

1973

Toward A Naturalistic Theory Of Meaning

Hugh Thompson Wilder

Follow this and additional works at: <https://ir.lib.uwo.ca/digitizedtheses>

Recommended Citation

Wilder, Hugh Thompson, "Toward A Naturalistic Theory Of Meaning" (1973). *Digitized Theses*. 650.
<https://ir.lib.uwo.ca/digitizedtheses/650>

This Dissertation is brought to you for free and open access by the Digitized Special Collections at Scholarship@Western. It has been accepted for inclusion in Digitized Theses by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlsadmin@uwo.ca.

The author of this thesis has granted The University of Western Ontario a non-exclusive license to reproduce and distribute copies of this thesis to users of Western Libraries. Copyright remains with the author.

Electronic theses and dissertations available in The University of Western Ontario's institutional repository (Scholarship@Western) are solely for the purpose of private study and research. They may not be copied or reproduced, except as permitted by copyright laws, without written authority of the copyright owner. Any commercial use or publication is strictly prohibited.

The original copyright license attesting to these terms and signed by the author of this thesis may be found in the original print version of the thesis, held by Western Libraries.

The thesis approval page signed by the examining committee may also be found in the original print version of the thesis held in Western Libraries.

Please contact Western Libraries for further information:

E-mail: libadmin@uwo.ca

Telephone: (519) 661-2111 Ext. 84796

Web site: <http://www.lib.uwo.ca/>



**NATIONAL LIBRARY
OF CANADA**

**CANADIAN THESES
ON MICROFILM**

**BIBLIOTHÈQUE
NATIONALE
DU CANADA**

**THÈSES CANADIENNES
SUR MICROFILM**

1 4 9 8 8

TOWARD A NATURALISTIC THEORY OF MEANING

by

Hugh Thompson Wilder

Department of Philosophy

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Canada

February 1973

© Hugh Thompson Wilder 1973

ABSTRACT

One major source of tension in current philosophy of language is the apparent conflict between three widely discussed theories of meaning. The three theories are what may be called the naturalistic theory, the intentional theory, and the speech act theory. There are two main tasks of this thesis: first, to clarify relations obtaining between particular examples of these three theories of meaning, and second, to attempt to relieve a source of tension in philosophy of language by arguing that one version of the naturalistic theory of meaning is more adequate than any of the other theories discussed.

The three theories of meaning discussed are those of Quine, Grice, and Searle. The theories are described in detail, and it is argued that they are competing theories of meaning. Difficulties with each theory are discussed, and a revision of Quine's theory is proposed which meets certain objections raised against Quine's theory. It is argued, in conclusion, that this

revised version of Quine's naturalistic theory of meaning is more adequate than either Grice's intentional theory or Searle's speech act theory.

TABLE OF CONTENTS

	page
CERTIFICATE OF EXAMINATION	ii
ABSTRACT	iii
TABLE OF CONTENTS	v
CHAPTER 1. INTRODUCTION	1
1.1 Three Levels of Meaning	1
1.2 Three Theories of Meaning	8
1.3 On What is to Count as a Theory of Meaning	30
CHAPTER 2. MEANING: QUINE'S ANALYSIS	35
2.1 Quine's Rejection of the Mentalistic Theory of Meaning	35
2.2 Language and Dispositions to Verbal Behavior	60
2.3 Stimulus Meaning	70
2.4 Private Languages and Private Meanings	76
2.5 Stimulus Meanings, Internal States, and Verbal Stimuli	77
2.6 Stimulus Meaning and Naturalism ...	94
CHAPTER 3. MEANING: GRICE'S ANALYSIS	103
3.1 Meaning _{nn}	103
3.2 Grice's Rejection of the Causal Theory of Meaning	111
3.3 Meaning _{nn} and Speakers' Intentions	125
3.4 Objections to Grice's Analysis of Meaning _{nn}	128

CHAPTER 4.	MEANING: SEARLE'S ANALYSIS	145
	4.1 Language as Rule Governed Behavior	145
	4.2 Speech Acts	162
	4.3 Meaning	167
CHAPTER 5.	"LEVELS OF MEANING" AND COMPETING THEORIES OF MEANING	184
	5.1 Harman's Characterizations of Grice's, Searle's, and Quine's Theories of Meaning	184
	5.2 Three Competing Theories of Meaning	194
	5.3 Criteria for Evaluating Competing Theories of Meaning	200
	5.4 A Defense of the Criteria for Evaluating Competing Theories of Meaning	211
CHAPTER 6.	TOWARD A NATURALISTIC THEORY OF MEANING	218
	6.1 The Naturalistic Theory of Meaning: Stimuli vs. Stimulations	218
	6.2 The Naturalistic Theory of Meaning: Speakers' Internal States	233
	6.3 The Naturalistic Theory of Meaning: Verbal Stimuli	236
	6.4 The Naturalistic Theory of Meaning Defended	242
REFERENCES		264
VITA		268

CHAPTER 1

INTRODUCTION

1.1 Three Levels of Meaning

In his paper, "Three Levels of Meaning," Gilbert Harman describes three different approaches to the theory of meaning; each of the approaches is held to constitute an approach to a different part or "level" of the theory of meaning. "Theory of level 1," as Harman calls it, takes "meaning to be connected with evidence and inference, a function of the place an expression has in one's 'conceptual scheme'" (Harman, 1968, p. 590). "Theory of level 2" takes "meaning to be a matter of the idea, thought, feeling, or emotion that an expression can be used to communicate" (1968, p. 590). And, "theory of level 3" takes "meaning to have something to do with the speech acts the expression can be used to perform" (1968, p. 590).

After briefly describing these three different approaches to the theory of meaning, Harman goes on to review some of the standard objections to each type of theory. His purpose, in reviewing the objections, is neither to defend nor to attack any one of the three types of theory; rather, his purpose is to show that each of the standard objections reviewed has systematically misinterpreted the types of theory in a particularly instructive way. Specifically, Harman claims that each of the objections assumes "that the three approaches to the

theory of meaning are approaches to the same thing" (Harman, 1968, p. 592). The main argument of Harman's paper is concerned to show that this common assumption is false.

Harman's strategy in arguing against the assumption that the three approaches are approaches to the same thing is, as the title of his paper suggests, to distinguish between three "levels of meaning," and to argue that each one of the approaches described is an approach to a theory of a different level of meaning. That is, Harman's strategy is to ascribe different domains to each one of the three approaches to the theory of meaning. Harman's argument is, in other words, that there is no "experimental overlap" between the three approaches; each approach constitutes a theory attempting to explain a unique set of facts about meaning. Thus, Harman contends that theory of level 1 is best understood as a theory "about the nature of meaning of thoughts and other psychological ('intentional') states" (Harman, 1968, p. 596). Theory of level 2 is not to be understood as a theory about the meaning of thoughts, but, rather, as a theory about "the meaning of messages" used in communication (Harman, 1968, p. 602). And, theory of level 3 is held to concern neither thoughts nor messages, but is, rather, a theory about "the meaning of speech acts" (Harman, 1968, p. 602), or a theory that attempts to explain how the existence of certain social and linguistic institutions and conventions can make certain acts possible (Harman, 1968, p. 597).

After describing these different domains of each of

the three approaches to the theory of meaning, Harman goes on to suggest that certain recent theories of meaning are best thought of as being theories of one or another level. Thus, Harman claims that the theories of meaning developed by Carnap, Ayer, Lewis, Firth, Hempel, Sellars, and Quine are all level 1 theories of meaning: i.e., theories about the meaning of thoughts or other psychological states. Theories of meaning developed by Morris, Stevenson, Grice, and Katz are level 2 theories: i.e., theories about the meaning of messages used in communication. And, theories of meaning developed by Wittgenstein, Austin, Hare, Nowell-Smith, Searle, and Alston are level 3 theories: i.e., theories about the meaning of speech acts.

Let us call Harman's claim that each of the three approaches to the theory of meaning is an approach to a theory of a different level of meaning, his "reconciliation thesis." The thesis is a reconciliation thesis because it claims that the three approaches to the theory of meaning are not, in some sense, competing approaches. The reconciliation thesis does not say that theories of meaning of different levels are compatible with one another, but, rather, that theories of different levels are not competing theories in a way in which they might be thought to be competing theories. Thus, for example, it is a common assumption, I think, that Quine's theory of meaning attempts to explain some of the same facts that both Grice's and Searle's theories attempt to explain. The assumption is, in other words, that

Quine's, Grice's, and Searle's theories of meaning are in some sense competing theories. But, Harman is suggesting that this assumption is mistaken. Harman claims that Quine's, Grice's, and Searle's theories of meaning are best thought of, not as competing theories of meaning, but as theories of different levels of meaning--i.e., as theories with different domains.

Harman's paper is important for two reasons: first, he has confronted one of the major sources of tension in current philosophy of language, and, second, he has proposed an ingenious way to resolve this tension. The source of tension that Harman has confronted is the belief that representative theories of each of the three approaches to the theory of meaning are competing theories. And, Harman's proposal for resolving the tension is, not to argue that one theory or general approach is superior to the others, but that each approach has its own domain, and that therefore, e.g., Quine's, Grice's, and Searle's theories of meaning are not, in some sense, competing theories.

Because of the importance of Harman's paper, I propose to examine his arguments in more detail. In this chapter, I will describe, very briefly, certain methodological and substantive errors in Harman's paper, which render it doubtful that Harman has successfully resolved the tension just noted in the philosophy of language. I will not argue yet, however, that Harman's reconciliation thesis is either tenable or not tenable. Rather, I will contend simply that

for anyone to definitively resolve the tension, more work needs to be done than has been done in Harman's paper.

The major methodological reason why I think it is doubtful that Harman can have definitively established his reconciliation thesis in his paper is simply that Harman's canvas is too small for his subject; although Harman's paper is suggestive, it is far too sketchy. This sketchiness becomes apparent in two ways. First, Harman argues for a general reconciliation of the three approaches to the theory of meaning. I suggest that in order to establish his general thesis, viz., that theories of level 1, 2, and 3 are not competitive theories, it is necessary to examine cases of theories of each level. Harman's general thesis can only be established if particular cases of each level can be reconciled in the sense of Harman's reconciliation thesis. Harman has begun this task of considering cases, but it remains to be completed in detail.

The second way in which the sketchiness of Harman's presentation is made apparent is the way in which theories of meaning get grouped together. Harman's main argument is for a general reconciliation of theories of level 1, 2, and 3; in order to establish his general thesis, then, Harman is led to group different theories of meaning into one or another of the three general approaches he sees to the theory of meaning. For example, Harman groups Quine's theory of meaning together with Carnap's, Lewis', and Sellars'. My complaint is that any such grouping must either restrict

attention to only the most general (and trivial) features of the theories being grouped, or distort the theories for the sake of neat grouping. For example, I have argued elsewhere¹ that Lewis' theory of meaning contains important elements of the mentalistic theory of meaning which is explicitly rejected by Quine in Word and Object. Thus, although I think there is a sense in which Harman is justified in linking Quine's theory to Carnap's, Sellars', and even Lewis' theories, I also think that in doing so Harman has glossed over certain crucial differences between the theories--differences which seem to be relevant to his reconciliation thesis.

My general point is, therefore, that if Harman's reconciliation thesis is to be established, then particular theories, such as Quine's, must be shown to be not competing with others, such as Grice's and Searle's. In my thesis, therefore, I abandon Harman's program of examining the relations between the three general approaches to the theory of meaning, in favor of considering the more modest question of what logical relations hold between certain cases of each of Harman's three general approaches; the cases I shall consider are the theories of meaning held by Quine, Grice, and Searle. Of course, much of the interest and significance of Harman's paper lies in the generality of the reconciliation thesis; I have already indicated, however, that accuracy of detail seems to have been sacrificed for the sake of this generality in Harman's treatment. In my own treatment, I hope to

¹In Wilder, 1971.

redress the balance between accuracy and generality; because each of the particular theories of meaning I examine is representative of an important and distinctive approach to the theory of meaning, my conclusions should not be devoid of significance for those interested in the relations between the general approaches to the theory of meaning.

My substantive quarrel with Harman's paper centers on his ascription of different domains to each of the three approaches to the theory of meaning. As I have indicated, Harman claims that the three approaches are not competing theories of meaning, because they are not attempts to explain the same set of facts. I shall argue, in Chapter 5, that this claim of Harman's is simply false. I shall argue that each of Harman's characterizations of the three approaches to the theory of meaning is incomplete, and that once the characterizations are completed it becomes clear that Quine's, Grice's, and Searle's theories of meaning are attempts to explain (at least some) of the same facts about meaning. I shall argue, further, that Quine's, Grice's, and Searle's theories are each inconsistent with the others. From these claims, that Quine's, Grice's, and Searle's theories are mutually incompatible attempts to explain (some of) the same facts about meaning, it follows that Quine's, Grice's, and Searle's theories are in fact competing theories of meaning.

If the three theories of meaning with which I am concerned are in fact competing theories, then it follows that Harman's reconciliation thesis, as it is applied to the

three theories in question, is false. After showing Harman's thesis to be false, in the way just sketched, I will proceed to outline my own resolution of the conflict between Quine's, Grice's, and Searle's theories of meaning. My resolution will not be an attempt to reconcile the theories to each other, as was Harman's; rather, I will describe certain conditions of adequacy for assessing competing theories of meaning, and argue that, on these conditions, an extension of Quine's theory of meaning is more adequate than either Grice's intentional theory or Searle's speech act theory. The extension of Quine's theory which will be sketched and defended in the final chapter of the thesis is a naturalistic theory of meaning, in a sense to be indicated later in this chapter.

The task of this thesis is, then, twofold. First, a description of three widely-held theories of meaning will be given, along with an analysis of the logical relations obtaining between them. Second, an extension of one of the theories--Quine's--will be developed, and an argument will be presented for the claim that it is more adequate than either of the other theories. Along the way, the theory of the structure of language in which is embedded the naturalistic theory of meaning being defended here will be contrasted with the currently ascendant linguistic theory, transformational grammar.

1.2 Three Theories of Meaning

At this point, it will be helpful to introduce, very briefly, some of the different theories of meaning which will

receive more detailed consideration in succeeding chapters. In this section, I will first show why Grice's and Searle's theories each represent important and distinctive approaches to the theory of meaning; second, I will describe the principles of what I call a naturalistic theory of meaning. In the first two parts of this section I will be concerned to make good my earlier claim, that in sorting theories of meaning into different groups Harman has distorted or obscured the theories being sorted: my claim will be supported with reference to Grice's theory and to Searle's theory.

(a) Grice's Theory

Harman claims that Grice's theory is best understood as taking "meaning to be a matter of the idea, thought, or emotion that an expression can be used to communicate"; and, the special domain of Grice's theory is, according to Harman, "to show what communication is and what is involved in a message's having a particular meaning" (Harman, 1968, p. 596). Now, while I think that Harman's emphasis on the notion of communication in his characterization of Grice's theory is helpful, I also think that his characterization ignores that aspect of Grice's theory which has been most emphasized by Grice and which has been at the center of the controversy over the adequacy of Grice's theory. The aspect to which I am referring is Grice's reliance on the notion of intention.

Grice begins his original account of meaning² by

²In Grice, 1967.

distinguishing between the "natural" and the "non-natural" senses of "meaning." Roughly, Grice thinks that his distinction isolates the sense of "means" which is connected with communication from the sense of "means" which is not connected with communication.³ It is the non-natural sense of "means"--the sense which is connected with communication--with which Grice is concerned; in this sense, therefore, Harman is right to emphasize Grice's concern with communication.

