Alexander the Great's most famous teacher[10]—they have sense. Frege is led to say that two people whose modes of presenting Aristotle differ, at least in this respect, speak different languages ([14], p. 58). My characterization leads me to the opposite conclusion, as my earlier remarks on the contrast between proper and common nouns have shown.

My theoretical characterization requires a putative theory of sense, such as our first approximation, to be extended whenever a semantic property or relation cannot be explained with existing notational and definitional apparatus. In this vein, consider the semantic properties and relations of superordination, analyticity, and analytic entailment, as illustrated, respectively, by:

(11) dwelling/cottage
(12) Southpaws are left-handed pitchers
(13) Freud had a nightmare
Freud had a dream.

These semantic properties and relations are basically different from meaningfulness, ambiguity, and synonymy. Superordination, analyticity, and analytic entailment are *nonexpressional* semantic properties and relations in virtue of applying to senses. One can speak of expressions and sentences also as superordinates, analytic, and analytic entailments but not *simpliciter*, only as superordinates on a sense, analytic on a sense, and analytic entailments on a sense. In contrast, the semantic properties and relations we have been considering up to this point are *expressional*. They apply to linguistic constructions. It is absurd to speak of a sense as meaningful or ambiguous. Expressional semantic properties and relations are simply a count of the number of senses correlated with a linguistic construction in the language. The definitions (D_6), (D_7), and (D_8), respectively, say that meaninglessness is a count of zero senses and meaningfulness a count of at least one, ambiguity a count of two or more, and synonymy a count of the same sense more than once.

The problem for our first approximation is that definition in the case of nonexpressional semantic properties and relations has to refer to the separate parts of the structure of senses differentially but its notation scheme, being adapted for a count of whole senses, does not reflect sense structure. To define the superordination relation illustrated in (11), it is necessary to have a notation scheme that reflects the decomposition of the sense of "cottage" into the two concepts of a dwelling and of a certain kind of dwelling (different from, say, a dormatory or a skyscraper). Without such a decompositional scheme, it is not possible to formulate the defining condition that the sense of the subordinate expression is a proper part of the sense of the superordinate expression. Similarly, the notational and definitional apparatus in our first approximation cannot reveal the details of sense structure necessary

to define analyticity and analytic entailment. Thus, we have to go beyond a notational scheme consisting, in effect, of numericals functioning as bare names of senses. We require a scheme that describes the structure of senses in the decompositional way that chemical diagrams describe the molecular structure of compounds. Our first approximation is inadequate because it only provides names for senses taken as unanalyzed wholes.

In going beyond this first approximation, we are going beyond both Fregean and standard predicate calculi theories of the logical structure of natural languages (for Fregean theories, see [7]). Such theories represent the so-called "extralogical" vocabulary of a language–the entire stock of its nouns, verbs, adjectives, etc.–with such names. That is, their notation schemes for such vocabulary consists exclusively of symbols I will call *designations*: symbols, functioning as individual or predicate constants, that do not represent the internal structure of senses they name because their orthographic form serves only to determine the assignment of symbol tokens to symbol types (see [29]).

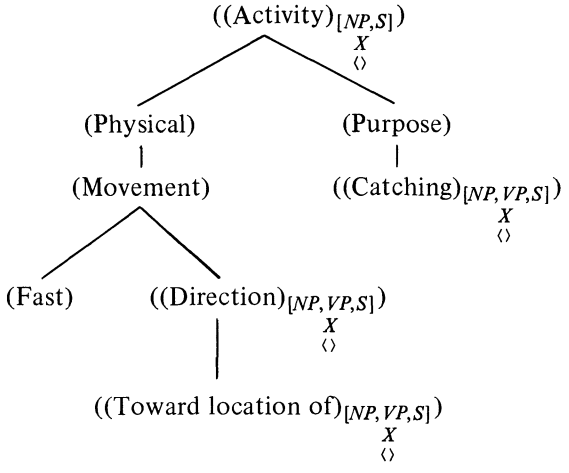
Frege does not seem to have recognized the need to replace such schemes with ones whose symbols represent sense structure, although, of course, he recognized complex concepts (in his sense), whose component concepts, which he calls "marks", are contrasted with properties of concepts because marks are properties of the objects falling under them (see [15], pp. 42-55). But a concept is the reference of a linguistic predicate for Frege, not its sense; furthermore, even if one were to read Frege's account of concept-complexity into his notion of sense, the result would be an account of sense-complexity in which the component senses of an expression are restricted to the senses of its syntactically distinguishable subexpressions. (Frege's examples are all of this sort, e.g., [15], pp. 51-52.) Such composition of senses can be handled within the notation scheme of our first approximation by assigning syntactically complex expressions a sequence of sets of numerals whose members are assigned to their syntactic subexpressions. *The significant point is that there is here no general notion of composition, no notion that applies to structure within the senses of primitives like "cottage", "southpaw", and "nightmare".* A further point is that, even if Frege were supplied with a general notion of sense composition, he would have no place to use it insofar as his rejection of the Kantian notions of analyticity and analytic entailment in favor of his much broader notions of logical truth and logical implications rules out definitions using the containment relation for senses generally (see [16], pp. 2-5).

Going beyond our first approximation, then, requires at least replacement of a vocabulary of designations by a vocabulary of what I shall call *descriptions*, that is, symbols whose formal structure describes the structure of the objects they symbolize. The special class of descriptions needed for semantic theory is composed of symbols that I have called "semantic markers" (see [25] and [28]). Semantic markers decompose senses into their component concepts and the relations between them. They are called "semantic markers" because they mark the semantic structure of senses in the manner that phrase markers mark the phrase structure of syntactic constructions.

The formal structure of a semantic marker is a representation of the

conceptual structure of a sense under certain conventions. Current theory takes the formal structure of semantic markers to be trees with labeled nodes. The branching reflects the complexity of the conceptual composition, the labels identify the conceptual components, and the assignment of labels to nodes specifies the logical relations that compose the complex sense out of its components. These conventions may be illustrated with the semantic marker (14) taken as the semantic representation of the sense of "chase".

(14)



The label on the topmost or root node specifies the category of the concept represented: (14) specifies that the concept of a chase belongs to the category of activity concepts. The labeled branching under the topmost node provides qualifications that distinguish the concept of a chase from other activity concepts. The first subbranch on the left distinguishes it from all concepts of mental activity, such as that of thinking, remembering, planning, and so on. The next longer subbranch on the leftmost full branch adds a further qualification that distinguishes the concept of a chase from that of physical activities not involving change of position like that of doing push ups. The next longer subbranch distinguishes it from concepts of physical movement which have no inherent speed specification like that of walking and from concepts of physical movement which are specified as slow like that of creeping.