Grice takes as his basic analysandum the expression "By uttering x U meant something,"⁴ where x stands for any utterance, U is a human agent, and "meant" is taken in its non-natural sense. Grice acknowledges that he is using "the terms 'uttering' and 'utterance' in an artificially extended way, to apply to any act or performance which is or might be a candidate for nonnatural meaning" (Grice, 1969, p. 151). Specifically, non-linguistic acts or objects (e.g., hand signals) as well as linguistic units (e.g., words, phrases, and sentences) are candidates for non-natural meaning--i.e., are utterances. Thus, Grice is not interested in simply the meaning of sentences in languages; he is not aiming to develop simply a theory of what is called in Chapter 6 the meaning of linguistic units. Rather, he is concerned to develop a general theory of what it is for an agent to mean x by uttering U, for variable x and U; since "U" may range over linguistic units as well as non-linguistic units, the

³V Grice, 1969, p. 147.

⁴V Grice, 1967, p. 43 and Grice, 1969, p. 151.

theory of the meaning of linguistic units will be a special case of Grice's general theory of the meaning of utterances.

Harman's specification of communication and the meaning of messages as the domain of Grice's theory of meaning is, therefore, correct, so far as it goes. I do not believe, however, that Harman's characterization of Grice's theory is sufficient to distinguish it from certain other theories, which conflict with Grice's in important respects; this claim may be established as follows. Grice is concerned to analyze the expression "By uttering x U meant something." There seem to be, however, at least three general ways in which this expression could be analyzed, each of which is incompatible with the others. Grice devotes much of his effort to showing that two of the ways are inadequate, while the third is adequate. And, Harman's characterization of Grice's theory does not serve to distinguish Grice's theory from its competitors in these respects.

The three approaches one could take toward giving an analysis of the expression "By uttering x U meant something" are, as I shall call them, the naturalistic, the intentional, and the illocutionary act approaches. (These approaches are not to be identified with Harman's three approaches to the theory of meaning; they are not, as Harman's are claimed to be, "approaches to different things.") A naturalistic analysis of the above expression would, very roughly, assert that the expression "By uttering x U meant something" is true if either x tends to produce in an audience some effect or

response, or x tends to be elicited by some particular stimulus situation. An intentional account, very roughly, asserts that the expression "By uttering x U meant something" is true iff x was intended by U to produce in an audience some effect or response. And, an illocutionary act account asserts that the expression "By uttering x U meant something" is true iff x was intended by U to produce in an audience some illocutionary effect (with "illocutionary effect" suitably defined as certain effects of utterance acts generally, and not merely effects of linguistic acts).

Of course, each of these three kinds of analysis of meaning would have to be considerably refined, in order to render them at all plausible. My point, however, is that the distinguishing mark of Grice's theory is not its adoption of the expression "By uttering x U meant something" as the basic notion to be analyzed; for, in the terminology just adopted, naturalistic theories, intentional theories, and illocutionary act theories may each adopt as their basic analysandum the expression "By uttering x U meant something." That Grice thinks his intentional theory to be the most adequate is the crucial point, obscured in Harman's characterization, which must be considered in any attempt to describe the relations between Grice's theory and others.

Grice argues, then, that what I have called naturalistic and illocutionary act theories of meaning are inadequate. The main purpose of each of his papers on meaning has been to show that any adequate analysis of the expression "By

uttering x U meant something" must "involve a reference to an intended effect of, or response to, the utterance of x" (Grice, 1969, p. 151). And, Grice has suggested that the intended effect need not be, as illocutionary act theories would have it, an illocutionary effect. The main concern of my chapter on Grice's theory of meaning will be, therefore, to elucidate the intentional aspect of Grice's theory, and to evaluate Grice's defense of his theory against both naturalistic and illocutionary act theories of meaning.

(b) Searle's Theory

Searle's theory of meaning is, of course, a representative of what I have called an illocutionary act theory of meaning. Harman is correct in interpreting Searle's theory as a theory about the meaning of speech acts, and a theory which takes "meaning to have something to do with the speech acts the expression can be used to perform." However, to say that Searle's is a theory about speech acts is not enough to distinguish it from other competing theories of meaning. For, as Searle himself is aware,⁵ speech acts are a species of Gricean utterances, and therefore a part of Grice's theory of meaning may be used in an analysis of the meaning of speech acts. Since, as I shall argue, Grice's theory is inconsistent with Searle's, Harman's simple characterization of Searle's theory as a theory of the meaning of speech acts is misleading insofar as it does not specify

⁵V Searle, 1969, pp. 43ff., and Searle, 1971, pp. 8ff.

any important differences between Searle's and Grice's theories. As was the case with Harman's characterization of Grice's theory of meaning, in the present case Harman's characterization of Searle's theory fails to distinguish it from competing theories of meaning.

Searle's theory of meaning is, nevertheless, a theory about the meaning of speech acts. Unlike Grice, Searle is concerned with the meaning of specifically linguistic units (e.g., sentences): "speech acts" are linguistic acts, and the set of all speech acts is a proper subset of the set of all Gricean utterances (proper, because Gricean utterances can be both linguistic and non-linguistic acts and performances, while speech acts are linguistic acts). Speech acts are linguistic acts, while Gricean utterances are acts of communication generally, linguistic or otherwise. Therefore, Searle is concerned with language in a way that Grice is not. Several other differences between Searle's and Grice's theories may be clarified via an examination of Searle's conception of language.

Searle conceives of languages and speech acts as analogous to games and moves in the games. Roughly, a game is created or made possible, and subsequently regulated, by a set of rules; and language is "rule governed" in the sense that it, like a game, is constituted and regulated by a set of rules. To speak a language is to perform linguistic acts: i.e., to make moves in a game which is constituted and governed by a set of rules.

Thus, language is conventional for Searle in the sense that it is rule governed. Specifically, "the semantic structure of a language may be regarded as a conventional realization of a series of sets of underlying constitutive rules" (Searle, 1969, p. 37). And, just as the rules of baseball create new forms of behavior (such as hitting a home run, or striking out), so the rules of language make possible a variety of linguistic acts. According to Searle, the rules of language make possible four general kinds of linguistic acts: utterance acts, propositional acts, illocutionary acts, and perlocutionary acts.

Although there will be occasion to discuss these four types of linguistic acts in more detail later, it will be helpful to characterize them briefly now. Let us suppose a speaker utters the following sentence:

(1) Sam smokes habitually.

Now, in uttering this string of phonemes the speaker is, according to Searle, "saying something" in English, and not merely "mouthing words." Searle calls this "uttering of words (morphemes, sentences)," utterance acts. When the speaker utters (1), he also refers to Sam, and, according to Searle, "predicates the expression" "smokes habitually" of Sam. In thus referring and predicating, the speaker performs a propositional act. And, when the speaker utters the sentence in question, he also states or asserts that Sam smokes habitually. In doing so, the speaker performs the illocutionary act of making an assertion (Searle, 1969, pp. 22-24).

Searle also believes that "correlated with the notion of illocutionary acts is the notion of the consequences or effects such acts have on the actions, thoughts, or beliefs, etc. of hearers" (Searle, 1969, p. 25). For example, when I warn somebody about something (i.e., perform the illocutionary act of warning), I might scare or alarm him; and when I request something of somebody (i.e., perform the illocutionary act of requesting), I might get him to do something. Searle believes that scaring or alarming someone, and getting someone to do something, are perlocutionary acts. Thus, the effect of performing the illocutionary act of warning somebody about something might be the perlocutionary act of scaring or alarming the hearer; and, the effect of performing this perlocutionary act might be to get him to do something. Illocutionary and perlocutionary acts seem to be distinguished on the grounds that their effects are different; this point will be examined later.

For now, though, it is sufficient to notice that Searle uses his distinction between illocutionary effects and perlocutionary effects in his first main criticism of Grice's theory of meaning. Searle's criticism of Grice's theory is that Grice identifies meaning with, roughly, intended perlocutionary effect. Grice does not use either "illocutionary" or "perlocutionary" in discussing the kinds of effects that are intended when someone means something by uttering some utterance; but, Searle claims that Grice intends the effects to be perlocutionary effects, which is,

according to Searle, mistaken. According to Searle, meaning is to be analyzed in terms of the intended illocutionary effects of speech acts.

The questions of whether Searle's first criticism of Grice's analysis is justified, and, if it is, whether Grice's analysis can and ought to be revised to accommodate it, will be dealt with later. For now, the point is simply that while Grice apparently believes that meaning can be analyzed without referring to either perlocutionary or illocutionary effects, Searle believes that the analysis must be given in terms of illocutionary effects. Thus, one crucial element of Searle's theory of meaning which is obscured in Harman's characterization of the theory is the emphasis on illocutionary effects. As already indicated, Grice's theory may be interpreted (and is so interpreted by Searle) as, in part, a theory about the meaning of speech acts; one difference between Grice's and Searle's theories, obscured by Harman's oversimplified characterizations, is Searle's emphasis on illocutionary acts and Grice's alleged emphasis on perlocutionary acts.

Searle's second criticism of Grice's analysis of meaning is that while meaning has both intentional and conventional aspects, Grice's analysis obscures or ignores the conventional aspect. As we noted, according to Grice the expression "By uttering x U meant something" is true iff x was intended by U to produce in an audience a certain effect or response. And, as we shall see, Grice believes that the question, What did U mean by uttering x? can be answered by

specifying the intended effect or response (Grice, 1969, p. 151). Thus, for example, let us suppose that a madman utters the string of phonemes "Gleeg, gleeg, gleeg!"⁶ The madman could be said to mean, on Grice's account, that it is snowing in Tibet, if the madman had what may be called "the right Gricean intention" in uttering "Gleeg, gleeg, gleeg!" On Grice's account, the madman means something in this case in exactly the same sense in which I mean something when I say "It is snowing in Tibet." And, on Grice's account, the madman means the same thing when he says "Gleeg, gleeg, gleeg!" as I do when I say "It is snowing in Tibet."

Searle's complaint is that the madman can't mean that it is snowing in Tibet when he says "Gleeg, gleeg, gleeg!" in the same sense in which I mean that it is snowing in Tibet when I say "It is snowing in Tibet," because "Gleeg, gleeg, gleeg!" does not mean anything in any language. According to Searle, my utterance "It is snowing in Tibet" means something because of both linguistic conventions and my intentions; and the madman's utterance means nothing because it is not a linguistic act governed by any linguistic conventions, even though it was uttered with the right intentions.

Searle is claiming, then, that Grice's account of meaning is too weak: that it ascribes meaning to some utterances that (obviously) have no meaning. Searle believes that Grice's analysis can be made acceptable by revising it

⁶This example is adapted from Ziff, 1967, p. 5.

to include (among other things) some condition concerning the conventional aspect of meaning. Again, whether Searle's criticism of Grice's analysis is justified, and, if it is, whether Grice's analysis can and ought to be revised in the way suggested by Searle, are problems which remain to be solved.

(c) Principles of a Naturalistic
Theory of Meaning

So far, I have described, very briefly, Grice's attempt to analyze meaning in terms of intentions and Searle's attempt to analyze meaning in terms of illocutionary acts. These theories of meaning will receive more detailed treatment in subsequent chapters, as will Quine's theory. I describe Quine's theory of meaning in order, first, to contrast a naturalistic theory with the theories of Grice and Searle, and, second, to introduce the naturalistic theory of meaning to be defended in the final chapter of the thesis. At this point, therefore, it will be more helpful to sketch the principles of what I call a naturalistic theory of meaning, rather than proceed to introduce Quine's theory itself. I will argue, later, that Quine's theory is a naturalistic theory in the relevant sense, and that the theory to be defended in the final chapter is a consistent extension of it.

Naturalism, as I understand it, is a methodological position rather than an ontological or metaphysical

doctrine.⁷ Or, more precisely, naturalism is committed to a particular ontological doctrine only insofar as its methodology dictates such a commitment.

Following Danto, we may note certain "tenets of naturalism"; any naturalistic theory of meaning must, at least, subscribe to these tenets.

(1) Naturalism is, first of all, a commitment to a kind (or kinds) of explanation--namely, that kind (or kinds) employed in the natural sciences (Danto, 1967, p. 448). Naturalism is committed to the view that the models of explanation used in the natural sciences--whatever these models may be--are sufficient as models for all phenomena, whether in the natural sciences or not. Thus, while there may be debates about what the models of explanation in the natural sciences are, naturalism--even though it is a methodological position--need not take sides in this debate. Naturalism adopts, in every sphere of scientific inquiry, the kind of explanation adopted in the natural sciences, whatever this kind may be. It is important to realize that no assumption is made here that the model of explanation in the natural sciences is either simple or monolithic. Rather, what is being denied is that there is any kind of behavior, linguistic or otherwise, that cannot, in principle, be explained by the same kind of explanation used in the natural sciences. Thus, for example, Winch's thesis (1958) that explanation in the social sciences must be fundamentally different from

⁷My account of naturalism is based on Danto, 1967.

explanation in the natural sciences is, according to this conception of naturalism, an anti-naturalistic thesis. On the other hand, a thesis which questioned the adequacy of, e.g., Hempel's characterization of explanation in the natural sciences would not necessarily be anti-naturalistic.

(2) What exists are natural objects: i.e., all those objects whose changes are susceptible to explanations not different in kind from the explanations given in the natural sciences--or, what I shall call "naturalistic explanations"--and only those objects which are required in order to explain other natural phenomena.⁸ According to naturalistic principle (1) any change in a natural object or system of natural objects is explicable by naturalistic explanations; and, according to the first part of principle (2), only those objects whose changes are explicable by naturalistic explanations exist. Thus, any ontological thesis that asserted that there are basically two kinds of objects and processes in the universe, the one kind being explicable by the kind of explanation used in the natural sciences, and the other kind not being explicable by naturalistic explanations would be an anti-naturalistic thesis.

Implicit in the second part of naturalistic principle (2) is a "principle of ontological parsimony," which expresses the naturalist's desire, in Quine's words, "to dream of no more things in heaven and earth than need be" (Quine, 1970a, p. 3).

⁸The second part of this principle goes beyond anything said by Danto; cf. Danto, 1967, p. 448.

According to this principle, entities (i.e., natural objects) are not to be posited without good reason, the only good reason being that they are required in order to explain other phenomena. Thus, any ontological thesis which posited entities which were useless in explaining other naturalistic phenomena would be an anti-naturalistic thesis.

An example of an application of this naturalistic principle of ontological parsimony may be drawn from Bloomfield's discussion of what we may call naturalistic and anti-naturalistic theories of meaning.⁹ Bloomfield argues that the naturalist and the anti-naturalist must have precisely the same data available, on which to proceed in their investigation of the meanings of utterances: specifically, the data can only be drawn from the speakers' situations (which may include internal neurophysiological states, as well as "outer" behavioral situations), the actual utterances being investigated, and the hearers' responses. All these data are constituted by natural objects and/or processes; no extra data are available to either the naturalist or the anti-naturalist that are not available to the other.

Nevertheless, Bloomfield observes, if the naturalist is asked "What does utterance U mean?" he will answer with reference to the speaker's situation and the hearer's response; if the anti-naturalist is asked the same question, he will answer with reference to the speaker's situation,

⁹Bloomfield calls them mechanistic and mentalistic theories; y (Bloomfield, 1935, pp. 142-144).

the hearer's response, plus certain mental events alleged to occur in the speaker and hearer in connection with the utterance. The point is that the naturalist and the anti-naturalist seem to have different definitions of meaning: the naturalist defines the meaning of a linguistic form (according to Bloomfield) simply as the speaker's situation and the hearer's response; the anti-naturalist, however (according to Bloomfield), defines the meaning of a linguistic form "as the characteristic mental event which occurs in every speaker and hearer in connection with the utterance or hearing of the linguistic form" (Bloomfield, 1935, p. 142).