The occurrences of the variable "*X*" with brackets above and angles below mark positions in a semantic marker at which other semantic markers may be embedded. The notation scheme of semantic markers must provide a set of semantic representations suitable for describing the meaning of every sentence in a language. Since the compositional mechanisms in a language compose the meaning of syntactically complex constructions from the meanings of their parts and since the syntactic mechanisms construct sentences with no limit on their complexity, semantic representations must be built up recursively in a way that reflects the compositional process over the infinite class of sentences. Hence, the notation scheme in semantic theory must contain primitive semantic markers and recursive means of constructing an infinite set of nonprimitive semantic markers out of the primitive semantic markers. One such recursive means is a *projection rule* which substitutes semantic

representations for occurrences of variables like those in (14) and a *dictionary* which provides the initial semantic representations for such operations.

The dictionary is a list of the morphemes (or syntactic primitives of the language) each of which is paired with a set of semantic representations (describing its senses as a lexical item).[11] The first step in accounting for the compositional meaning of a sentence is to correlate semantic representations from the dictionary with occurrences of morphemes in a syntactic description of the sentence. The next and further steps involve projection rule substitutions of semantic representations for occurrences of categorized variables. The brackets above a categorized variable determine the range of potential values for such substitution. The syntactic symbols in the brackets, called grammatical functions (see [5], pp. 68-74), specify the semantic representations which may be used in substitutions. The grammatical functions pick out semantic representations in terms of their assignment to constituents with those functions. For example, the function [*NP*, *S*] specifies that the variables over which it appears have as values semantic representations that are assigned to subjects of sentences, and [*NP*, *VP*, *S*] specifies that the variables over which it appears have as values semantic representations that are assigned to direct objects of sentences. In the case of:

(15) The police chased the demonstrators

the projection rule would substitute the semantic representation assigned to "the police" for the highest variable occurrence in (14) and the semantic representation assigned to "the demonstrators" for all the other variable occurrences in (14).

The angles below a categorized variable contain a constraint on the substitution of semantic representations. Unlike (15), in the case of:

(16) Truth chased falsehood

we do not want the semantic representations of the subject and direct object to combine. If they do, the sentence (16) will receive a semantic representation, and in accord with (D_6), it will be marked as fully meaningful in the language. Thus, constraints are required to block the assignment of semantic representations to constituents like (16) that are semantically deviant. Such constraints take the form of conditions on the semantic marker content of semantic representations that are potential values of a variable. For example, the concept of a chase is one in which the agent and the recipient of the action are inherently spatio-temporal particulars, and, hence, when abstractions like truth and falsehood are expressed by the senses of the subject and direct object, there is no sense of the whole sentence deriving from their combination.[12] Accordingly, we want to construct the semantic marker for "chase" with conditions that restrict substitution to just those semantic representations of subjects and objects that contain the semantic marker '(Physical)'. Such conditions would block substitution of the semantic representations of "truth" and "falsehood" in the case of (16), but allow substitution of the semantic representations of "the police" and "the demonstrators", since the latter contain the semantic marker '(Physical)' to account, *inter alia*, for the

meaningfulness of sentences like "The fat policeman jumped on and crushed the demonstrator".

Such conditions on substitution, called *selection restrictions* (see [25], pp. 89-116), not only make it possible to distinguish fully meaningful sentences like (15) from deviant sentences like (16), but, in so doing, they provide a general filtering mechanism for blocking semantic combinations that makes it possible to account for the actual degree of ambiguity of sentences under (D_6)-(D_8'). In actuality, the number of senses of morphemes runs, on the average, between five to ten, while the number of senses possible for, say, a twenty-word sentence, runs, on the average, well into the hundreds. Since the actual degree of ambiguity of a sentence is far greater than the degree on sheer combinatorial grounds there has to be severe filtering at work in the formation of sentence meanings from the meanings of their parts. To see that selection restrictions meet this need, note that semantic deviance is just the extreme case where the filtering mechanism spares nothing: meaninglessness is just zero degree of ambiguity. Thus, the same filtering mechanism, operating in less extreme cases spares some combinations and not others, giving rise to senses with degree of ambiguity from 1 to n. Hence, if the dictionary is set up properly, the projection rule will assign sets of semantic representations to sentences in such a way that (D_6)-(D_8') predict the meaningful ones, the meaningless ones, the ambiguous ones, and the synonymous ones.

We come now to the question of how a notation scheme consisting of semantic markers enables us to define the further semantic properties and relations of superordination, analyticity, and analytic entailment. Frege suggests that his rejection of the Kantian notion of analyticity in favor of the broader notion of logical truth is justified by two shortcomings of the Kantian notion of the containment of a predicate concept in the subject concept, namely, the metaphorical character of containment and the restriction of analyticity to subject-predicate sentences (see [16], pp. 99-102). These grounds for abandoning Kant's notion have been echoed down to the present. Quine, for example, uses both to motivate his taking analyticity to be, potentially, a class of logical truths ([34], p. 21). But neither Frege nor his followers have established that the replacement of the containment notion of analyticity with the notion of truths that follow from laws of logic plus definitions (see [16], pp. 3-4) does not simply replace one form of a priori truth with a different one having only a temporary advantage in technical explication.

What they need to show, but have not, is that their notion of logical truth is a better formulation of the *same* notion or replaces an *inadequate* notion. By themselves, the cited shortcomings mean nothing. Instead of taking them as deep difficulties calling for abandonment of the Kantian notion, these shortcomings might just as well be taken as nothing more than temporary features of the sort that one naturally expects in any as yet unexplicated notion.

Therefore, the task for us is to show that both shortcomings are just such temporary features of the Kantian notion and disappear once this notion is explicated within semantic theory, and that the notion of analyticity resulting from such an explication is both a legitimate and distinct form of a priori truth. It is clear that the first shortcoming disappears if the second

does, since the availability of formal descriptions of sense structure in semantic theory enables us to replace the metaphor of containment with a precise inclusion relation, provided a broad enough one can be stated. Thus, we have to show that descriptions in the form of semantic markers like (14) enable us to construct a definition of analyticity that covers analytic non-subject-predicate sentences as well as analytic subject-predicate sentences.

What are the analytic non-subject-predicate sentences that need to be covered? Frege and his followers stack the deck in favor of their proposal to replace analyticity with logical truth by citing cases of logical truths like "If John is a bachelor, then John is a bachelor or John is rich". But, from a common sense standpoint, sentences like:

(17) John walks with those with whom he strolls

are the genuinely natural counterparts of analytic subject-predicate sentences. This observation puts the question in a new form: How do we formulate a condition (in terms of the formal relations in semantic representations) that captures analytic transitive verb sentences like (17) as well as analytic subject-predicate sentences like (12)?