Now, either the anti-naturalist's "mental events" (at least, those connected with particular utterances) are part of the speaker's situation or the hearer's response, or not. Bloomfield argues that if they are part of the speaker's situation or hearer's response, then the naturalist will have included them in his definition of meaning, in which case there is no disagreement between the naturalist and the anti-naturalist concerning the definition of meaning (Bloomfield, 1935, p. 144). On the other hand, if the mental events are not part of either the speaker's situation or the hearer's response, then it is hard to see what evidence the anti-naturalist has for positing their existence. For, if there is any evidence for the existence of the mental events allegedly connected with the meanings of linguistic forms, such evidence must lie in speakers' situations and hearers' responses; this point is a consequence of the anti-naturalistic

thesis being considered, viz., that the meanings of linguistic forms are the characteristic mental events which occur in every speaker and hearer in connection with the utterance or hearing of the linguistic form. Therefore, there can be no evidence, independent of speakers' situations and hearers' responses, for the existence of the mental events. Bloomfield concludes that the mental events alone--without additional reference to speakers' situations and hearers' responses--cannot be used to explain meanings.

Bloomfield's point is, then (in our terms), that to identify meanings with mental events is anti-naturalistic, because such mental events (if there is any evidence at all for their existence) explain no phenomena that are not adequately explained by reference simply to speakers' situations and hearers' responses. The mental events are just so much extra baggage: they do not explain anything that cannot be explained in terms of the very things that constitute the only available evidence for the existence of the mental events--viz., speakers' situations and hearers' responses. Since the mental events themselves are useless in explaining meanings, they need not be posited; any theory of meaning that posits them will be an anti-naturalistic theory.¹⁰

¹⁰Quine has given a similar argument, on several occasions, for the claim that mentalistic "ideas" ought not be posited, because they are either redundant (i.e., they explain nothing that cannot be explained without them), or they are in need of explanation themselves, and therefore cannot be used to explain anything else. Quine, 1961, pp. 47-48.

(3) Changes in natural objects or systems of natural objects are effects of natural causes. And "a natural cause is a natural object or episode in the history of a natural object which brings about a change in some other natural object" (Danto, 1967, p. 448). There are no changes in natural objects that are not effects of natural causes.

(4) A natural process is any change in a natural object or system of natural objects (Danto, 1967, p. 448).

(5) Natural processes are regular. Therefore, in order to explain natural phenomena, the naturalistic method seeks to establish general laws covering those phenomena (Danto, 1967, p. 448).

(6) Explanations of natural phenomena in terms of non-natural causes are non-natural explanations, and are not explanatory. A non-natural explanation is only used as a last resort, and the occurrence of such an explanation "merely underscores the fact that something cannot be explained or made intelligible at the moment--it does not provide an alternative kind of explanation or intelligibility. All non-natural explanations are . . . in principle replaceable with natural explanations" (Danto, 1967, p. 449).

(7) Human beings are natural objects, and human behavior is explicable by the naturalistic method. Human institutions and practices are natural objects, processes, or systems of natural objects, and are explicable by the naturalistic method. The goals and values of man are properties of natural objects, and are explicable by the naturalistic

method (Danto, 1967, p. 449). Thus, there are not "two worlds"--the world of facts and the world of values--but only one world--the natural world. This principle is a corollary of principle (2), according to which all and only those things exist whose changes are explicable by the method of naturalism.

According to these principles, a naturalistic theory of meaning must be one according to which the methods of investigating the meaning of utterances or sentences are the methods of natural science, and facts about meaning must be explained, if at all, by naturalistic explanations. Meanings must be considered as natural objects, natural processes, or properties of such objects or processes. As such, the meanings of utterances or sentences may be explained in terms of other natural objects or processes. Any non-natural explanation of phenomena germane to meaning is not explanatory, and must be replaceable, in principle, by a naturalistic explanation. Meanings must be considered to be regular, in the sense that in order to explain phenomena concerning meanings linguists must try to establish natural laws which cover the phenomena. And, language itself, as a human institution, must be considered as either a natural object or process, or property of natural objects. These principles of a naturalistic theory of meaning will be clarified as particular theories are shown to conform to them.

A word of warning concerning the meaning of the term "naturalistic theory of meaning" is in order. Because Quine's theory of meaning has also been called a behavioristic theory,

it will be helpful to specify that according to the usage adopted here not all behavioristic theories have been or need to be naturalistic theories, and not all naturalistic theories have been or need to be behavioristic theories. I have already indicated how "naturalism" is to be understood in this thesis; and since I am not concerned to show (or dispute) that Quine's or any other theory is behavioristic, I need not specify what might be meant by "behaviorism." The important point is simply that I do not take "behaviorism" to be synonymous with "naturalism." Quine says that his theory is not behavioristic if behaviorism is narrowly interpreted as being committed to a rejection of innate mechanisms of learning and/or a schematization of learning as conditioned response (Quine, 1969a, p. 96). Since behaviorism usually is associated with one or both of these doctrines, and since Quine rejects them both, I have avoided using the term "behaviorism" in connection with Quine's theory of meaning.

The above principles of naturalism are sufficient to distinguish any naturalistic theory of meaning from both Grice's and Searle's theories. This point will be argued in detail later, but for now it is sufficient to notice, first, that Grice does not consider meanings to be natural objects or processes (or properties of such objects or processes), in the sense just noted. This is so because he believes that meanings, but not certain other things, must be explained in

terms of intentions, and that intentions are neither natural objects nor processes (nor properties of such objects or processes). Moreover, he believes that explanations of meanings in terms of intentions are explanatory, and he is not concerned with the question of whether his non-natural explanations of meanings are replaceable by explanations in terms of natural objects or processes. Thus, Grice's theory of meaning conflicts with our naturalistic principle (1) insofar as Grice believes that explanations of meanings in terms of intentions are explanations of a different kind than those that are given of phenomena in the natural sciences. His theory conflicts with principle (2) insofar as he posits two kinds of objects in the world--entities that may be explained without reference to intentions or other intensional concepts, and entities that can only be explained with reference to intentions and other intensional concepts. And, his theory conflicts with principle (6) insofar as he thinks that his explanations in terms of intentions are explanatory, and not in need of replacement.

Searle's illocutionary act theory also conflicts with certain principles of naturalism. First, Searle adopts a modified version of Grice's analysis of meaning; this modified version still relies on the notion of intention, and Searle conceives of intentions as does Grice, as non-natural objects (or properties of non-natural objects). Therefore, Searle's theory conflicts with our naturalistic principles in at least the same ways that Grice's theory conflicts with

them. In addition, Searle's basic assumption, that language is rule governed behavior, conflicts with certain naturalistic principles as well. Thus, Searle believes that there are two general kinds of phenomena in the world--i.e., what have become known as "brute" and "institutional" facts.¹¹ "Institutions are systems of constitutive rules" (Searle, 1969, p. 51), and institutional behavior is rule governed behavior. Brute facts are not rule governed. There are regularities of institutional behavior as well as of brute facts; but explanation of institutional regularities must be in terms of rules, while explanation of regularities of brute facts must be in terms of covering laws (Searle, 1969, p. 55). Thus, Searle's theory conflicts with naturalistic principle (1) insofar as he relies on two different and irreducible kinds of explanation; it conflicts with principle (2) insofar as it posits two kinds of entities--brute facts and institutional facts; it conflicts with principle (6) insofar as it contends that explanations in terms of rules, which are non-natural explanations, are explanatory, and do not need to be replaced; and it conflicts with principle (7) insofar as it assumes that language, as a human institution, is not explicable by the naturalistic method.

It remains to be shown in detail why Grice's and Searle's theories cannot be modified, without completely distorting them, to make them consistent with our principles of naturalism. It also remains to be shown that the theory of

¹¹Searle, 1969, pp. 50ff.; cf. also Anscombe, 1958.

meaning developed by Quine is a naturalistic theory of meaning, and that the naturalistic theory of meaning to be sketched in the last chapter is more adequate than either Grice's theory or Searle's theory.

1.3 On What is to Count as a Theory of Meaning

So far, the notion of a theory--as in "theory of meaning"--has been taken for granted; it was indicated that one purpose of this thesis is to describe several theories of meaning, but the sense in which such theories are theories has not been discussed. It should be helpful, at this point, to indicate what is to be counted as a theory of meaning.

Roughly, any more or less systematic attempt to either make sense of everyday talk about meaning, or show why such talk (or some of it) is nonsense, is to be counted as a theory of meaning. To make sense of everyday talk about meaning is to give an analysis of the concept of meaning: as may have been noticed, I use the terms "analysis" and "theory" interchangeably. To call the subject matter of this thesis "analyses of meaning" is to emphasize (a) that what is being investigated is a concept in common usage, and (b) that the result of the investigation will be an account of meaning couched in terms which are more precise and perspicuous than those in which preanalytical beliefs about meaning are expressed. To call the subject matter of this thesis "theories of meaning" is to emphasize that the ultimate aim of the analyses/theories is explanatory: what is to be

explained are facts about meaning.

There are at least two kinds of facts about meaning which require explanation. The first kind concern the meaning simpliciter of linguistic units: we talk about the meaning of utterances or inscriptions, words, phrases, sentences, or extended passages; we say that such units have particular meanings, and that sometimes meanings are unclear or ambiguous. The second kind of facts about meaning that requires explanation concern synonymy relations: we say that two words or phrases are synonymous, that a direct quotation preserves the meaning of (is synonymous with) the passage quoted, and that indirect quotations, paraphrases, and translations preserve the meanings (to some degree or other) of the passages quoted, paraphrased, or translated. In addition, a theory of meaning must give an account of concepts thought to be definable in terms of either meaning simpliciter or synonymy--e.g., analyticity. The task of the theory of meaning, then, is to explain particular phenomena occurring within the contexts of everyday talk about meanings: e.g., to explain why sentence S_1 means x_1 , why the meaning of S_2 is ambiguous, why S_3 and S_4 are synonymous, etc.

Although the form of such explanations is left unspecified in this account of what is to count as a theory of meaning (because one theory of meaning may employ covering law explanations, while others may employ rules rather than laws), any adequate explanation will provide, I shall assume, a functional account of certain phenomena germane to meaning.

Meanings are taken as the dependent variable of such an account, and are shown to be a function of other independent variables.¹² To say that the theory of meaning provides a functional account of phenomena germane to meaning is to make no commitment as to the kind of entity meanings are, or to the kind of entities the independent variables are: meanings may be abstract entities or natural objects, and the independent variables may be natural objects, speaker's or hearer's intentions, or institutional facts. The characterization of the theory of meaning as providing a functional account of meaning as a dependent variable does not preclude any particular analysis of the ontological status of meanings and the independent variables in the account.

Obviously, however, some analysis of the notion of meaning is presupposed by any explanation of phenomena germane to meaning. And, in our terms, to identify the independent variables of which meaning is a function is to provide the required analysis. For example, an explanation of the fact that utterance U_1 of S_1 means x might be given in terms of some general law according to which whenever conditions C_1, \dots, C_n obtain, then utterance U_1 of S_1 means x . Such an explanation, in effect, picks out certain elements of the situation in which S_1 was uttered, and asserts that these circumstances determine the meaning of S_1 . What the independent variables are which are relevant to meaning depends on

¹²Cf. Skinner's treatment of meaning as an independent variable in his functional account of verbal behavior, (Skinner, 1957, pp. 10-14).

one's analysis of the notion of meaning; the choice of the conditions relevant to meaning (the independent variables) constitute one's analysis of meaning. Therefore, although the prime objective of the theory of meaning is explanatory, one of the crucial ingredients of the theory must be an analysis of the notion of meaning. That a crucial ingredient of the theory of meaning is an analysis of the concept of meaning, in the sense just spelled out, does not distinguish the theory of meaning from any other scientific theory: a functional account of any phenomena must include a statement of what the independent variables are taken to be.

The conditions of adequacy used in assessing theories of meaning will be described in a subsequent chapter; it is important to bear in mind, however, that one's analysis of the notion of meaning may in fact preclude one from explaining certain phenomena which were, pre-analytically, thought to be germane to meaning. For example, one might give an analysis of the notion of meaning which rendered it impossible for any two linguistic forms to be synonymous. In this case, the theory of meaning could be interpreted as asserting that part of our everyday talk about meanings--that part which relies on the assumption that pairs of linguistic forms can be synonymous--is nonsense. In general, then, it must be remembered that the conditions of adequacy for an analysis of meaning need not include a requirement that the analysis "save" all our preanalytic

beliefs about meaning. While the main objective of the theory of meaning is to explain phenomena germane to meaning, it may be that the phenomena deemed by the theory to be germane to meaning are not identical with the phenomena thought preanalytically to be germane to meaning.

CHAPTER 2

MEANING: QUINE'S ANALYSIS

2.1 Quine's Rejection of the Mentalistic Theory of Meaning

In this chapter, Quine's theory of meaning will be described. The main thesis of the chapter is that Quine's theory is a naturalistic theory of meaning, as defined in 1.2. Quine says that his theory of meaning is a naturalistic theory--or, in his terms, that semantics is nothing other than a chapter of natural science (and, in particular, of psychology).¹ The main task of this chapter is, therefore, to show how Quine's theory of meaning is a naturalistic theory in the relevant sense.

Before describing Quine's theory of meaning, it will be helpful to describe a theory of meaning which Quine has explicitly rejected. This theory is an extreme version of what we may call the mentalistic theory of meaning. According to Quine, the prevalence of the mentalistic theory of meaning has blinded philosophers of language to the fact that semantics is really a chapter of natural science. Therefore, a brief examination of the mentalistic theory of meaning, and of Quine's criticisms of it, should provide an instructive introduction to Quine's own theory.

¹Quine, 1969c, pp. 26-30 and 83-89; y also Quine, 1970a, p. 5, and Quine, 1970b, pp. 3-5.

In the first section of this chapter, each of Quine's three main criticisms of the mentalistic theory of meaning is examined. I argue that Quine's first argument against the mentalistic theory is sound, while each of the other two arguments, considered separately, beg the question at issue, concerning the adequacy of the mentalistic theory of meaning. I argue that these criticisms of the mentalistic theory rely on a prior acceptance of Quine's own theory of meaning, and therefore do not give anyone who has not already adopted a Quinian theory of meaning any reason for abandoning the mentalistic theory; only Quine's first objection to the mentalistic theory provides such a reason.

Whether the whole of the mentalistic theory of meaning, as it is being understood here, has been held by any one philosopher is an open question; certainly, parts of it have been held by many.² At the heart of the mentalistic theory of meaning lies an identification of meanings with propositions. Mentalists have been led to make this identification, according to Quine, because of, first, a series of inferences based on certain idioms of ordinary talk about meanings, and, second, a mistaken conception of the mind as what Quine has called a "mental museum."

First, mentalistic philosophers have, says Quine, taken a cue from ordinary language, in developing their theory of meaning. Starting from our talk about words and

²Quine, 1960, pp. 200-206. V also Wilder, 1971, in which a substantial part of the mentalistic theory rejected by Quine is attributed to Lewis.

sentences as "having meanings," Quine reports the mentalistic philosophers as being led to infer that there must be meanings which words and sentences have (Quine, 1960, p. 206). And, taking another cue from our ordinary talk about "expressing meanings" in our utterances of words and sentences, mentalistic philosophers have been led to locate the meanings which words or sentences have in speakers' minds. In this way, reports Quine, the theory is developed that meanings are certain abstract entities--often called "ideas"--that are located in speakers' minds, and that are expressed by utterances. Quine calls this theory the "museum myth" of meaning (Quine, 1970a, p. 8; 1969c, p. 27; and 1970b, pp. 5, 7, and 9); it is a version of what I am calling the mentalistic theory of meaning. According to the museum myth, the mind is like a museum in which the exhibits are meanings (ideas), and words and sentences are the labels of the exhibits. "It is as if there were a gallery of ideas, and each idea were tagged with the expression that means it" (Quine, 1970a, p. 8).