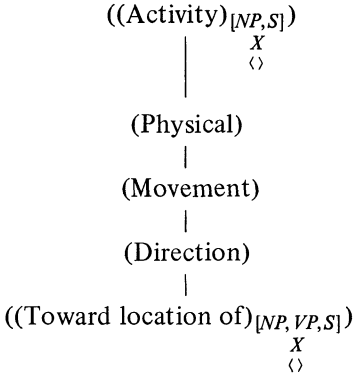
Once we have an example like (17) in front of us, we can see that the concept containment in the Kantian notion is not essentially a matter of the subject-predicate form of a sentence like (12). Rather what is essential to the Kantian notion, what makes (17) an intuitively natural counterpart of (12), is that the components of the meaning of an analytic sentence which pick out the object(s) it is about already contain the component of its meaning that expresses the attribution it makes to these object(s). In the case of a subject-predicate sentence, the component of the meaning which picks out the object(s) the sentence is about is the sense of the subject; in the case of a transitive verb sentence, there is more than one component of the meaning which picks out the object(s) the sentence is about, the sense of the subject, direct object, etc. Thus, Kant, having begun with a special case of analyticity, failed to generalize it, whereas Frege and Kant's other critics, in their zeal to generalize it, failed to take account of true character of the special case.

Because semantic markers mirror the relations between components of sense structures in the formal relations between their parts, it is a straightforward matter to state such a general condition within semantic theory. Without considering all of the details of this explication, we can convey the basic idea as follows. The concept expressing the attribution of a sentence is represented by a skeletal semantic marker like (14) whose occurrences of categorized variables have not yet been replaced by semantic representations. The components of the meaning of the sentence that pick out the object(s) the sentence is about are represented by the semantic representations that substitute for these occurrences in the projection process. Since both the components picking out the domain of attribution and the components expressing the attribution are formalized in the above tree notation, the formal explication of our general condition for the analyticity of (a sense of) a sentence is that the semantic representation into which such substitutions are made be a subtree of a semantic representation that is substituted into it.[13]

The formalization of the conditions for superordination and analytic

entailment are essentially the same. The sense of one expression is superordinate to the sense of another just in case the semantic representation of the latter is a subtree of the semantic representation of the former; and the sense of one sentence analytically entails the sense of another just in case the semantic representation of the latter is a subtree of the semantic representation of the former. To provide an illustration, suppose that the semantic representation for the verb "follow" is:

(18)



Then, since (18) is a subtree of (14), the sentence (15) is predicted, correctly, as analytically entailing the sentence:

(19) The police followed the demonstrators

on the appropriate senses of these sentences.

Having shown how to overcome both of the shortcomings in Kant's notion of analyticity, we must now show that the notion explicated constitutes a *legitimate* form of a priori truth and also that it is *distinct* from logical truth and logical implication.

Establishing legitimacy requires showing that the analyticity of a sentence and the analytic entailment of one sentence by another can be determined without knowledge of states of affairs in the world, and further, that the analyticity of a sentence is a sufficient condition for the statement it makes to be true (in all worlds where the object(s) it is about exists) and that the analytic entailment of one sentence by another is a sufficient condition for the entailed sentence to be true whenever the entailing sentence is.

Given how analyticity and analytic entailment have been explicated, determining that a sentence is analytic or that one sentence analytically entails another is entirely a matter of reasoning from intuitive judgments about the semantic properties and relations of their constituents. Since such judgments concern only the structure of senses and the intragrammatical relations between senses and constituents and not relations between language and the world, the determination of analyticity and analytic entailment is a priori. Furthermore, since what is determined in determining analyticity is that the conditions expressing what the sentence is about include the condition for the attribution to be true, and since what is determined in determining analytic entailment is that the truth conditions of the entailed sentence are included in the truth conditions of the entailing sentence, analyticity is sufficient

for truth (in all worlds where the object(s) the sentence is about exists) and analytic entailment is sufficient for the truth of the entailed sentence in all worlds where the entailing sentence is true.

Establishing distinctness requires showing that there is a genuinely logical difference between analytic truths and logical truths and between analytic entailments and logical entailments. A complete argument for distinctness would first establish the proof-theoretic difference and then exhibit the corresponding model-theoretic difference. The task of exhibiting the model-theoretic difference is beyond the scope of this discussion and will be presented elsewhere.[14]

Frege's notions of logical truth and logical implication provide an account of inferences like (15) to (19) only when the underlying logic is supplemented with an appropriate set of meaning postulates (see [2]). In order to provide such an account, Carnap adapted the postulational treatment of logical implication in predicate calculus to the broader class of implications based on the extralogical vocabulary by introducing meaning postulates in analogy to logical postulates. Predicate calculi were thus expanded, first, by the addition of an infinite list of designations which serve as individual and predicate constants in the regimentation of sentences from natural languages containing extralogical vocabulary, and second, by the addition of a finite list of new postulates, expressed as formulas of first-order logic but employing such designations. Such postulates restrict the class of admissible models for a calculus in the same way as the logical postulates. An inference like (15) to (19) is accounted for in terms of a meaning postulate like:

(20) $(x)(y)(C_{x,y} \supset F_{x,y})$,

where C and F are the designations that regiment "chase" and "follow", respectively.

The claim that analytic entailment is a special case of logical implication in Frege's sense can be put as the claim that this meaning postulate account of inferences like (15) to (19) is correct in taking analytic entailments to be continuous with first-order implications generally. But there is the striking discontinuity that a first-order implication has to be justified on the basis of principles of deduction which sanction the step from premiss(es) to conclusion (as one that never leads from truth to falsehood), whereas an analytic entailment requires no such justification. Such a justification is required in the case of first-order implications by virtue of the character of the truth conditions associated with the premiss(es) and conclusion in them: the truth conditions of the conclusion are not explicit parts of the truth conditions of the premisses. Thus, the satisfaction of the truth conditions of the conclusion must be shown to obtain whenever the truth conditions of the premiss(es) are satisfied by subsuming the relation between them under a principle, like modus ponens, which exhibits the satisfaction of the former as a necessary consequence of the satisfaction of the latter. Analytic entailments, in contrast, require no such justification because the truth conditions of the conclusion *are* an explicit part of the truth conditions of the premiss. Thus, there is no deductive step from premiss to conclusion that makes it necessary to appeal to a principle of logic. Satisfaction of the truth conditions

of the premiss is in itself satisfaction of the truth conditions of the conclusion: the latter is not merely a necessary consequence of the former.

This difference can be illustrated using Lewis Carroll's philosophical fable about Achilles and the Tortoise ([3], pp. 1225-1229). The Tortoise's strategy is to challenge Achilles each time he tries to infer a conclusion on the grounds that the principle sanctioning the inference has not yet been written down in Achilles' book of accepted truths. The case in the fable is a Euclidean proof, but we may illustrate the Tortoise's strategy in the case at hand. Let (15) be (A), (20) be (B), and (19) be the elusive conclusion (Z). Achilles reasons that if (A) and (B) are true, then (by modus ponens) so is (Z). The Tortoise points out that the principle invoked to sanction the inference has not yet been written down in the book. Achilles dutifully writes it down as (C), thinking that he has done what is required. As Lewis Carroll has Achilles remark at this point, referring to the conclusion,

> "You should call it D, not Z. It comes *next* to the other three. If you accept A and B and C, you *must* accept Z."