Many mentalistic philosophers have gone on to identify meanings with propositions. Propositionalists still depend, according to Quine, on the museum myth of the mind; it is just that meanings are said to be propositions, and each meaningful expression is the tag or label for the proposition that it means. According to this sophisticated version of the mentalistic theory of meaning, each declarative sentence is held to express a proposition; each sentence expresses

the proposition that it means. Not each proposition, however, is held to be the meaning of one and only one declarative sentence. As we shall see, because of phenomena such as synonymy, there are some sentences, which are not identical with one another, which are held to express the same proposition.

The advantage to the mentalist in identifying meanings with propositions is that the mentalistic theory of meaning can be supplemented, and, apparently, supported, with a considerable amount of philosophical talk about propositions. Propositions are useful to the semantic theorist, because they seem to allow him to link such notions as meaning, truth, belief, and synonymy. Thus, in addition to being the meanings of sentences, propositions are also held by the mentalistic theorist to be truth vehicles and the objects of our beliefs, desires, etc. (i.e., the objects of what are known--prejudicially, from a Quinian point of view--as the "propositional attitudes"). Propositions are also held by the mentalist to be what "stand in the logical relation of implication" (Quine, 1970a, p. 2). Propositions are held to be shared by synonymous words, sentences, and extended passages: they are the constants of unilingual synonyms, multilingual translations, and philosophical analyses (Quine, 1960, p. 206). And, finally, propositions (or their absence) are held to account for the phenomenon of meaninglessness. Meaningless linguistic forms are held to be meaningless because they do not express propositions; in the language of the museum myth,

meaningless expressions are simply "unattached labels": labels for non-existent exhibits (Quine, 1970b, p. 7). When "the mentalistic theory of meaning" is referred to in this thesis, reference is always being made to that theory according to which meanings are taken to be propositions, which are considered as abstract entities located in speakers' minds.

Quine's three main criticisms of the mentalistic theory of meaning are each directed against propositions, in all of their various roles just listed. In his first objection to the theory, Quine applies a Cartesian principle of clarity and distinctness, and argues that propositions ought to be dispensed with because they do not satisfy the principle. The principle appealed to by Quine is that entities which lack clear identity conditions are to be avoided. Quine argues that propositions lack clear identity conditions, and, therefore, they are to be avoided (Quine, 1969c, p. 140).

Mentalists have traditionally attempted to state the identity conditions of propositions--or, in other words, to state the principle of individuation for propositions--in terms of notions such as synonymy and cognitive equivalence. Thus, each declarative sentence is held, by the mentalist, to express a proposition; but, propositions cannot be identified with the sentences, because in cases of synonymy two or more sentences are held to express the same proposition. The mentalist argues, therefore, that if we could develop an adequate definition of synonymy or cognitive equivalence of sentences, then we could identify a proposition as that

which is expressed by the declarative sentence that expresses it (or, means it) and by all cognitively equivalent or synonymous sentences.

Quine has argued on several occasions³ that any attempt to state identity conditions of propositions in terms of synonymy or cognitive equivalence is doomed to failure. I shall only consider here Quine's argument against what is perhaps the most recent and sophisticated attempt to define propositions as shared information expressed by all cognitively equivalent sentences--viz., the definition provided by possible world semantics. According to this theory, propositions are identified with the "objective information" expressed by particular sentences; and "objective information" is defined in terms of possible worlds in which the sentences are true. Thus, consider a space-time matrix over which elementary particles may be distributed in various patterns: "each distribution of elementary particles of specified kinds over total space-time may be called a possible world; and then two sentences mean the same proposition when they are true in all the same possible worlds" (Quine, 1970a, p. 4). And, "the class of possible worlds in which a sentence comes out true is, we might say, the sentence's objective information--indeed, its proposition" (Quine, 1970a, p. 4).

The above may be taken as a sophisticated attempt to specify identity conditions for propositions, in terms of the concepts of objective information and possible worlds--

³Most notably, in Quine, 1961, pp. 20ff.

successors to the notions of synonymy and cognitive equivalence. Quine's objection to this identification of propositions in terms of possible worlds, as his objection to all earlier attempts to identify propositions in terms of synonymy and cognitive equivalence, is based on his claim that meaning (or "objective information") is not something possessed by individual sentences.

Quine says that the idea that two sentences express the same proposition (and are therefore synonymous or cognitively equivalent) if they are true in the same possible worlds "affords us no general way of equating sentences in real life. For surely we can never hope to arrive at a technique for so analyzing our ordinary sentences as to reveal their implications in respect of the distribution of particles" in possible worlds (Quine, 1970a, pp. 4-5). The reason why we will never find such a technique for specifying the implications of individual sentences--i.e., for specifying the possible worlds in which individual sentences are true--is that meaning or objective information is not something possessed by individual sentences or even fairly large sets of sentences. If Quine is right, then it is a mistake to think that individual sentences may be analyzed in terms of the possible worlds in which they are true.

According to Quine, the reason why individual sentences and sets of sentences cannot be said to express separable funds of objective information is that there is an "empirical slack" or degree of underdetermination to our

physical theories. Quine explains this "empirical slack" as follows. Our physical theory is underdetermined by all past data (because a future observation can conflict with it); it is underdetermined by all past and future data combined (because some unobserved piece of data can conflict with it); and it is also underdetermined by all possible observations.⁴ Quine explicates "all possible observations" extensionally, as all those observations reported by "all the observation sentences of the language" (to which dates and positions are applied in all combinations, "without regard to whether observers were at the place and time"--1970c, p. 179).

Quine supports his thesis that physical theory is underdetermined by all possible observations with the claim that "there can be a set H of hypotheses, and an alternative set H' incompatible with H, and it can happen that when our total theory T is changed to the extent of putting H' for H in it, the resulting theory T' still fits all possible observations just as well as T did" (Quine, 1970a, p. 6). In other words, as Quine says, H and H' can be "logically incompatible and empirically equivalent" (1970c, p. 179). The point is, then, that insofar as "empirical information" can be apportioned to H and H' individually, they convey the same empirical information; and, if meaning is equated with empirical information expressed, then H and H' have the same meaning. But H and H' are logically incompatible. Since it cannot be right to say that (sets of) sentences which have the same

⁴Quine, 1970c, p. 179; y also Quine, 1970a, p. 6.

meaning are logically incompatible, Quine rejects the identification of meaning with empirical information expressed by individual sentences or sets of sentences. According to Quine, two (sets of) sentences cannot have the same meaning and be logically incompatible; yet two (sets of) sentences can be both empirically equivalent and logically incompatible. If Quine is right, any identification of meaning with empirical information expressed by (sets of) sentences--even if this is done in terms of possible worlds--cannot be right.

Quine's objection to the definition of "proposition" provided by possible world semantics rests on an inclusion of theoretical sentences in the set of sentences making up any given language. Once it is admitted that theoretical sentences are part of our language, Quine's thesis follows that we cannot uniquely specify the possible worlds in which the individual sentences of our language are true. This is so, as Quine says, "if only because the observational criteria of theoretical terms are commonly so flexible and fragmentary" (1970c, p. 179). Because theories are neither mere sets of sentences about observable entities, nor sets of sentences making simply inductive generalizations about such entities, there is what Quine calls an "empirical slack" in scientific theories. Because of this slack, it is possible for two hypotheses, or sets of hypotheses, to be logically incompatible and empirically equivalent. And because two hypotheses, or sets of hypotheses, can be logically incompatible and empirically equivalent, it cannot be right to identify meanings or propositions with

empirical information expressed by individual sentences or sets of sentences--even if "empirical information" is explicated in terms of possible worlds.

Quine's objection to the identification of propositions with empirical information expressed by individual sentences seems sound, because (a) his inclusion of theoretical sentences in the set of sentences making up any given language is certainly justified, and (b) his analysis of the empirical slack of scientific theories has as yet received no serious criticism (if anything, such criticism as there has been suggests that the empirical slack infects even the observation sentence level of theories, as well as the theoretical level). However, one apparent objection to Quine's argument against the attempted identification of propositions with empirical information expressed by individual sentences may be mentioned. The objection centers on Quine's explication of "all possible observations" in terms of "all the observation sentences of the language." This explication would be unproblematic if "language" were defined extensionally, as being constituted by a corpus of utterances, and if the subset of this corpus which consists of the observation sentences of the language were clearly specifiable. As we shall see in the second section of this chapter, however, there is no way to specify the observation sentences of a language according to Quine's theory of meaning, because Quine's definition of "observation sentence" is inconsistent with other parts of his theory. And, insofar as Quine's definition of "observation sentence" is inadequate, so is his explication

of "all possible observations" inadequate.

This objection to Quine's argument against the definition of "proposition" given by possible worlds semantics is, however, undamaging, because Quine's theory can be amended so that it escapes the objection. In fact, in Chapter 7 I show how unproblematic definitions of "language" and "observation sentence" can be given in a consistent extension of Quine's theory. The result is that Quine's argument given above against the proposed definition of "proposition" is sustained, within a theory of meaning which is a consistent extension of Quine's.

Quine's second argument against the mentalistic theory of meaning is that even if the identity conditions of propositions or meanings were clear, they would be useless. Quine argues that even if propositions were clear, they need not, and therefore ought not, be used in constructing a theory of meaning.

Quine may be interpreted as claiming, in this second criticism of the mentalistic theory of meaning, that propositions ought to be dispensed with, because they can be dispensed with: i.e., because they explain nothing that cannot be explained without them. More precisely, he thinks that the phenomena which propositions are purported to explain fall into two categories: those that ought not to be explained, by propositions or anything else, and those that can be explained in terms of sentences (and certain other concepts defined in terms of sentences), without recourse to propositions. And, if the phenomena which propositions are

purported to explain do in fact fall into either one or the other of these categories, then, argues Quine, propositions are useless, and can be dispensed with.

Quine claims that the first kind of phenomena--those that ought not to be explained--are not genuine phenomena; they are only thought to be phenomena in need of explanation by philosophers beguiled by the mentalistic theory of meaning. The first phenomenon thought to be explained by propositions is "our undeniable intuitions of synonymy and analyticity" (Quine, 1960, p. 207). Quine recognizes these intuitions, but argues that on examination it becomes obvious that our intuitions of synonymy "do not sustain a synonymy concept suited to identity of propositions, or meanings" (Quine, 1960, p. 207). That is, Quine thinks that we cannot talk in general about any two linguistic forms x and y (where x and y range over words, phrases, and sentences) as being synonymous or not; at most, we can talk about the synonymy or non-synonymy of observation sentences. Quine argues that his theory of meaning shows that the only synonymy concept which is warranted on behavioral grounds is a restricted synonymy concept applicable only to observation sentences (and to occasion sentences generally, when we are investigating intrasubjective synonymy). Our preanalytic talk of synonymy of non-observation sentences simply does not make sense on behavioral grounds, and therefore such an unrestricted concept of synonymy is simply not something that ought to be explained, by propositions or anything else.

It seems clear that Quine's criticism of the general notion of synonymy relies on his own theory of meaning; the most he claims is, as was noted, that his theory of meaning shows that the only synonymy concept which is warranted on behavioral grounds is a synonymy concept applicable only to observation sentences. Quine's arguments for this claim will be assessed shortly.

The second phenomenon that Quine thinks ought not to be explained, by propositions or anything else, is the general notion of "correct translation," as determined on grounds independent of the manual of translation used in constructing the translation in question. As with synonymy, he recognizes that we can speak of correct translations of observation sentences (Quine, 1960, p. 68), but that, as a general notion intended to apply to any pair of linguistic forms, the notion makes no empirical sense. Quine thinks that his thesis of the indeterminacy of translation shows that there is no such thing as a correct translation (of non-observation sentences), as judged on criteria independent of the particular manual used in constructing the translation. Since there is no general notion of correct translation, neither propositions nor anything else need be posited in order to explain what is shared by correct translations (Quine, 1960, pp. 207-208). As Quine says, what the thesis of the indeterminacy of translation occasions "is a change in prevalent attitudes toward meaning, idea, proposition" (1969d, p. 304); if the thesis is correct, neither meanings, nor ideas, nor

propositions are needed to explain the notion of correct translation.

The third phenomenon that Quine thinks ought not to be explained, by propositions or anything else, is the alleged synonymy of paraphrase and philosophical analysis. Quine argues that these alleged cases of synonymy are not genuine phenomena in need of explanation. He says that "even if the notion of synonymy as such were in the best of shape," "synonymy claims would generally be out of place" in connection with either paraphrase or philosophical analysis (Quine, 1960, p. 208). Because he believes it is a mistake to think that either paraphrase or analysis aims at providing synonyms, Quine argues that there is therefore no need to posit propositions as that which is shared by analysans and analysandum and by paraphrase and paraphrased.

The fourth phenomenon that Quine thinks ought not to be explained, by propositions or anything else, is the alleged meaninglessness of certain linguistic forms. It will be recalled that the mentalist claimed that all those declarative sentences which do not express propositions are meaningless. Thus, sentences like "This stone is thinking about Vienna," and "Quadruplicity drinks procrastination" are held to be meaningless, in the sense that they are "unattached labels" in the mentalist's museum: they are expressions which express no proposition (Quine, 1970b, p. 7).

Quine has argued, however, that once we see that semantic notions such as meaning must be explained, if they

can be explained at all, by reference solely to verbal behavior, it will become clear that it is impossible to distinguish meaningfulness from what Quine calls verbal "extravagance." That is, Quine argues that "on the strength of verbal behavior," without relying on any evidence not given in overt behavior, "we could as well say that" sentences like "This stone is thinking about Vienna" "are meaningful but just too extravagantly false to be worth denying" (Quine, 1970b, p. 7). In other words, Quine argues that, seen naturalistically, sentences which are alleged by the mentalist to be neither true nor false, but meaningless, are actually meaningful and extravagantly false. Quine thinks that meaningfulness is not a genuine phenomenon that requires explanation in terms of propositions; rather, what needs explanation are meaningful sentences which are extravagantly false. And, according to Quine, these sentences can be adequately explained without recourse to propositions.