But, again, the Tortoise resorts to his strategy. Lewis Carroll has him query, "And why *must* I?" To which Achilles, not yet realizing his situation, replies,

> "Because it follows *logically* from them. If A and B and C are true, Z *must* be true. You don't dispute *that* I imagine?"
>
> "If A and B and C are true, Z *must* be true . . . that's *another* hypothetical, isn't it? And, if I failed to see its truth, I might accept A and B and C, and *still* not accept Z, mightn't I?"

Lewis Carroll concludes this dialogue

> "You might," the candid hero admitted; "though such obtuseness would certainly be phenomenal. Still, the event is *possible*. So, I must ask you to grant one more hypothetical."
>
> "Very good. I'm willing to grant it, as soon as you've written it down."

Now, I wish to compare this with a fable of my own. Again, (15) is (A) and (19) is (Z). But in my fable Achilles had once time traveled to the twentieth century where he acquired the rudiments of semantic theory. Reviewing the conditions of the contest proposed by the Tortoise, my Achilles says,

> I'm to force you to accept Z, am I? And your present position is that you accept A. I may write A down in my book. Well, in that case force isn't necessary. For, you see, having written down A, Z, too, is written down! When you asked in our previous contest what else I had written down in my book, all I could reply was, "Only a few memoranda of the battles in which I have distinguished myself." But, now, after my visit to the future, I see that I can reply, "Z". Surely you won't balk because I've written A and Z in markerese which won't be invented for some two thousand years.

What is the nature of the difference between treating an inference like (15) to (19) as an analytic entailment and treating it as first-order implication

in an applied predicate calculus with meaning postulates like (20)? Treating an inference as an analytic entailment explains its validity in terms of the meanings of the sentences: the use of descriptions exhibits the meaning of the premiss and conclusion in fine enough detail to show that the truth conditions of the latter are part of the truth conditions of the former. Treating an inference as a first-order deduction with a meaning postulate as a further premiss makes no claim about a semantic source of its validity. Being constructed with designations, meaning postulates do not describe meaning at all. They state an extensional condition but not as reflecting meaning relations. Such a condition, being a restriction on admissible models, may reflect any relations that hold in all possible worlds, for example, mathematical relations like that between being the number two and being the only even prime or metaphysical relations like being an event and having a cause. The name "meaning postulate" is a misnomer! Such statements express a relation between the extensions of the predicates designations that they contain, but they do not make an explanatory link between the extensional relation and the meaning of the premiss and conclusion. Indeed, semantic theories containing a notation scheme with only designations can make no such link.

In effect, we have just presented one example of how a theory of meaning, because of the greater explanatory power afforded by a notation of descriptions (in the form of semantic markers), can throw light on problems in the philosophy of logic and language. Another problem on which this theory sheds light is the problem of deciding what the categories of our conceptual system are. It has been a persistent complaint about systems of categories like Aristotle's that they retain an arbitrary character as long as there is no rational basis for judging one way of arranging the objects of thought as preferable to others. Kant complained in the *Prolegomena* that in the case of Aristotle's categories

> . . . we are not able to give a reason why each language has just this and no other formal constitution, and still less why any precise number of such formal determinations in general, neither more nor less, can be found in it. . . . This rhapsody must be considered (and commended) as a mere hint for future inquirers, not as an ideal developed according to rule; and hence it has, in the present more advanced state of philosophy, been rejected as quite useless.

Kant's own explanation of why each language has just the formal constitution of the Kantian categories and no other was that these categories are the conceptual means by which our understanding imposes order on the impressions we receive of things in themselves. But this rationale is too deeply embedded in his theory of mind to be an acceptable basis for comparing the ways different philosophical theories arrange the objects of thought, one which enables us to evaluate the arrangements and also the theories in which they appear.

Such a basis is available within a semantic theory of the kind we have described (see the discussion in [24], pp. 224-239). Recall that a semantic theory is a metatheory consisting of a notation scheme and definitions and that the theories in question are semantic components of grammars (dictionaries, pairings of morphemes of a language with semantic representations

of their senses). Lexical semantic representations will have to contain a distinct semantic marker for each component concept into which the sense it represents breaks down. For example, the sense of the noun "chair" will have to be assigned a semantic representation containing the semantic markers '(Object)', '(Physical)', '(Artifact)', and '(Furniture)', among others. Such marker content is necessary to predict the analyticity of sentences like "Chairs are physical objects", "Chairs are artifacts", and "Chairs are furniture". Similarly, the dictionary representation for related words like "table", "bench", "stool", etc., will have to contain these same four semantic markers. Further, on the same considerations, the dictionary representation of morphemes like "truck", "automobile", "bus", etc., will contain the semantic markers '(Object)' '(Physical)', '(Artifact)', and '(Vehicle)'. Thus, we find the same pattern, in which the semantic markers '(Object)' and '(Physical)' are found whenever the semantic marker '(Artifact)' occurs, repeated in one dictionary entry after another. This, moreover, is only a small fraction of the cases where such regularities appear: they exist on an enormous scale across the various entries of a dictionary and over the range of dictionaries for all natural languages.

From a methodological viewpoint, such regularities, since they are avoidable without loss in predictive power, constitute redundancy in the formulation of dictionaries. Dictionaries having such redundant semantic markers like '(Object)' and '(Physical)' are stating facts about the occurrence of semantic markers separately for each semantic representation in which '(Artifact)' occurs when these facts could be stated just once. If, furthermore, they are stated just once, then the dictionary rule which states them will capture a generalization about lexical structure that the uneconomical dictionaries fail to capture.

Hence, methodological considerations force us to simplify the fomulation of dictionaries by introducing rules that express such generalizations about lexical structure. In previous discussions ([24], pp. 230-233, and [25], pp. 44-46), we have, accordingly, expanded semantic theory to contain *redundancy rules*, rules of the form '$(M_i) \rightarrow (M_j)$' which are interpreted to say that if the semantic marker '(M_i)' appears in a semantic representation, then the semantic marker '(M_j)' also appears in it.

Now, putting aside methodological considerations, let us look at the generalizations that redundancy rules express. It is clear that the generalization expressed by the redundancy rule '(Artifact) → (Object)' holds because the concept of an artifact is superordinate to the concept of an object. That is, the concept of an artifact arises from the more abstract concept of an object when this concept is qualified as coming into existence as a product designed to serve a function. Since individual redundancy rules express superordination relations between two concepts, chains of redundancy rules of the form:

$$(21)\ (M_1) \rightarrow (M_2), (M_2) \rightarrow (M_3), \ldots, (M_{n-1}) \rightarrow (M_n)$$

express a transitive 'less abstract than' relation from 1 to n. More importantly, if (21) is a maximal chain with respect to a complete set of redundancy rules for semantic theory, then (M_n) is maximally abstract. If, therefore, there are

k maximal chains with respect to a complete set of redundancy rules, the set of terminal members of these k chains, viz., (M_{n_1}), (M_{n_2}), . . ., (M_{n_k}), are the maximally abstract semantic markers.

The maximally abstract logical divisions in the sense system of natural language–the concepts represented by the semantic markers (M_{n_1}), (M_{n_2}), . . ., (M_{n_k})–are a plausible explication of the notion 'categories of language'. The explication clearly satisfies our original demand for a nonarbitrary way of arriving at a set of categories that plays no favorites among philosophical theories, for the considerations that motivate the adoption of redundancy rules, apart from the predictive adequacy of the semantic representations, are just a matter of the simplicity of theories.

My next example of a philosophical problem which is illuminated by the theory of meaning is that of stating the conditions for inferences by substitution into opaque contexts. Simple referential theories of meaning like Russell's and Davidson's stand no chance of capturing just the cases where such inferences go through, namely, cases like the inference from (22) to (23):

(22) Simon believes that his daughter will marry an adult unmarried human male

(23) Simon believes that his daughter will marry a bachelor

and so their supporters have denied the legitimacy of such inferences, usually appealing to Quinian skepticism on meaning to criticize the synonymy relation on which the inferences depend (see [35], pp. 141-156, 206-221). This shortcoming of simple referential theories has prompted possible world semanticists to claim that their theory enjoys an advantage because it does not have to flatly claim that there are no valid inferences by substitution into opaque contexts (see [21], pp. 87-111). But is clear that, since co-extensitivity in all possible worlds does not adequately reconstruct synonymy, possible world semantics cannot completely capture this class of inferences. Possible world semantics cannot systematically distinguish between cases like the valid inference (22) to (23) and cases like the invalid inference (24) to (25):

(24) Simon believes that Ali has seven wives

(25) Simon believes that Ali has a number of wives equal to the even prime plus the square root of twenty five.

It is thought by those who reject both Quine's criticisms of meaning and the possible worlds semanticist's reconstruction of these inferences that Frege's account of such inference solves all our problems. But this thought is mistaken. The same difficulties we found above with Frege's theory of sense prevent it from providing a satisfactory account of inference by substitution into opaque contexts. I will first explain why this is and then show how a semantic theory can provide a satisfactory account.

Frege thought that if he introduced senses to be the reference of expressions and sentences in opaque contexts, he could characterize both the conditions for inferences by substitution into such contexts and the conditions for inference by substitution into transparent contexts as coreferentiality of the expression or sentence substituted and the expression or sentence substituted for ([14], pp. 64f); in effect, if we substitute an expression or

sentence for one that is synonymous with it, then the sentence resulting from the substitution follows validly from the sentence into which the substitution was made. This principle nicely distinguishes cases of valid inference like (22) to (23) from cases of invalid inferences like (24) to (25), and because of the vast improvement that this represents in comparison with previous theories, it has been concluded that Frege's account is adequate in general.

But Frege's principle fails to capture an indefinitely large class of valid inferences that a sense theorist is committed to accepting (see [25], pp. 265-267). For example, assuming it is valid to infer (23) from (22), it must also be valid to infer (27) from (26):

(26) Simon believes that his daughter will marry a bachelor.
(27) Simon believes that his daughter will marry someone unmarried.

Yet Frege's principle does not apply to cases where the expression or sentence substituted and the expression or sentence substituted for it are nonsynonymous. Moreover, it is clear that the relations required to handle such cases are superordination and analytic entailment, and, as we have already shown, Frege does not have the apparatus required to handle these relations.

We have the apparatus and so can formulate the condition for such inferences, but some care is needed in formulating it. If we formulate it in the way Frege formulates his, as a condition on the expression or sentence substituted and the expression or sentence substituted for, we shall run into trouble with "translucent verbs" ([25], pp. 275-280). Even though the embedded sentence in:

(28) Simon doubts that his daughter will marry a bachelor

analytically entails the embedded sentence in:

(29) Simon doubts that his daughter will marry a male

the inference from (28) to (29) is not valid. Simon's suspicions may only concern bigamy.

Translucent verbs have meanings that operate at the verb-phrase level in the compositional process whereby the meaning of the full sentence is formed from the meanings of its parts to change the object of the propositional attitude from the proposition expressed by the sentence occurring in the context of the verb. Thus, a Fregean condition for inference by substitution fails in the case of analytically entailed sentences in translucent contexts because the object of the propositional attitude (expressed by the verb creating the context) is determined at a level higher than that at which the condition applies. There is a negative operator in the meaning of the translucent verb "doubts" that changes the object of doubt from the sense of the sentence substituted to something like the proposition that Simon questions whether his daughter will marry an adult or questions whether his daughter will marry a single person or questions whether his daughter will marry a male. Having neglected the details of the compositional process, Frege formulates his condition for such substitutional inferences as a condition on the pair of the

sentence substituted and the sentence it substitutes for. The condition thus applies too early in the compositional process to consider all the semantic relations within the sentences that determine whether one follows validly from the other.

To provide a better account than Frege's, we have to formulate our condition for these inferences at a high enough point in the compositional process to guarantee that the condition will miss no semantic operations on which implications of a sentence depend. The surest way to do this is to formulate the condition to apply to full sentences. This, moreover, has the further advantage that it is now unnecessary to construct a special substitution principle. The definition of analytic entailment, which applies to full sentences, already provides the principle we require. The definition not only handles cases like the inference from (22) to (23) which Frege's principle handles (but the principles of referential theories do not) and cases like the inference from (26) to (27) which Frege's principle does not handle, but, in addition, the definition correctly treats cases of invalid inferences like that from (28) to (29). Because the definition of analytic entailment applies to the semantic representations of full sentences, it applies after all compositional operations have taken place, and, hence, in a case like the invalid inference from (28) to (29), it applies to semantic representations that do not meet the condition for analytic entailment.

The ability of semantic theory to adequately state the conditions for inference by substitution into opaque contexts is not just a matter of following Frege in recognizing sense as well as reference but a matter of going beyond Frege in introducing a notation of descriptions which makes it possible to treat the aspects of sense structure that are involved in all phases of the compositional formation of sentence meaning.

The last problem I want to consider illustrates another application of the increased explanatory power offered by a notation which enables semantic representation to describe the structure of senses. The problem concerns a general misrepresentation of modification in natural language which has resulted from relying solely on the representational apparatus in reductionist conceptions of meaning. The problem arises even in such otherwise sound accounts of modification as Davidson's account of the logic of action sentences (see [11]). Davidson's treatment of the logical form of sentences like (30)-(33)

(30) Jones strolled into the bathroom at midnight.
(31) Jones strolled into the bathroom.
(32) Jones strolled.
(33) Jones walked.

enables him to account for inferences like (30) to (31) and (31) to (32), but not inferences such as (30) to (33), (31) to (33), and (32) to (33). Indeed, he explicitly exempts such inferences. Right at the beginning Davidson says:

> I am not concerned with the meaning analysis of logically simple expressions in so far as this goes beyond the question of logical form. I am not concerned with the meaning of 'deliberately' as opposed, perhaps, to 'voluntary'. ([11], pp. 105-106)

Now, it is more than a bit confusing for Davidson to say "not concerned with" insofar as his general theory, as we have seen, claims that, beyond a characterization for the language of the predicate 'is true', there is nothing "essential to the idea of meaning that remains to be captured". Clearly, Davidson's replacement of '*s* means *p*' form of analysis with the '"*s*" is true if, and only if, *p*' form and his commitment to standard logic as a complete theory of logical structure together *commit* him to excluding inferences like that from (32) to (33) from the class of logical entailments.