Among those phenomena that Quine thinks need explanation, but that can be explained in terms of sentences rather than propositions, is that of the bearers or vehicles of truth. Propositions were wanted, by the mentalist, as truth vehicles. It was claimed that particular sentences cannot be either true or false, because, if they are, then their truth values will change both as the world changes and as the sentences get uttered by different speakers. For example, the sentence

(1) The door is open

may be true of the door located at p_1 at time t_1 , while false of the door located at p_2 at t_2 . Propositions were wanted as truth vehicles, because then what gets counted as either true or false--namely, the propositions--will not change in truth value as the world changes. But, Quine claims, we can appeal to eternal sentences as truth vehicles, rather than propositions, and achieve the same results. An eternal sentence is "a sentence whose truth value stays fixed through time and from speaker to speaker" (Quine, 1960, p. 193). For example, the sentence

(2) The door located at p_1 at t_1 is open
is an eternal sentence. According to Quine, it is eternal sentences which are either true or false; non-eternal sentences are neither true nor false. What is true or false, for Quine, are not the propositions expressed by sentences, but eternal sentences; and, in cases in which speakers utter non-eternal sentences, what is true or false is not the actual sentence uttered, but the eternal sentence "that the original speaker could have uttered in place of his original utterance in those original circumstances without detriment, so far as he could foresee, to the project he was bent on" (Quine, 1960, p. 208).⁵

The second phenomenon that Quine thinks can be adequately explained without recourse to propositions is

⁵This analysis is slightly revised by Quine, 1969c, pp. 141-144; the basic point remains, however, that propositions are not needed as truth-vehicles.

that of implication. It will be recalled that according to the mentalist propositions were wanted as the things "that stand in the logical relation of implication" (Quine, 1970a, p. 2). Now, Quine claims, quite simply, that sentences rather than propositions can be taken as the things which stand in the logical relation of implication (Quine, 1970a, pp. 48-49). Quine claims that his definition of implication--namely, that "one sentence logically implies another . . . when [the first is] logically incompatible with the other's negation" (1970a, p. 49)--is simpler and therefore preferable to any definition of implication in terms of propositions. Quine's definition is simpler, because if implication is defined in terms of propositions, then, as we have seen, propositions (or, at least, their identity conditions) must be explained in terms of sentences; and, the case of implication is one in which we may simply go with the sentences themselves, and dispense with the extra baggage of propositions.

The third phenomenon Quine thinks can be explained without recourse to propositions concerns what are known as the propositional attitudes. Propositions were wanted as the objects of these propositional attitudes. Quine has offered several different explanations of propositional attitudes, none of which have relied on the notion of proposition. In Word and Object, Quine offered a "no-object" analysis of propositional attitudes, according to which we simply "dispense with the objects of the propositional attitudes" (Quine, 1960, p. 216). According to this analysis, neither propositions

nor sentences are needed as the objects of propositional attitudes. This is not to say that we need to stop speaking about believing things, desiring things, etc., but that we formulate the propositional attitudes with the help of intensional abstraction . . . but just cease to view these notations as singular terms referring to objects" (Quine, 1960, p. 216). Thus, the "thing" I believe, desire, etc., is simply not an object, abstract or concrete. We may continue to talk about believing "things," but only so long as we recognize that such talk has a "second-grade status": it has "the status of useful vernacular having no place in the austere apparatus of scientific theory" (Quine, 1969c, p. 146).

More recently, however, Quine has liberalized his conception of scientific discourse, and has admitted that the loss to scientific discourse of the idioms of propositional attitudes is a genuine loss. In this spirit, Quine has resuscitated the objects of some of the propositional attitudes, and claimed that stimulus meaning (v 2.3) may serve as the objects of these attitudes (Quine, 1969c, p. 156). Thus, Quine still believes that propositional attitudes can be adequately explained without recourse to propositions.

Quine's third and final criticism of the mentalistic theory of meaning is that the talk about ideas and propositions lends to concepts such as synonymy and translation an air of determinacy which is not really there. As we saw, "according to the museum myth the words and sentences of a

language have their determinate meanings. To discover the meanings of the native's words we may have to observe his behavior, but still the meanings of the words are supposed to be determinate in the native's mind" (Quine, 1969c, pp. 28-29). Quine's claim is that concepts such as synonymy and correct translation are in fact indeterminate, and that the mentalistic theory of meaning merely masks the underlying indeterminacy of such notions. And, Quine argues that his own naturalistic theory accurately reveals the indeterminacy surrounding the notion of meaning.

Quine's arguments for the thesis that the notions of synonymy and translation (incidentally, bilingual translation here is construed merely as bilingual synonymy--i.e., interlinguistic synonymy--and "translation in the home language" [Quine, 1969c, p. 47] is mere synonymy--i.e., intralinguistic synonymy) are indeterminate have already been introduced, and will be dealt with shortly; they are identical with his arguments for the claim that synonymy and translation are not phenomena that need to be explained, by propositions or anything else. Quine's general claim is simply that the notions of synonymy and translation won't wash; one problem with propositions is that they are not needed (because the notions they are supposed to explain, such as unrestricted synonymy and correct translation, won't wash), and another problem with propositions is that they make the notions of synonymy and translation look like they will wash.

In evaluating Quine's arguments from indeterminacy of

synonymy and translation for the thesis that propositions ought to be dispensed with, it is not necessary to examine the arguments in any detail. It is sufficient for our purposes to note that translation and synonymy can only be shown to be indeterminate if propositions have been previously dispensed with and ruled as unavailable. This fact constitutes a major difficulty for Quine's third criticism of the mentalistic theory of meaning.⁶

As we have seen, Quine says that one trouble with propositions is that they lend the notions of synonymy and translation an air of determinacy. This is misleading, according to Quine, because synonymy and translation are, in fact, indeterminate. And, because a belief in propositions obscures this indeterminacy, propositions ought to be dispensed with. But, it is also the case that synonymy and translation can only be shown to be indeterminate if propositions are dispensed with:

For an uncritical mentalist, no such indeterminacy threatens. Every term and every sentence is a label attached to an idea, simple or complex, which is stored in the mind. When on the other hand we take a verification theory of meaning seriously [i.e., dispense with propositions], the indeterminacy would appear to be inescapable (Quine, 1969c, p. 80).

And again:

In the old days when the idea idea was supreme, there was a comfortable illusion of determinacy of translation. To understand a language was to get its labels on to the right ideas . . . As soon as we recognize that there is nothing in meaning that

⁶This difficulty was pointed out to me by D. J. Hockney.

is not in behavior, on the other hand [i.e., as soon as we dispense with ideas and propositions], we are bound to expect ultimate indeterminacies of translation (Quine, 1970b, p. 10).

Quine has stressed the point several times, that the "crucial consideration" behind his argument for indeterminacy of synonymy and translation is the assumption that the basic hypothesis of propositionalism--viz., that a statement about the world has a "separable fund of empirical consequences [i.e., meaning] that it can call its own" (Quine, 1969c, p. 82)--is false.⁷ Furthermore, in his original arguments for indeterminacy of translation, in Chapter II of Word and Object, Quine acknowledges that he is assuming, ex hypothesi, that propositions are unavailable: he says that "in this chapter we shall consider how much of language can be made sense of in terms of its stimulus conditions [i.e., without recourse to propositions], and what scope this leaves for empirically unconditioned variation [i.e., indeterminacy] in one's conceptual scheme" (Quine, 1960, p. 26).

Thus, propositions are deplored by Quine as lending determinacy to the notions of synonymy and translation; but, the only way to show that synonymy and translation are indeterminate is to dispense with propositions. It follows that Quine cannot, on pain of circularity, base his case against propositions solely on his case for indeterminacy of translation, because his case for indeterminacy of translation relies on a prior dismissal of propositions.

⁷V Quine, 1970a, p. 3 and Quine, 1969c, pp. 28, 29-30, 80, and 82.

If Quine's claim that propositions ought to be dismissed is well-founded, then there must be other reasons, independent of his arguments for the indeterminacy of translation, for dismissing propositions. We have already seen that Quine's first objection to propositions--that they lack identity conditions--was sound; therefore it provides one reason, independent of his indeterminacy thesis, for dismissing propositions and the whole mentalistic theory of meaning. In the remainder of this section, I argue that Quine's second objection to propositions is similar to his third objection, insofar as it does not provide a reason, independent of Quine's own theory of meaning, for dismissing propositions.

Quine's second objection was that propositions ought to be dispensed with, because they are not needed; they are not needed, because the phenomena they are thought to explain are either not genuine phenomena at all, or are adequately explained without recourse to propositions. The four phenomena that Quine contends are not genuine phenomena in need of explanation are unrestricted synonymy, correct translation (as judged on grounds independent of the manual of translation used in constructing the translation), synonymy with respect to paraphrases and philosophical analysis, and meaninglessness. As was noted, however, Quine relies on his arguments for indeterminacy of translation and synonymy (i.e., translation "in the home language") in his arguments for the thesis that synonymy and translation

are not genuine phenomena in need of explanation; in fact, his arguments for indeterminacy of translation are identical with his arguments for the claim that unrestricted notions of synonymy and correct translation need not be explained, by propositions or anything else.⁸ And, since Quine's arguments for indeterminacy of translation rely, as we have seen, on a prior rejection of propositions, it follows that his arguments for the claim that general notions of synonymy and correct translation need not be explained by propositions or anything else also relies on a prior rejection of propositions.

Quine also claims that the relation of synonymy alleged to obtain in paraphrases and philosophical analyses is not a genuine phenomenon in need of explanation. His argument for this claim was that "even if the notion of synonymy were in the best of shape" synonymy claims would be out of place in connection with either paraphrase or philosophical analysis. But, this argument is undamaging to the propositionalist. Granted, it may remove one motivation for positing propositions--namely that they can be used to explain the synonymy of paraphrase with paraphrased and of analysandum with analysans--but it does not show there to be anything wrong with propositions themselves. Quine says that synonymy claims are out of place with respect to paraphrase and philosophical analysis, even if the notion of synonymy were in the best of shape; and, the propositionalist

⁸v pp. 46 and 47 above.

may agree, adding that synonymy is in the best of shape, thanks to propositions. Quine's argument against synonymy with respect to paraphrase and philosophical analysis is undamaging to the theory of propositions.

The fourth phenomenon that Quine thinks is not a genuine phenomenon in need of explanation is the alleged meaninglessness of certain declarative sentences. It will be remembered that the mentalist claims that sentences like "This stone is thinking about Vienna" are meaningless because they express no propositions. According to the museum myth of meaning, meaningless expressions could be thought of as "unattached labels": as long as the museum myth persisted, the meaninglessness of certain declarative sentences could be thought of as a phenomenon explained by the fact that the sentences expressed no proposition (Quine, 1970b, p. 7). But, argues Quine, once propositions are dispensed with, and it is agreed that there is nothing to meaning (or meaninglessness) that is not in verbal behavior, then "we cannot easily distinguish it [meaninglessness] from mere extravagance" (1970b, p. 7). As we have seen, Quine argues that "on the strength of verbal behavior we could as well say that [sentences like "This stone is thinking about Vienna"] are meaningful but just too extravagantly false to be worth denying" (1970b, p. 7).

It should be clear, however, that Quine has merely shown that according to his theory of meaning meaninglessness is indistinguishable from extravagance. That is,

meaninglessness can only be shown to be not a genuine phenomenon in need of explanation if one has dispensed with propositions. Only if propositions have been rejected, and only if meanings have been identified, a la Quine, with aspects of verbal behavior, will it follow that verbal extravagance, and not meaninglessness, is the relevant phenomenon in need of explanation.

Finally, Quine claims that there are three phenomena which are in fact in need of explanation, but which can be adequately explained without recourse to propositions. These three phenomena concern truth, the objects of the propositional attitudes, and the logical relation of implication. As we have seen, both the propositionalist and Quine have their own theories as to what to count as truth vehicles, as the objects of the propositional attitudes, and as the things related in the logical relation of implication. Now, although it is impossible to substantiate this claim by any means save inspection of the relevant texts,⁹ it is fair to say that in developing his own theories as to what to count as truth vehicles, as the objects of propositional attitudes, and as the things related in the relation of implication, Quine adds no arguments to those already reviewed against propositions. That is, Quine claims that

⁹Quine, 1960, pp. 207ff. "Let us review the situations that prompted the positing of propositions, and consider what can be done without that expedient;" and Quine, 1970a, pp. 48-49, in which it is merely assumed that the things related in the logical relation of implication are sentences and not propositions.

his theories as to what to count as truth vehicles, the objects of propositional attitudes, and the things related in the relation of implication, are, in some sense, more adequate than the mentalist's theories; but, he adduces no grounds against the mentalist's identification of propositions with truth vehicles, the objects of propositional attitudes, and the things related in the relation of implication, other than the grounds adduced in his other objections to propositions. Thus, in his defense of his own theories as to what to count as truth vehicles, the objects of propositional attitudes, and the things related in the relation of implication, Quine offers the mentalist no new reasons for abandoning his propositions.

I conclude this section, therefore, with the observation that two of Quine's three objections to the mentalistic theory of meaning are undamaging. Neither Quine's argument from the contention that propositions are not needed, nor his argument from his indeterminacy thesis, provide any reason for abandoning the mentalistic theory of meaning. Only Quine's argument for the thesis that propositions lack clear identity conditions supports his rejection of the mentalistic theory.

2.2 Language and Dispositions to Verbal Behavior

It should be clear, then, that Quine assumes, in his attempt to develop a theory of meaning, that propositions and mentalistic ideas will be of no use to his theory. As

Quine says, "surely one has no choice but to be an empiricist so far as one's theory of linguistic meaning is concerned" (1969c, p. 81); and, this may be understood as claiming that ideas and propositions are useless for the theory of meaning. More specifically, Quine assumes, with Bloomfield, that the only evidence the linguist has to work with in developing a theory of meaning (or of phonetics, syntax, or anything else connected with language, for that matter), consists in (1) a corpus of utterances, (2) the speakers' situations, and (3) the hearers' responses (Bloomfield, 1935, pp. 23ff.).

As we have seen, Quine begins his chapter in Word and Object on "Translation and Meaning" with the acknowledgment that his purpose in the chapter is to "consider how much of language can be made sense of in terms of its stimulus conditions." In other words, Quine's program in much of Word and Object is to determine how much of our everyday conception of language can be made sense of in behavioral terms. Quine relies on certain kinds of behavioral evidence in developing his theory of meaning; and if, on these terms, some of our preanalytic notions concerning meaning cannot be made sense of, then so much the worse for our preanalytic notions (Quine, 1969c, p. 29).

The first notion that Quine attempts to make sense of in terms of the kind of behavioral evidence which he admits is the notion of language itself. Quine's analysis of the notion of language will serve as an instructive introduction to his analysis of stimulus meaning, because,

as I shall show, his theoretical notion of language shares many problems with his notion of stimulus meaning. Quine himself seems to have doubts about whether he has succeeded in giving an adequate analysis of the notion of language; thus, for example, Quine says that "a trouble with the notion of a language is that it, like the very notion of proposition or meaning that I have complained about, has been given no satisfactory principle of individuation" (1969c, p. 142).¹⁰ Nevertheless, Quine goes on to rely on some notion of language [for example, in his discussion of translation (1960, p. 28, and elsewhere), in his discussion of the eternal sentences of a language as truth vehicles (1970a, p. 14), and in his discussion of the objects of propositional attitudes (1969c, p. 144)], and suggests that a principle of individuation for languages can be given, which is sufficient for his purposes.

Basically, Quine defines the language of a speaker as the speaker's dispositions to respond verbally to stimulations during a certain time. Thus, Quine says that "a man's current language" is constituted by "his current dispositions to respond verbally to current stimulation" (1960, p. 28); and again, language is "the complex of present dispositions to verbal behavior" (1960, p. 27). What to count as "current stimulation" is determined by what Quine calls the "modulus

¹⁰L. J. Cohen has been convinced by Quine's disclaimer; Cohen claims that Quine has simply offered no principle of individuation for languages--a claim that I shall show is false; v Cohen, 1962, p. 359.

of stimulation." All stimulations that fall within the modulus count as current: a man's current language is constituted by his current dispositions to respond verbally to stimulations within the modulus. His current verbal dispositions constitute his "language in use"; all past verbal dispositions to respond to past stimulations outside the modulus constitute his "language in acquisition" (Quine, 1960, p. 28).