It is still more confusing when Davidson goes on to say:

> . . . we need not view the difference between 'Joe believes that there is life on Mars' and 'Joe knows that there is life on Mars' as a difference in logical form. That the second, but not the first, entails 'There is life on Mars' is plausibly a logical truth; but it is a truth that emerges only when we consider the meaning analysis of 'believes' and 'knows'. ([11], p. 106)

If "the meaning analysis of logically simple expressions . . . goes beyond the questions of logical form", how is it "plausibly a *logical* truth" (emphasis added) that the sentence "Joe knows that there is life on Mars" entails "There is life on Mars"? We are not helped out to be told next

> Admittedly there is something arbitrary in how much of logic to pin on logical form. But limits are set if our interest is in giving a coherent and constructive account of meaning: we must uncover enough structure to make it possible to state, for an arbitrary sentence, how its meaning depends on that structure, and we must not attribute more structure than such a theory of meaning can accommodate. ([11], p. 106)

Wherever the arbitrary line is drawn, either the meaning analysis of *some* "logically simple expressions" gives rise to logical truths, in which case such analysis does go beyond logical form in the sense Davidson is concerned with, or the meaning analysis of *no* "logically simple expression" gives rise to a logical truth, in which case the plausible logical truth that "If Joe knows that there is life on Mars, then there is life on Mars" cannot count as a logical truth on Davidson's view. Conceding such arbitrariness does not give one a license to have it both ways.

There is "something arbitrary in how much of logic to pin on logical form" for someone like Davidson. This arbitrariness is not, however, a fact that everyone has to face. It exists only for those who have bought the idea that the words of a language divide into a logical vocabulary and an extralogical vocabulary. For, by the only criterion of the logical that there is, namely, what is logical is that upon which valid argument rests, all the words of a language can be counted as logical vocabulary. Since there is no other criterion, making a logical/nonlogical distinction among the uniformly logical words of a language is just the familiar arbitrariness of a distinction without a difference.

Everyone will agree that they want a theory that "uncovers enough structure to make it possible to state, for an arbitrary sentence, how its meaning depends on that structure". The issue is what will count as *structure* and what as *structured*. Davidson, except for one or two confusing remarks,

counts only first-order structure as logical structure and only the logical particles as logical vocabulary. For him, the meanings of "*logically* simple expressions" (italics mine) like "deliberately" or "voluntary" do not contribute to logical form, while, for me, such expressions are *syntactically* simple but logically complex (in that aspects of the structure of their meaning support valid arguments).

Davidson must, therefore, treat inferences like (32) to (33) as nonlogical in the manner of an inference like (34) to (35):

(34) Brutus stabbed Caesar
(35) Brutus killed Caesar

Inferences like (32) to (33) have to be represented as depending on contingently true synthetic propositions in the way that the inference from (34) to (35) depends on (36).

(36) Brutus's stabbing of Caesar is Brutus's killing of Caesar.

But, although it is perfectly possible for Brutus to have stabbed Caesar without killing Caesar, it is perfectly *im*possible for Jones to stroll into the bathroom at midnight without walking, just as it is impossible for Jones to stroll into the bathroom at midnight without strolling. Hence, Davidson's classification of the inference from (30) to (33) with the nonvalid inference (34) to (35) is mistaken. And, likewise, the claim that the sentence corresponding to (36), namely,

(37) Jones's strolling is Jones's walking.

is merely a contingently true synthetic proposition like (36) is mistaken: (37) is a necessarily true analytic proposition.

The upshot, so far, is nothing terribly surprising: the analysis of verb-adverbial modification in Davidson's framework is arbitrarily limited, restricted not by the intrinsic boundaries of logical structure in natural language but by the shortcomings of the conception of analysis used to represent it. But something more surprising emerges if we view this arbitrariness from a slightly difference perspective.

Since "stroll" is synonymous with "walk leisurely", the qualification of the concept of Jones strolling that is brought about by the adverbial modifier "into the bathroom at midnight" is, logically speaking, no different than the qualification of the concept of Jones walking that is inherent in the concept of Jones walking in a leisurely way. The fact that the meaning of this latter concept is *un*reflected in syntactic structure in the case of "stroll" but is reflected in syntactic structure in the case of "walk leisurely" is logically irrelevant. Davidson's use of "logically" in his phrase "logically simple expressions" ("deliberately", "know", etc.), which wrongly identifies syntactic structure with logical structure, obscures this important point. Similarly, it is logically irrelevant that a semantically complex sense is not reflected in the syntax of the primitive "stroll" but is in the syntax of the verb phrase of (30). Hence, the arbitrary limitation divides semantically equivalent cases in a way that mistakes syntactic reflection for logical substance. Accordingly,

to treat like cases in a like manner, we have to drop the arbitrary limitation and treat inferences that depend on the meaning of syntactically complex expressions and inferences that depend on the meaning of syntactically simple expressions in the same way.

Now, Davidson treats action sentence inferences like

(38) Brutus killed Caesar with a knife.

(38) to (35) as first-order inferences by, first, hypostatizing events, and second, taking the logical form of the conclusion (35) to assert the existence of a killing of Caesar by Brutus, that is, to be:

(39) $(\exists x)(\text{Kills}_{x,\ \text{Caesar},\ \text{Brutus}})$

and, correspondingly, the logical form of the premiss (38) to assert that the killing is accomplished with a knife, that is, to be:

(40) $(\exists x)((\text{Kills}_{x,\ \text{Caesar},\ \text{Brutus}}) \,\&\, (\exists y)(\text{With}_{x,y} \,\&\, \text{Knife}_{y}))$.

Accordingly, to handle the inference from (32) to (33) in this manner, Davidson will have to introduce the meaning postulate:

(41) $(\exists x)((\text{Stroll}_{x,\ \text{Jones}}) \,\&\, (\exists y)((\text{Walk}_{y,\ \text{Jones}}) \,\&\, (x = y)))$

But, as we have already seen, this way of handling analytic entailments is inadequate.

Thus, it seems that not only Davidson's neglect of inferences like (32) to (33) but his entire treatment of the logical form of action sentences is wrong. For if we are right about the continuity of semantic structure from the decompositional meanings of primitives to the compositional meanings of syntactic complexes,[15] and further, are right in our earlier argument to show that inferences like that from (32) to (33) are fundamentally different from first-order implications based on meaning postulates (and are properly handled as analytic entailments), then even the cases that Davidson treats, namely, inferences like (30) to (31), (31) to (32), and (38) to (35), are improperly treated and should be handled as analytic entailments, too.