Obviously, in any given modulus, a speaker will (presumably) manifest only a few of the dispositions to verbal behavior he has in that modulus. Quine has suggested that one way a linguist could determine a speaker's complete (i.e., manifested and non-manifested) dispositions to verbal behavior during a given modulus would be to gather "indirect evidence" for the dispositions, "from which we reason according to plausible psychological theories and generalizations regarding the persistence of habits and other matters" (1969c, p. 144). According to this method, a speaker's language in use during a certain modulus can be reconstructed on the basis of his actual utterances during that modulus plus indirect evidence for what might have been uttered during that modulus (were the stimulations different), which is inferred from past utterances with the help of certain psychological theories.

We may assume, then, that a speaker's current language consists of more than just the utterances actually uttered during the modulus; it consists of his dispositions to

respond verbally to current stimulations. Therefore, the identity conditions for languages are as follows: two speakers (or one speaker in two different moduli) are speaking the same language if and only if during particular moduli they have the same dispositions to verbal behavior to respond to the same stimulations.

Quine notes that it is important to construe the "stimulations" that prompt utterances in speakers as stimulations of nerve receptors rather than with the sensed objects in the world. Thus, for example, what prompts Quine's natives to utter "Gavagai" are "stimulations and not rabbits" (1960, p. 31). There are, therefore, as many different generic kinds of stimulations as there are senses: e.g., visual stimulations, auditory stimulations, etc. "A visual stimulation is perhaps best identified, for present purposes, with the pattern of chromatic irradiation of the eye" (Quine, 1960, p. 31); and the other kinds of stimulations get identified with the other senses (Quine, 1960, p. 33).

If two speakers (or one speaker in two moduli) are to speak the same language, they must have during particular moduli the same verbal dispositions to respond to the same stimulations. And, Quine has come to realize that this requirement concerning same stimulations renders it implausible, if not impossible, that two speakers ever speak the same language. The difficulty is that since stimulations are identified with firings of nerve ends, the notion of "same stimulation" presupposes complete homology of these

nerve ends. And, as Quine says, "if we construe stimulation patterns my way, we cannot equate them without supposing homology of receptors; and this is absurd, not only because full homology is implausible, but because it surely ought not to matter" (1969c, p. 158). Because of the implausibility of assuming intersubjective homology of nerve receptors, it is clear, as John Wallace has remarked, "that no two persons ever receive the same pattern of stimulation" (1971, p. 148). Because, according to Quine, languages are identical if and only if speakers have the same verbal dispositions to respond to the same stimulations, and because no two speakers ever receive, in Quine's sense, the same stimulations, it follows that no two people can ever speak the same Quinian language.¹¹ We shall have occasion to examine this conclusion in detail later; as we shall see, the implausibility of two speakers ever receiving the same Quinian stimulations, and the attendant implausibility of believing that two speakers could ever speak the same Quinian language, render highly suspect Quine's repeated claim that according to his theory there can be no private languages (y, e.g., 1969c, p. 27; and 1970b, p. 5).

I see no way of extricating Quine's theory from its involvement with private languages, so long as stimulations are identified with firings of nerve receptors. As we saw, Quine thinks that there are two things wrong with presupposing

¹¹This is so, despite Quine's assumption to the contrary; cf. 1960, p. 27.

homology of nerve ends in any judgment of stimulus (and hence language) identity: (1) such a presupposition is implausible, and (2) such a presupposition ought not to be relevant to judgments of linguistic identity. Quine is ready to act on his first worry, that a presupposition of complete homology of nerve ends is implausible; he acknowledges that, because of this implausibility, it would be better to speak of similarity rather than identity of stimulations (and hence languages) (Quine, 1969c, pp. 159-160). But, oddly enough, Quine is not willing to act on his second worry, that homology of nerve ends ought not to be relevant to the identity conditions of languages. Quine is not willing to give up his identification of stimulations with firings of nerve ends; and, so long as he maintains this identification, homology of nerve ends will be relevant to the identity conditions of languages. In Chapter 6 I will propose a revision of Quine's theory of meaning which, among other things, rejects Quine's identification of stimulations with firings of nerve ends.

One problem with Quine's definition of language, therefore, concerns his construal of "stimulations": because stimulations are identified with firings of nerve receptors, it is implausible to suppose that any two speakers ever speak the same language, according to Quine's definition.

A second problem, which has been noted but misinterpreted on occasion, is that the identity conditions for languages implicit in Quine's definition of language do not

classify as multilingual those people which we want to say, preanalytically, are multilingual. According to Quine's definition of language, no one is multilingual in the sense that no one can speak two or more different languages during the same modulus (even though, as we shall see, everyone is multilingual in the sense that it is highly likely that the language one speaks in one modulus is different from the languages one speaks in all other moduli). No one can speak two or more languages in the same modulus simply because everything that a speaker says or is disposed to say during that modulus constitutes his language for that modulus. But, there are certain circumstances in which a person we want to call (e.g.) a bilingual will be disposed to utter, in one modulus, what we want to say are utterances in two different languages. As Cohen has argued, what Quine needs to do, if he wants to bring his theoretical notion of language into conformity with our preanalytic notion, is "to define the conditions under which a speaker is neither disposed to innovate nor to speak a foreign language; and this would be in effect to define the conditions under which a speaker uses the same or a different language" (Cohen, 1962, p. 359). Cohen says that Quine has failed to "find a principle of individuation for languages"--i.e., to state their identity conditions. But this claim of Cohen's is simply false; Quine has, as we have seen, stated identity conditions for his theoretical notion of language. Cohen's complaint should be that Quine has failed to define "language" in a

way that conforms to ordinary usage.

A second way in which Quine's theoretical notion differs from the preanalytic notion of language is that it is highly unlikely that one speaker can ever speak the same Quinian language twice (i.e., in two different moduli). This problem with Quine's notion of language stems, again, from his identification of stimulations with firings of nerve receptors. As we have seen, for one person to speak the same language twice, he must, among other things, receive the same set of stimulations twice; and as Wallace has pointed out, "if 'pattern of stimulation' is defined in terms of geometrical arrangement of firings in the nerve net, then it is clear that . . . no person ever receives the same pattern twice" (Wallace, 1971, p. 148).¹² And, if no person ever receives the same pattern of stimulation twice, then no one can speak the same Quinian language twice. This conclusion supports my earlier suggestion that Quine's theory of meaning is involved with the notion of private languages: there seem to be private languages for Quine in the sense that, first, no two people can ever speak the same Quinian language, and, second, no one person can ever speak the same Quinian language twice.

It might seem that one way in which Quine could have

¹²Although this conclusion seems plausible enough, it would be more accurate to say that the chances of anyone ever receiving the same pattern of stimulation twice are almost nil; although it is not impossible that anyone could speak the same Quinian language twice, it is highly unlikely.

given a theoretical account of language, which conforms more closely with our intuitions on the matter, would be to define languages in terms of meanings. Thus, it would seem that Quine could have relied on the plausible generalization that in a language like linguistic forms are alike in meaning (except in cases of ambiguity, in which linguistic forms have more than one meaning). According to this general principle, a linguistic community could be defined as, roughly, that set of speakers which, for any given linguistic form, whenever a speaker in the set is disposed to utter that form, the form has the same or a similar meaning (except in cases of ambiguity) that it has for any other speaker in the set. And, a language could be defined, roughly, as the set of dispositions to verbal behavior of all members of a linguistic community during a given modulus.

Such a definition of language is not without its problems, of course. It has the advantage, however, which Quine's definition lacks, of making the notion of language depend on the notion of meaning. That is, it rests on the insight that two people are speaking the same language (or, that one person is speaking the same language twice) if, whenever they (or he) utter(s) the same linguistic form, that linguistic form has the same or similar meaning for both speakers (or the speaker). Thus, it might seem that once meaning has been defined in terms congenial to Quine, then definitions, more adequate than Quine's original ones, of linguistic community and of language could be given in

terms of meaning. I shall now show, however, that this promising approach is not open to Quine; I shall argue that Quine cannot use his notion of stimulus meaning to develop an analysis of language which makes languages enduring objects with boundaries extending over many people over a long period of time.

2.3 Stimulus Meaning

Basically, the reason why Quine cannot use his notion of stimulus meaning in an analysis of language of the kind wanted is that stimulus meanings do not lend themselves to intersubjective and inter-modular (in the intrasubjective case) comparison any better than do languages. And, the reason why neither stimulus meanings nor languages are neither intersubjectively nor inter-modularly comparable is that both rely on Quine's construal of stimulations as firings of nerve ends.

Quine takes sentences to be the basic unit of meaning, and defines the stimulus meaning of a sentence as the ordered pair of the affirmative stimulus meaning and the negative stimulus meaning of that sentence. And, the affirmative stimulus meaning of a sentence for a speaker is defined as "the class of all the stimulations (hence evolving ocular irradiation patterns between properly timed blindfoldings) that would prompt his assent" to the sentence when queried (Quine, 1960, p. 32). And,

a stimulation σ belongs to the affirmative stimulus meaning of a sentence S for a given speaker if and only if there is a stimulation σ' such that if the speaker were given σ' then were asked S, then were

given S, and then were asked S again, he would dissent the first time and assent the second (Quine, 1960, p. 32).

Negative stimulus meaning is defined analogously, with "assent" and "dissent" interchanged.

Quine's definition of stimulus meaning is essentially Bloomfieldian in approach. As we have seen, Quine construes acts of verbal behavior in the same way as does Bloomfield. For both, all episodes of verbal behavior consist of three distinct parts: "A, the speaker's situation; B, his utterance of speech sounds and its impingement on the hearer's eardrums, and C, the hearer's response" to the utterance (Bloomfield, 1935, p. 74). And, Bloomfield identifies the meaning of an utterance as "the important things with which the speech utterance . . . is connected, namely the practical events (A and C)" (1935, p. 27). Although strictly speaking the meaning of an utterance is constituted by the speaker's situation and the hearer's response, Bloomfield acknowledges that in practice the meaning of an utterance may be identified simply with the speaker's situation (Bloomfield, 1935, p. 139). Thus, in practice, the meaning of an utterance is, for Bloomfield, the situation--or, the "stimulus," since as for Quine situations are thought of as causally related to utterances (Bloomfield, 1935, p. 23)--which prompts the utterance. And, just so for Quine, for whom the affirmative stimulus meaning of a sentence for a speaker is the class of all those stimulations that would prompt the speaker's assent to the sentence when queried.

Bloomfield recognizes, however, that "the situations which prompt us to utter any one linguistic form, are quite varied; philosophers tell us, in fact, that no two situations are ever alike" (1935, p. 140). Because of this fact, and because Bloomfield identifies meanings with the situations which prompt speech utterances, it follows that no two utterances (including no two utterances of the same linguistic form) ever have the same meaning for Bloomfield. Bloomfield attempts to resolve this difficulty by saying that we must distinguish between what he calls the "non-distinctive" and "distinctive" (or "semantic") features of any situation if we are to identify in any fruitful way the meaning of an utterance (1935, p. 141). Thus, Bloomfield argues that not all features of a situation will be relevant to the meaning of the utterance which is prompted by that situation; only the semantic features are relevant. And, even though no two situations are identical, Bloomfield claims that the semantic features of different situations may be identical. In cases in which two utterances have been prompted by situations whose semantic features are identical, the utterances may be said to have the same meaning.

Bloomfield proposes that the meaning of any given utterance can be identified with the semantic features which are in common to all the diverse situations which prompt the utterance. Whether Bloomfield's methods for distinguishing between the "distinctive" and "non-distinctive" features of situations are adequate is a question which lies beyond the

scope of the present thesis. It is sufficient to note that Bloomfield's distinction between the "distinctive" and "non-distinctive" features of situations is drawn in order to make it possible for two or more utterances (of one or more linguistic forms) to have the same meaning.

It should be clear, because Quine identifies stimulus meanings with the stimulatory situations of utterances, that his notion of stimulus meaning faces the same problems as does Bloomfield's notion of meaning: viz., that since no two stimulatory situations are identical, no two stimulus meanings can be identical. Bloomfield construes stimulatory situations as consisting of what I call "stimuli"--i.e., sensed spatio-temporal objects. Quine, on the other hand, construes stimulatory situations as consisting of what I call "stimulations"--i.e., firings of nerve receptors. Thus, the problem of equating stimulatory situations for Quine reduces to the problem of intersubjective and inter-modular comparison of firings of nerve receptors. That the notions of intersubjective and inter-modular comparison of firings of nerve receptors are problematic renders problematic the claim that stimulus meanings may be compared and found to be either the same or different. Although both Quine and Bloomfield face the problem that since no two stimulatory situations are identical, no two meanings (or stimulus meanings) can be identical, the reasons why both Quine and Bloomfield face this problem are different. For Quine, the problem arises from his construal of stimulus situations

as stimulations rather than stimuli.

Quine claims that there are cases in which we can talk, with justification, about stimulus meanings being intersubjectively the same or different. In fact, Quine's definition of observation sentence relies on a notion of intersubjective comparison of stimulus meanings; and, since, as Quine says, "the observation sentence is the cornerstone of semantics" (1969c, p. 89), it is of particular importance to Quine's semantic theory that intersubjective comparison of stimulus meanings be possible. I shall argue, however, that Quine's claim that there are cases in which it is possible to intersubjectively compare stimulus meanings is inconsistent with his definition of stimulus meaning.

Quine says that "the behavioral definition of an observation sentence is as follows: an observation sentence is a sentence whose stimulus meaning is the same for just about all speakers of the language" (1972, p. 451). That is, the distinguishing mark of observation sentences is constancy of stimulus meaning over a large number of speakers of one language (assuming, for the moment, that there can be a large number of speakers of one Quinian language). Obviously, this mark for distinguishing between observation and non-observation sentences presupposes an effective method of intersubjective comparison of stimulations. But, as we have seen, intersubjective comparison of stimulations presupposes full homology of nerve ends, which Quine admits is an "implausible myth" (1969c, pp. 158-159). Therefore, Quine's

definition of "observation sentence" is no less mythical than is the presupposition of full homology of nerve ends. Specifically, the implausibility of the presupposition of full homology of nerve ends renders implausible Quine's claim that one sentence can have the same stimulus meaning for two or more speakers.

The general point is, then, that because stimulations are identified with firings of nerve ends, it is impossible for any two linguistic forms to have the same meaning for two or more speakers. That is, Quine's definition of stimulus meaning (and especially his construal of stimulations as firings of nerve ends) precludes the possibility of any two linguistic forms (which may be phonetically identical) being intersubjectively stimulus synonymous (i.e., have the same stimulus meaning). And, because observationality of sentences is defined in terms of constancy of stimulus meaning over a wide range of speakers, it follows that no sentences can be identified as observation sentences. Quine's distinction between observation and non-observation sentences, and, more generally, any notion of intersubjective stimulus synonymy, is inconsistent with his definition of stimulus meaning.

Not only is it impossible for any two linguistic forms to be intersubjectively stimulus synonymous; in addition, it is highly unlikely that any one linguistic form will have the same stimulus meaning (for one or more speakers) twice--i.e., in two different moduli. As we have

seen, since stimulations are identified with firings of nerve receptors, it is highly unlikely that anyone would ever receive the same stimulation twice. Because of this fact, it is also unlikely, on Quine's definition, that any utterance will ever have the same stimulus meaning twice. Uttered in one modulus, the utterance will be prompted by a certain stimulation; and it is highly unlikely that the speaker will ever again receive the same stimulation. Therefore, if the original utterance should, somehow, come to be uttered again, it is highly unlikely that it will have the same stimulus meaning it had on its first occasion of utterance.