We may check this claim that it is wrong to take these inferences to be first-order inferences as follows. Davidson represents these inferences as depending essentially on an appeal to principles of logic for their justification, even when, as in the case of inferences like that from (38) to (35), such inferences are at their simplest, depending only on an appeal to simplification. If this is correct, then there ought be no way to prevent Lewis Carroll's Tortoise from frustrating Achilles's attempt to infer (Z) = (35) from (A) = (38). But the Tortoise can be frustrated here in exactly the same way in which my Achilles frustrates him in the case (Z) = (19), (A) = (15). If inferences like (30), (31), or (32) to (33) can be defended against the Tortoise on the grounds that there is no deductive gap to be bridged by a principle of logic—because, appearing as an explicit component of the truth conditions of the premiss, the truth condition of the conclusion is shown to be satisfied in the satisfaction of the truth condition of the premiss—then there is the same defense against the Tortoise in cases like (30) to (31), (31) to (32), or (38) to (35), which

differ only in that the critical aspects of their meaning is reflected more fully in their syntax.

The conclusion, then, is that Davidson's hypostatization of events and proposal to represent the logical form of inferences like (38) to (35) in the manner of (30) and (41) has to be replaced by a uniform treatment of such inferences with other analytic entailments.

The lines along which such a treatment can be given have already been suggested in the presentation of the notation of descriptions in terms of semantic markers and of the projection apparatus. A sentence like

(42) The police followed the demonstrators speedily and with the purpose of catching them.

has the same sense as the sentence (15), and hence, the senses of the adverbial modifiers "speedily" and "with the purpose of catching them" must combine compositionally with the sense of "follow" to form the sense of "chase". Since the sense of "chase" in question is represented in (14) and the sense of "follow" in question is represented in (18), the semantic representations of "speedily" and "with the purpose of catching them" must be something like:

(43) (Fast)

(44) (Purpose)
|
$((\text{Catching})_{[NP,VP,S]})$
X
$\langle\rangle$

and the projection apparatus must attach (43) and (44) to (18) so that the resulting tree is that in (14). The projection rule attaches the semantic representation of a modifier as a branch to the semantic representation of its head by connecting the topmost semantic marker in the representation of the modifier with the most immediate superordinate semantic marker in the semantic representation of the head, where (M_i) is superordinate to (M_j) in case there is a radundancy rule $(M_j) \rightarrow (M_i)$ or (M_i) appears in a later rule than (M_j) in a chain of the form (21).[16]

NOTES

1. See [9], especially p. 455. Something has to be said about Davidson's grounds for taking the crucial step from the 'means that' form of semantic analysis to the '"*s*" is true if, and only if, *p*' form. Davidson says two things. One is that this step enables us to escape being "enmeshed in the intension". Presumably, this is a cryptic reference to the difficulties Quine raises with intensionalism. If so, we shall turn to them later. The other thing Davidson says is that this step is the only way he knows of to deal with the difficulty that "we cannot account for even as much as the truth conditions of belief sentences and others containing intensional contexts on the basis of what we know of the meanings of the words in the belief" (pp. 453-455). But Davidson simply says this, giving no argument to back up the claim. In particular, he does not even consider the line of argument in [8]. I should make clear that, in observing that Davidson does not consider the line of response Church introduced, I do not mean to suggest that

I think Church's work on the topic is the last word or that it is without problems. I think it is the right line. See [25], pp. 267-280 and [32], pp. 138-142.

Davidson's disciples have done no better than Davidson in motivating this step. Typical is Foster's criticism of the 'means that' form of analysis (see [13], p. 6). Foster's first criticism amounts to nothing more than a bare claim to the effect that "it is hard to envisage an acceptable theory which handles this intensionality". But since there is no examination of attempts to construct an acceptable theory to back up this claim, the first criticism can be ignored. Foster's second criticism is that use of the form '*s* has the meaning . . . ' is circular. He says: "[insight into the nature of meaning] will be diminished if, to gain it, we have to take intensional idiom for granted, for to understand such locutions as 'means that' . . . requires an implicit grasp of the very concept of meaning which we hope to explicate". This criticism is of special interest in connection with our discussion of reductionism in semantics because it quite unabashedly reads reductionism into nonreductionist approaches to explaining what sentences mean. For unless reductionist goals are assumed, there is nothing wrong with presupposing "an implicit grasp of the very concept . . . we hope to explicate". No exception is taken to such a presupposition outside of semantics–in, say, syntax, phonology, or logic, no one would think it circular to use the '*x* has the syntax . . . ', '*x* has the phonological structure', or '*x* has the logical structure' form of analysis in syntax, phonology, or logic. Only if one supposes in advance that the insight one hopes to gain into the nature of meaning will come by a form of analysis that reduces meaning to something else will it be circular to presuppose an implicit grasp of meaning.

2. They are clearly different in meaning on the face of it. Note also that they make different semantic contributions to the sentences in which they occur as constituents: compare "Polar animals are white because snow is white" with "Polar animals are white because grass is green".

3. If they are relevant, we can, of course, exhibit the grotesqueness in cases where everyone is certain by employing arbitrary pairs of simple arithmetic truths in place of (7) and (8).

4. For further discussion of Davidson, including examples of the kind indicated in the text, see [26].

5. A semantic theory of language in general, accordingly, will only have to define the semantic properties and relations in natural languages on the basis of the representational apparatus that enables us to construct correct semantic theories of every natural language.

6. Hintikka claims that the " . . . whole concept of meaning (as distinguished from reference) is very unclear and usually hard to fathom". Well, if so, then the whole concept of possibility is equally unclear and hard to fathom. See [21], pp. 87-88. For further discussion, see footnote 42, p. 107 of [30].

7. Whether the definitions are given in this way or in terms of functions from possible worlds on to extensions of expressions and sentences in them does not matter in the present context.

8. Another form of skepticism about meaning has to be considered, namely, that initiated in Putnam [33]. See my rejoinder in [27] and [31].

9. Frege only pays lip service to such compositionality in that he makes no attempt to go beyond the "thought-building-block" metaphor and formally specify the architectural blueprints.

10. This feature of Frege's account of *Sinn* does not remove it as a case of what we have called a first approximation, since whether or not proper nouns have meaning is a question about the grammar of English and other languages rather than a question about the notation scheme or definitions.

11. Technically, lexical items are given in a notation called "features"; see [5].

12. To say that a sentence like (16) has no sense in the language is not to say that it has none in the use of language. Metaphorical uses of semantically deviant sentences to make meaningful statements, on my view, work in a manner similar to that described for conferring meaning on a meaningless expression like "grrr" (see above).

13. I have avoided stating the actual definitions because of the considerable complexity of notation and explanation that would be necessary. Earlier statements are found in [25], ch. 4; in simplified form in [29], p. 391. See [38] for a full statement.

14. What is required is a model theory for semantic representations that defines 'admissible model' in such a way that there exist admissible models on which the logical truths are untrue and the logical implications invalidated, while the analytic truths are true on all admissible models and the analytic entailments validated on all admissible models. See [38].

15. We might illustrate the point by observing that it is merely a historical accident that English does not contain a single verb meaning "kill with a knife" in the way it does contain a single verb meaning "kill with a gun", namely, "gun down".