2.4 Private Languages and Private Meanings

It was noted earlier that Quine's construal of stimulations as firings of neural receptors renders somewhat suspect his claim that according to this theory of language there can be no private languages; this point may now be clarified. When Quine claims "that there cannot be, in any useful sense, a private language" (1969c, p. 27), he is echoing Dewey's point that there is nothing to language--and, in particular, nothing to meaning--that is not given in overt behavior (or in dispositions to overt behavior). Meanings are public for Quine in the sense that, in Dewey's terms "meaning . . . is not a psychic existence; it is primarily a property of behavior" (Quine, 1969c, p. 27, quoting Dewey, Experience and Nature, p. 179). That is, meanings are not unobservable entities located in the minds of speakers;

there are no private meanings or private languages (i.e., languages made up of sentences having (possibly idiosyncratic) meanings in speakers' minds) in this sense, for Quine.

There does seem to be a sense, however, in which meanings and languages are private, for Quine. As we saw, languages are private in the sense that, although languages are constituted by sets of publicly observable sentences, it is impossible for two people to ever speak the same language, and unlikely that any one person will ever speak the same language twice (i.e., in two different moduli). And, meanings are private in the sense that, although meanings are not unobservable mental entities, it is impossible for two utterances to ever be intersubjectively stimulus synonymous, and unlikely that any one linguistic form will have the same stimulus meaning on two different occasions of utterance.¹³ As I have already argued, I see no way to extricate Quine's theory of meaning from its involvement with private languages and meanings, in the above sense, so long as stimulatory situations are interpreted as firings of nerve receptors.

2.5 Stimulus Meanings, Internal States, and Verbal Stimuli

One standard objection to Quine's theory of language is that it is too narrowly conceived, in the sense that it

¹³The situation described in Quine's thesis of the indeterminacy of radical translation seems to constitute another way in which Quine's theory of meaning is committed to a behavioristic "analogue" to mentalistic private languages; v Quine, 1960, p. 79 and Wilder, 1971, pp. 34ff.

places unrealistic and unwarranted restrictions on the type of evidence held to be available to the linguist and language learner. Most such objections are raised against Quine's theory of linguistic acquisition; for example, Chomsky has argued that modern empiricist theories of language learning--of which Quine's is an example--given in terms of conditioning and certain principles of generalization, are inadequate to account for the facts of language learning.¹⁴ Empiricist theories are alleged to be inadequate because they fail to incorporate certain kinds of mechanisms which are held to be necessary to account for language learning--i.e., they deny access to the language learner of evidence or mechanisms held to be necessary to language learning.

I believe it is also possible to criticize Quine's theory of meaning, as well as his theory of linguistic acquisition, for placing unreasonable restrictions on what is included in the notion of stimulus meaning. In this section I describe two kinds of data that Quine either ignores or rejects in his description of stimulus meaning, and argue that these two kinds of data must be included in any plausible naturalistic account of meaning. In Chapter 7, I propose a revision of Quine's theory of meaning which extends the theory to include reference to the kinds of data discussed here, which are properly included in the notion of stimulus meaning.

The first kind of data rejected by Quine concerns the internal states of individual speakers, which must be

¹⁴Chomsky, 1965, pp. 47ff.

taken into consideration in a study of the meanings of utterances.¹⁵ That Quine rejects such data is made clear in his definition of stimulus meaning: stimulus meanings consist of sets of stimulations, which are distinguished from both sensed spatio-temporal objects and neurophysiological events "deep inside the body" (Quine, 1969c, p. 158). Quine's reasons for rejecting speakers' internal states as elements of the stimulatory situations of utterances are more obscure, for two reasons. First, and most importantly, Quine nowhere argues that there are no distinctive internal naturalistic causes of verbal behavior.¹⁶ And, second, Quine seems to acknowledge that in fact the internal states of speakers may be dissimilar from speaker to speaker. Such an acknowledgment appears in statements such as the following:

Different persons growing up in the same language are like different bushes trimmed and trained to take the shape of identical elephants. The anatomical details of twigs and branches will fulfill the elephantine form differently from bush to bush, but the overall outward results are alike (Quine, 1960, p. 8).

Recognizing, as he does, that the internal states of speakers may be different, and that such internal states are connected somehow with verbal behavior, it seems odd that Quine does not admit that the connection between the internal states and the outward behavior is causal. If the connection is

¹⁵That such internal states must be included in the notion of stimulus meaning is argued by Karl Schick in his recent paper "Indeterminacy of Translation" (1972); the same point was made by C. A. Hooker in criticism of an earlier version of this thesis.

¹⁶y Schick, 1972, p. 829.

causal, then speakers' internal states must be included in the notion of stimulus meaning--since stimulus meanings are supposed to include those elements of speakers' situations which prompt assent or dissent to sentences when queried.

That speakers' internal states are causally related to dispositions to assent or dissent to sentences when queried is shown by the fact that two speakers under identical current stimulation (in Quine's sense) may utter different sentences; also, one speaker of a language may dissent from a sentence when queried, in conjunction with a certain stimulation, while a second speaker assents to the same sentence when queried under the same stimulation. Such phenomena must be attributed to differences in internal states of the speakers, and it must be agreed that these different internal states prompt the different utterances--since something prompts the utterances, and the Quinian stimulations do not prompt them, the stimulations being identical.

Having established that there are cases in which something other than Quinian stimulations prompt utterances, let us call such prompting non-Quinian stimulations "internal states" of speakers. Hence, no restriction is placed on the particular content or location of such internal states--with one important exception. The exception is that the internal states being discussed here are not mental phenomena; there is no reason to suppose that the internal states are not naturalistic objects and/or processes, as defined in Chapter 1. The internal states are not fundamentally different from

the "outer" causes of verbal behavior, in the sense that changes in them are subject to the same kind of explanation as are changes in the "outer" causes.

The internal states of speakers which are the partial causes of verbal behavior are themselves determined in two ways. First, internal states are determined genetically: speakers' initial internal states may be different. Second, internal states are determined by successive "outer" stimulation (i.e., stimulation of nerve receptors), both verbal and non-verbal, and by earlier internal states. Therefore, verbal behavior is determined by current "outer" stimulation and inner states which have themselves been shaped (in part) by "outer" stimulation. Because of the differences in speakers' initial internal states, successive "outer" stimulations may be presumed to effect different speakers' internal states in different ways.

It was already noted that there is no need to construe speakers' internal states as mental states; neither is there any reason to suppose that they are private.¹⁷ Rather, the internal states which are part causes of verbal behavior are to be construed as neurophysiological states of speakers, which are (a) subject to the same physical laws as other natural objects and processes, and (b) in principle open to public inspection. Knowledge of speakers' internal states is not private to the speakers in the states, because the states are shaped publicly in known (or knowable) ways:

¹⁷As Schick assumes, without argument, 1972, p. 828.

i.e., by genetic factors, "outer" stimulations, and other internal states.

The basic point is, then, that the stimulations that prompt speakers to assent and dissent to sentences when queried are speakers' internal states in conjunction with "outer" stimulations. Speakers' internal states and "outer" stimulations are each partial causes of verbal behavior. Since there is nothing non-naturalistic about internal neurophysiological states of the sort discussed here, such internal states must be included in the notion of stimulus meaning.

The second kind of data which Quine rejects in his description of stimulus meaning concerns correlations of sentences with verbal stimulations. Quine makes the distinction between verbal and non-verbal stimulations, and argues, as I shall show, that since correlations of sentences with verbal stimulations cannot be preserved in translation, therefore such correlations ought to be excluded from the concept of stimulus meaning.¹⁸ Although I agree that a

¹⁸Hans Herzberger, in comments on an earlier draft of this thesis, has questioned this interpretation of Quine's position. According to Herzberger, Quine only denies the usefulness of verbal stimulations to the linguist engaged in radical translation; he does not exclude verbal stimulations from the concept of stimulus meaning. In what follows, I present textual evidence which supports my interpretation of Quine's position, and which tends to disconfirm Herzberger's interpretation. Nevertheless, since I argue that Quine ought to include verbal stimulations in the concept of stimulus meaning, if Herzberger is right, that Quine does include verbal stimulations, then so much the better for Quine's original theory. If, on the other hand, Quine does exclude verbal stimulations from the concept of stimulus meaning, then my recommendation that verbal stimulations ought not to be so excluded takes on the significance which I think it in fact has.

distinction can of course be made between verbal and non-verbal stimulations, I shall argue that verbal stimulations must be included in the concept of stimulus meaning. I shall argue that by including verbal stimulations in the concept of stimulus meaning, the notion of stimulus meaning can be considerably enriched.¹⁹ Although in this section the direction of this enrichment of the notion of stimulus meaning is suggested, the discussion remains sketchy; the details will be filled in later, after certain concepts necessary to the exposition have been introduced.

The affirmative stimulus meaning of a sentence for a speaker has been defined as the set of all stimulations which would prompt the speaker's assent to the sentence when queried. It was noted that stimulations are understood as firings of neural receptors; therefore, stimulations may be divided into different kinds, depending on which kind of neural receptors are being fired. Thus, visual stimulations, for example, are defined as "ocular irradiation patterns" (Quine, 1960, p. 31); and, we can "bring the other senses on

¹⁹Schick has supported my interpretation of Quine's position, arguing that according to Quine "stimulations must be nonverbal" (Schick, 1972, p. 820). Schick has also argued that Quine's notion of stimulus meaning can and ought to be enriched "by permitting affirmative and negative stimulus meanings to contain both nonverbal and verbal stimulations" (Schick, 1972, p. 825). This is the same contention which I make in this section and the final chapter of the thesis.

Due to the similarity of Schick's argument with mine, it is worth remarking that my arguments concerning the desirability of enriching Quine's notion of stimulus meaning in the way indicated were worked out independently of Schick's. The arguments presented here were worked out, with the help of D. J. Hockney, in 1970; Schick's paper was published in 1972.

a par with vision, identifying stimulations not with just ocular irradiation patterns but with these and the various barrages of other senses, separately and in all synchronous combinations" (Quine, 1960, p. 33). Thus far, then, auditory stimulations are "on a par" with visual stimulations; no distinction has been made between the two kinds of stimulation on grounds other than the kind of sensory receptors involved in each case.

One way to classify stimulations is, then, according to which neural receptors are being fired. A second way of classifying stimulations is according to source: thus, we might say that all stimulations are either verbal or non-verbal, depending on whether the (auditory) stimulations impinging on the linguist have verbal or non-verbal sources. It would seem that, from the linguist's point of view, the most notable kind of auditory stimulations are those which are verbal. And, since auditory stimulations and, in particular, verbal stimulations, are "on a par" with visual and all non-verbal stimulations, there is no apparent reason why verbal stimulations ought not be included in the notion of stimulus meaning, along with non-verbal stimulations. Quine thinks, however, that there is a reason why verbal stimulations must be excluded from the notion of stimulus meaning.

As suggested above, Quine argues that the reason why verbal stimulations must be excluded from the notion of stimulus meaning is because they are useless to the linguist; in particular, they are useless to the linguist engaged in

radical translation. Quine describes his notion of stimulus meaning initially as "the objective reality that the linguist has to probe when he undertakes radical translation" (Quine, 1960, p. 39). By "radical translation," Quine means "translation of the language of a hitherto untouched people" (1960, p. 28). Since the people are "hitherto untouched," there are no interpreters available to the linguist. Therefore, says Quine, "the utterances first and most surely translated in such a case are ones keyed to present events that are conspicuous to the linguist and his informant" (1960, p. 29). These sentences are most easily translated, because the linguist can easily observe the native's response to such conspicuous events. Thus, to use Quine's example, suppose a rabbit runs by in front of the linguist and his informant; the native may utter "Gavagai," and the linguist may tentatively equate the native's "Gavagai" with his own English "Rabbit" (or "Lo, a rabbit"). The linguist may go on to test this hypothesis (viz., that "Gavagai" may be translated as "Rabbit") by querying the native "Gavagai?" in various stimulatory situations, and noting the native's responses. Quine assumes that the linguist will at this point be able to recognize the native's signs for assent and dissent;²⁰ and, the linguist

²⁰Hintikka has argued that this assumption is either unwarranted or it renders dubious the difficulties Quine sees in "later" translations--i.e., translations of non-observation sentences (Hintikka, 1969, pp. 70-73); and, Quine has replied that translation of native signs of assent and dissent is just as difficult and uncertain as the "later" translations (since they can only be carried out with the help of analytical hypotheses) (Quine, 1969, p. 312). This reply casts some doubt on Quine's claim about the relative ease and certainty

notes whether the native assents, dissents, or makes neither response to the linguist's queries of "Gavagai?" in the different stimulatory situations (1960, pp. 29-30).

As we have seen, that class of stimulatory situations which the linguist finds prompts the native's assent to the query "Gavagai?" constitutes the affirmative stimulus meaning of "Gavagai" for the native. For the sentences that are first and most easily translated by the linguist, Quine claims that the stimulations that will make up their stimulus meanings will be non-verbal stimulations. He says that radical translation depends, in the early stages, "solely on correlations [of sentences] with non-verbal stimulation" (1960, p. 32). And, since no translation from one language into another can preserve correlations of sentences with verbal stimulations--i.e., with other sentences of the language being translated--such correlations of sentences with verbal stimulations are useless in all stages of translation.

In addition to saying that stimulus meaning is "the objective reality that the linguist has to probe when he undertakes radical translation," Quine also says that the stimulus meaning of a sentence is what he calls the "empirical meaning" of the sentence. Thus, Quine says that "the notion of stimulus meaning . . . isolates a sort of net empirical import of each of various sentences without regard

with which "early" translations--i.e., translations of observation sentences--may be carried out, since even these "early" translations depend on prior translation of signs of assent and dissent; *v*, Quine, 1960, p. 68.

to the containing theory . . ." (Quine, 1960, pp. 34-35).²¹

It is well known, however, that Quine believes that individual theoretical sentences (if not all sentences) do not have a separable fund of empirical consequences--i.e., separable empirical meanings. Put in the terminology of stimulus meaning, Quine's position is that the stimulus meanings of theoretical sentences are very sparse. Their stimulus meanings are sparse, because the sentences do not have separable empirical meanings; and they do not have separable empirical meanings because the sentences are not individually correlated with separate sets of non-verbal stimulations. The crucial point is, however, that even though the stimulus meanings of theoretical sentences are sparse, the sentences are, individually, correlated with separate sets of verbal stimulations. Thus, Quine says that

the less variable the standing sentences are in point of assent and dissent, the sparser their stimulus meanings will be [But] however sparse its stimulus meaning, a sentence retains its connections with other sentences and plays its distinctive part in theories (Quine, 1960, p. 63).

This seems to be conclusive evidence showing that Quine does exclude sentence-to-verbal-stimulation correlations from his notion of stimulus meaning: theoretical sentences have sparse stimulus meanings, even though they have rich connections with other sentences. If Quine admitted verbal stimulations in the notion of stimulus meaning, these sentence-to-

²¹See also Quine, 1969, p. 32, and Quine, 1969~~e~~, pp. 80-81, on the identification of stimulus meaning with empirical meaning.

verbal-stimulation connections could be exploited in the same way in which Quine exploits sentence-to-non-verbal-stimulation connections.

It might be replied to this argument that even if Quine's doctrine concerning the sparseness of stimulus meanings of theoretical sentences supports the view that Quine excludes verbal stimulations from the notion of stimulus meaning, other doctrines held by Quine tend to support the contrary view. Thus, it might be argued that Quine's claim that "the stimulus meaning of a very unobservational occasion sentence for a speaker is a product of two factors, a fairly standard set of sentence-to-sentence connections and a random personal history" (Quine, 1960, p. 45), contradicts both his claim that the stimulus meanings of theoretical sentences are sparse, and my claim that Quine excludes verbal stimulations from the notion of stimulus meaning.