16. Considering the importance of cases like "Jones is walking slowly but talking fast" for discussions of adverbials in Davidsonian and possible worlds frameworks, it is worthwhile saying something, even briefly, about how the decompositional approach handles them. On this approach, the sense of "walk" has a complex structure: the word expresses the concept of something moving at a moderate pace by placing one foot firmly on the ground before lifting the other. Similarly, the sense of "talk" has a complex structure: the word expresses the concept of moving one's mouth to make speech sounds to communicate one's ideas. The fact that these senses contain the concept of the movement enables the projection rule to attach the semantic representation of "slowly" at the appropriate point in the semantic representation of "walk" and to attach the semantic representation of "fast" at the appropriate point in the semantic representation of "talk". Since in the former case the movement is movement of the feet and in the latter case the movement is movement of the mouth, our account will represent the sentence as having a consistent sense in a particularly simple and natural way compared to other approaches. Also, the same decompositional analysis enables us to predict entailments of the sentence such as "Jones' feet are moving slowly (for walking)" and "Jones' mouth is moving rapidly (for a talker)".

REFERENCES

[1] Alston, W. P., "Meaning and use," *Philosophical Quarterly*, vol. 13 (1963), pp. 107-124.

[2] Carnap, R., "Meaning postulates," pp. 222-232 in *Meaning and Necessity* (enlarged edition), University of Chicago Press, Chicago, Illinois, 1956.

[3] Carroll, L., *The Complete Works of Lewis Carroll*, Random House (Modern Library series), New York, 1936.

[4] Chomsky, N., "A review of B. F. Skinner's *Verbal Behavior*," in *The Structure of Language: Readings in the Philosophy of Language*, eds., J. A. Fodor and J. J. Katz, Prentice-Hall Publishing Company, Englewood Cliffs, New Jersey, 1964.

[5] Chomsky, N., *Aspects of the Theory of Syntax*, M.I.T. Press, Cambridge, Massachusetts, 1965.

[6] Chomsky, N., *The Logical Structure of Linguistic Theory*, The Plenum Press, New York, 1975.

[7] Church, A., "A formalization of the logic of sense and denotation," pp. 3-24 in *Structure, Method, and Meaning*, eds., P. Henle, H. M. Kallen, and S. K. Langer, Liberal Arts Press, New York, 1951.

[8] Church, A., "Intensional isomorphism and identity of belief," *Philosophical Studies*, vol. 5 (1954), pp. 65-73.

[9] Davidson, D., "Truth and meaning," pp. 455-480 in *Readings in the Philosophy of Language*, eds., J. F. Rosenberg and C. Travis, Prentice-Hall Publishing Company, Englewood Cliffs, New Jersey, 1971.

[10] Davidson, D., "Thought and talk," in *Mind and Language*, ed., S. Guttenplan, Clarendon Press, Oxford, England, 1975.

[11] Davidson, D., "The logical form of action sentences," pp. 105-122 in *Essays on Actions and Events*, Clarendon Press, Oxford, England, 1980.

[12] Field, H., "Logic, meaning, and conceptual role," *Journal of Philosophy*, vol. 74 (1977), pp. 379-409.

[13] Foster, J. A., "Meaning and truth theory," in *Truth and Meaning*, eds. G. Evans and J. McDowell, Clarendon Press, Oxford, England, 1976.

[14] Frege, G., "On sense and reference," in *Translations from the Philosophical Writings of Gottlob Frege*, eds., P. Geach and M. Black, Basil Blackwell, Oxford, England, 1952.

[15] Frege, G., "On concept and object," in *Translations from the Philosophical Writings of Gottlob Frege*, eds. P. Geach and M. Black, Basil Blackwell, Oxford, England, 1952.

[16] Frege, G., *The Foundations of Arithmetic*, Basil Blackwell, Oxford, England, 1953.

[17] Frege, G., *The Basic Laws of Arithmetic*, ed. and trans., M. Furth, University of California Press, Berkeley, California, 1964.

[18] Grice, H. P. and P. Strawson, "In defense of a dogma," *Philosophical Review*, vol. 65 (1956), pp. 141-158.

[19] Grice, H. P., "Utterer's meaning, sentence meaning, and word-meaning," *Foundations of Language*, vol. 4 (1968), pp. 225-242.

[20] Harman, G., *Thought*, Princeton University Press, Princeton, New Jersey, 1973.

[21] Hintikka, J., *Models for Modalities*, Humanities Press, New York, 1969.

[22] Katz, F. and J. J. Katz, "Is necessity the mother of intension?" *Philosophical Review*, vol. 86 (1977), pp. 86-96.

[23] Katz, J. J., "Mentalism in linguistics," *Language*, vol. 40 (1964), pp. 124-137.

[24] Katz, J. J., *The Philosophy of Language*, Harper & Row Publishing Company, New York, 1966.

[25] Katz, J. J., *Semantic Theory*, Harper & Row Publishing Company, New York, 1972.

[26] Katz, J. J., "Logic and language: an examination of recent criticisms of intensionalism," pp. 63-76 in *Language, Mind, and Knowledge, Minnesota Studies in the Philosophy of Science*, vol. 7, ed., K. Gunderson, University of Minnesota Press, Minneapolis, Minnesota, 1975.

[27] Katz, J. J., "A proper theory of names," *Philosophical Studies*, vol. 31 (1977), pp. 1-80.

[28] Katz, J. J., *Propositional Structure and Illocutionary Force*, T. Y. Crowell, New York, 1977 (paperback reprint, Harvard University Press, 1980).

[29] Katz, J. J., "The advantages of semantic theory over predicate calculus in the representation of logical form in natural language," *The Monist*, vol. 60 (1977), pp. 380-405.

[30] Katz, J. J., "The theory of semantic representation," *Erkenntnis*, vol. 13 (1978), pp. 63-110.

[31] Katz, J. J., "The neoclassical theory of reference," pp. 103-124 in *Contemporary Perspectives in the Philosophy of Language*, eds., P. A. French et al., University of Minnesota Press, Minneapolis, Minnesota, 1979.

[32] Katz, J. J., *Language and Other Abstract Objects*, Rowman & Littlefield, Totowa, New Jersey, 1981.

[33] Putnam, H., "Is semantics possible?" *Metaphilosophy*, vol. 1 (1970), pp. 187-201.

[34] Quine, W. V. O., "Two dogmas of empiricism," in *From a Logical Point of View*, Harvard University Press, Cambridge, Massachusetts, 1953.

[35] Quine, W. V. O., *Word and Object*, M.I.T. Press, Cambridge, Massachusetts, 1960.

[36] Schiffer, S., *Meaning*, Oxford University Press, London, England, 1972.

[37] Searle, J. R., *Speech Acts*, Cambridge University Press, Cambridge, England, 1969.

[38] Smith, G. and J. J. Katz, *Intensionally Admissible Models*, forthcoming.

The Graduate Center
The City University of New York
33 West 42 Street
New York, New York 10036