I have two replies to this objection. The first is that the objection shows that Quine's position is, at best, self-contradictory, because Quine claims in one place that verbal stimulations are to be excluded from stimulus meanings (in his discussion of standing sentences, quoted above) and in another place that they are to be included in stimulus meanings (in his discussion of unobservational occasion sentences, also quoted). All I am recommending is that Quine's position be made consistent, and that verbal stimulations be consistently included in the notion of stimulus meaning.

My second reply to the objection is that when Quine says that the stimulus meaning of a very unobservation occasion sentence for a speaker is "a product of," in part, a "set of sentence-to-sentence connections," he is probably referring to the stimulations which instill the speaker's disposition to assent to and dissent from the sentence, rather than the stimulations which activate the disposition. (If he is, in fact, referring to the stimulations which instill the relevant dispositions, then his position is not self-contradictory, in the sense just noted.) The stimulations which instill speakers' dispositions to assent to and dissent from sentences may be verbal; but, Quine denies that the stimulations that instill the relevant stimulations are those that constitute the stimulus meanings of sentences (Quine, 1970, p. 34). The stimulations that constitute the stimulus meanings of sentences are those that activate rather than instill the relevant dispositions; and, as I have shown, these stimulations must, for Quine, be non-verbal.

The only reason Quine adduces for excluding verbal stimulations from the notion of stimulus meaning is that they are useless to the linguist engaged in translation. That they are useless in this project, however, is no reason for excluding them from the notion of stimulus meaning.

Furthermore, it is not altogether clear that verbal stimulations are, as Quine says, useless in translation. For, Quine seems to be ignoring the fact that among those events which are "conspicuous" to the linguist and his native

informant are events of verbal behavior, carried out by either the linguist, the informant, or other natives in the auditory field of the linguist and the informant. Thus, for example, a native may walk past the linguist and his informant, and say "Olleh." The informant might then say to the linguist "Gniteerg," and, as in the "Gavagai" case, the linguist might tentatively equate "Gniteerg" with "That is an expression for greeting." And, in this case, since the stimulatory situation of the informant's "Gniteerg" included a verbal aspect (viz., the native's utterance of "Olleh"), the linguist might also tentatively equate "Olleh" with "Hello."²² By querying the informant "Gniteerg?" in various stimulatory situations (all of which must contain the native who uttered "Olleh"), and by noting the verbal element of these situations, the linguist can test his two hypotheses that, first, "Gniteerg" may be translated as "That is an expression for greeting" and, second, that "Olleh" may be translated as "Hello." In addition, by noting the verbal elements of the stimulatory situations in which the informant assents to the query "Gniteerg?" the linguist can begin to construct hypotheses concerning which native sentences are intrasubjectively stimulus synonymous with the sentence "Olleh" (because the informant will presumably assent to "Gniteerg?" in situations in which utterances other than "Olleh" are uttered).

²²Of course, the linguist may not come to make such hypotheses until after witnessing a number of similar instances; the same is true of Quine's linguist developing his hypothesis concerning "Gavagai."

It would appear, then, that verbal stimulations can be of great help to the linguist in his initial efforts to discover the stimulus meanings of native sentences. In particular, the linguist may determine the stimulus meanings (with inductive uncertainty, of course) of the sentences "Gniteerg" and "Olleh," for the informant (in the case of "Gniteerg") and for the other native (in the case of "Olleh"). In addition, the linguist can begin to equate "Olleh" with other sentences uttered by the native: "Olleh" may be equated with each of the other sentences uttered by the native in the situations in which the informant assents to "Gniteerg?"

Quine does not seem to give any reasons for thinking that the linguist cannot use the methods described to determine the stimulus meanings of the sentences for the natives; in fact, Quine proposes a method for investigating the intrasubjective stimulus synonymy of non-observation sentences which is similar to the method just described for investigating the intrasubjective stimulus synonymy of "Olleh." And yet, the initial translations the linguist can give as a result of the procedures just described depend on correlations of sentences with verbal stimulations; this requires modification of what Quine says, that the first and easiest translations made by the linguist depend on sentence-to-non-verbal-stimulation correlations.

As just noted, Quine does propose a method for exploring intrasubjective stimulus synonymy of non-observation sentences. Thus, for example, a visiting Martian linguist

might test a tentative equation of "Unmarried man" with "Bachelor," by querying the sentences "in parallel under random stimulations until he either hits a stimulation that prompts assent or dissent to one sentence and not to the other, or else is satisfied at last that he is not going to" (Quine, 1960, p. 47). Quine says, however, that the flaw in the above method is that "there is no evident reason why it should occur to him [the linguist] thus blindly to try comparing 'Unmarried man' with 'Bachelor'" (1960, p. 47).

It seems, however, that the "flaw" Quine sees in his method for probing intrasubjective stimulus synonymy of non-observation sentences disappears, as soon as the Martian linguist is allowed to use sentence-to-verbal-stimulation correlations in his investigation. For, as in the "Olleh" case, the Martian linguist might be led to equate "Unmarried man" with "Bachelor" if he observed his informant make a particular response each time a third party uttered either "Unmarried man" or "Bachelor." And, once it had been suggested in this way to the linguist to compare "Unmarried man" with "Bachelor," the linguist could go on to test the stimulus synonymy of the two sentences for the native by the method described by Quine. Even though the linguist might not have yet discovered the stimulus meaning of either "Unmarried man" or "Bachelor," he could, by this method, discover that they are stimulus synonymous for the third party uttering them. And, in this case it is correlations of the informant's response with verbal stimulations--namely,

the third party's utterance of "Unmarried man" with "Bachelor." It would seem, therefore, that if Quine dropped his stricture against relying on sentence-to-verbal-stimulation correlations, at least the problem about investigating intrasubjective synonyms would be overcome.

The basic point is that verbal stimulations are relevant to the stimulus meanings of utterances; and yet, Quine claims that they are not relevant. Quine's claim is that only correlations of sentences (such as "Gavagai") with non-verbal stimulations will be useful in the initial stages of translation, and that therefore verbal stimulations are to be excluded from the notion of stimulus meaning. I have argued both that verbal stimulations are useful in the initial stages of translation, and that verbal stimulations ought not be excluded from the notion of stimulus meaning. Specifically, I pointed out that Quine ignores the role which can be played by sentences (such as "Gniteerg") which are correlated with verbal stimulations (such as the utterance of "Olleh"). As we have seen, sentence-to-verbal-stimulation correlations allow the linguist to not only begin constructing his native-English dictionary, but they also allow him to begin making hypotheses about which sentences are stimulus synonymous in the native's language. Once the linguist has begun to compile a list of stimulus synonymous sentences in the native's language, his task of discovering the stimulus meanings of native sentences will be considerably eased: because, once he has discovered the stimulus meaning for one

native sentence, he will have, ipso facto, discovered the stimulus meanings for all native sentences which are stimulus synonymous with that sentence.

That sentence-to-verbal stimulation correlations should be relevant to the stimulus meaning of sentences should not be surprising, since surely what we mean in a certain situation is at least partly a function of what is said in that situation. The stimulus meaning of a sentence is a function of not only the sentence uttered itself and its non-verbal stimulatory situation, but also of the other sentences uttered in the stimulatory situation. The above has been a very schematic indication of how linguists may use sentence-to-verbal-stimulation correlations in probing the stimulus meaning of sentences. The method is introduced here because it will be contended later that it is a naturalistic method for investigating meanings which is open to Quine, but which has been ignored by Quine. I shall argue later that if due attention is paid to sentence-to-verbal-stimulation correlations, then, the notion of stimulus meaning may be enriched in ways not foreseen by Quine.

2.6 Stimulus Meaning and Naturalism

Despite the inadequacies of Quine's notion of stimulus meaning, and, in particular, despite the involvement of his theory of meaning with the notion of private languages, there is a sense, I believe, in which Quine's theory of meaning is a naturalistic theory of meaning. In

this section, I shall show that Quine's theory of meaning is a naturalistic theory, in the sense spelled out in 1.2-1.3.

First, it should be obvious that Quine has provided a theory of meaning, in the sense that he has provided functional accounts of phenomena germane to meaning. As we have seen, the stimulus meaning of a sentence is, for Quine, a function of the stimulatory situation of that sentence. In analyzing the notion of meaning, Quine has explicitly argued that on his view the stimulatory situations of utterances are to be construed as the independent variables of which the stimulus meanings of the sentences are functions.

On this interpretation, Quine has provided a theory of meaning, in spite of the fact that, as we have seen, his notion of stimulus meaning (and of language) has important differences from the preanalytic notion of meaning (and of language). This situation is admissible, since it was noted in 1.3 that a theory of meaning need not "save" all our preanalytic beliefs about meaning.

Despite the discrepancies between stimulus meaning and meaning, Quine claims that there are good reasons for retaining his notion of stimulus meaning, at least for the theoretical purposes of describing the activities of linguists, language learners, etc. Very simply, Quine claims that his notion of stimulus meaning contains just what empirical content was present in the old notion of

meaning; stimulus meaning, not meaning as preanalytically conceived, is "the objective reality the linguist has to probe" (1960, p. 39): and if, on these standards, stimulus meaning does not live up to meaning, then so much the worse for meaning.

Quine's adoption of the theoretical notion of stimulus meaning, instead of the mentalistic notion of meaning, which might "save" more of our preanalytic beliefs about meanings, suggests the sense in which his theory of meaning is naturalistic. Quine's theory is naturalistic, first, because it is committed to the principle that phenomena germane to meaning can be explained, if at all, by those modes of explanation employed in the natural sciences. As we have seen, Quine says that his main purpose in Word and Object is to "consider how much of language can be made sense of in terms of its stimulus conditions" (1960, p. 26); he says that if there are elements of the mentalistic theory of meaning, or aspects of our preanalytic notion of meaning, which cannot be made sense of in terms of stimulus conditions, then so much the worse for those beliefs. We may interpret Quine's statement of purpose in Word and Object, then, as making explicit his conception of what kind of explanation of phenomena germane to meaning is acceptable (viz., explanation in terms of stimulus conditions), and his assumption that any phenomena which are not explicable by this method are to be ruled as pseudo-phenomena not in need of

explanation. As Schuldenfrei has observed, Quine is committed to the principle "that there is only one kind of knowledge in the world, and it is the kind that natural scientists have" (Schuldenfrei, 1972, p. 5). In our terms, this is to say that Quine is committed to our naturalistic principle (1), according to which the models of explanation used in the natural sciences are sufficient as models of explanation for all phenomena, whether in the natural sciences or not.

Quine's theory of meaning is also committed to our second principle of naturalism, according to which what exists are all those objects whose changes are susceptible to explanation by the method of naturalism, and only those objects which are required in order to explain other natural phenomena. This is simply to say that Quine is committed to the principle that natural objects, and only natural objects, exist. As Schuldenfrei says, again, "Quine's world view is . . . an anti-Cartesian materialist view . . . because it asserts . . . that there is fundamentally only one kind of entity in the world, and that is the kind studied by natural scientists" (Schuldenfrei, 1972, p. 5)--i.e., what we have called natural objects.

That Quine is committed to this naturalistic monism should have become apparent in the review of his objections to the mentalistic theory of meaning. It will be remembered that Quine argues, first, that only

those meanings exist which can be explained by the naturalistic method (viz., in terms of their stimulus conditions), and, second, that entities like propositions and ideas ought not be posited because they are either redundant or themselves in need of explanation. Quine argues that one reason why meanings and propositions ought not be posited is because they need not be; and, according to naturalistic principle (2), any theory which posits such needless entities is not a naturalistic theory. It would seem that Quine's inference from the premise that propositions need not be posited to the conclusion that therefore they ought not be posited, can only be licensed by something like the naturalistic "principle of ontological parsimony" of 1.3. According to this principle, entities are not to be posited without good reason, the only good reason being that they are required in order to explain other phenomena. In 2.1 we saw Quine objecting to propositions on the grounds that they are either redundant--i.e., that they explain no genuine phenomena that cannot be explained in terms of the very things that are needed to explain propositions--or they are in need of explanation themselves, in which case they explain nothing. In other words, Quine argues that propositions are not needed in an adequate theory of meaning; and, he claims, any theory which posits such useless entities will be, ipso facto, an untenable theory (so long as there is an adequate

alternative theory available, which does not rely on propositions). This is to say that Quine's theory of meaning is, in the sense defined in Chapter 1, a naturalistic theory.

The third sense in which Quine's theory of meaning is naturalistic is that it subscribes to the naturalistic principle that changes or processes of natural objects are effects of natural causes. As Paul Ziff has noted, what Quine has attempted to provide is a "causal account" of linguistic behavior in general, and of meaning in particular (Ziff, 1970, p. 73). Quine explicitly adopts the position that meanings are natural objects--i.e., sets of stimulations--and, as such, are causally related to other natural objects; he rejects the view that meanings (because they are "intentional" entities of some sort) are outside of the causal order of natural objects.

It will be recalled that for Quine the stimulus meaning of an utterance is a function of the speaker's behavior and situation; and, Quine says that instead of speaking of stimulus meanings of sentences as consisting of those stimulations under which the informant will assent to or dissent from the sentences when queried, it would be more accurate to "speak in a more causal vein of stimulations that will prompt the native [informant] to assent to or dissent from the queried sentence" (Quine, 1960, p. 30). It should be clear,

then, that Quine views the connection between the informant's assent or dissent and the situations which constitute the stimulus meaning of the sentence being queried as a causal connection. And, Quine construes the situations of which the stimulus meanings of sentences are functions as natural objects or properties of such objects. He explicitly repudiates the (Cartesian) view that meanings are functions of non-natural entities, such as minds, ideas, intentions, etc. (y Schuldenfrei, 1972, p. 6), and argues that meanings, if they are to be made sense of at all, must be made sense of in terms of overt behavior (or dispositions to overt behavior). Therefore, Quine seems to subscribe to the general naturalistic principle that phenomena germane to meaning are causally related to other natural objects, processes, or properties of such objects.

Fourth, Quine's theory of meaning is naturalistic in the sense that it subscribes to the principle that natural processes, and, in particular, processes germane to meaning, are regular, and that in order to explain such processes the linguist must seek to establish general laws covering the processes. The point is simply that Quine believes that verbal behavior is regular, in the sense that it, like other natural phenomena, can be explained in terms of certain statements of regularities or generalizations. Thus, Quine

rejects both the general principle that verbal behavior is not regular in the same sense as are phenomena studied by natural scientists, and the specific application of this principle according to which verbal behavior is rule governed behavior (and therefore not regular in the same sense as are phenomena studied by natural scientists).

As Ziff has observed, Quine has attempted to show that the only adequate method of investigation of verbal behavior is for the linguist to attempt to generate "viable regularities of the form 'If α is the case then a speaker does β '" (Ziff, 1970, p. 73), or, more specifically, "If α is the case then the (stimulus) meaning of a sentence for a speaker is β ." And, this is simply to say that according to Quine linguists involved in investigating stimulus meanings are to proceed in exactly the same manner as are natural scientists engaged in investigating any natural phenomena: they are to "gather inductive evidence" in attempts to confirm or disconfirm general laws covering the phenomena being investigated (Quine, 1960, p. 30).

Fifth, and finally, we have also seen how Quine's theory of meaning is naturalistic in the sense that it is committed to the principle that explanations of phenomena in terms of non-natural causes are not adequate explanations. Thus, for example, Quine claims that if propositions are not explicable in terms of sentences (or other natural objects or processes or properties of objects, such as stimulus meanings), then propositions cannot be used to explain

phenomena in need of explanation. Since, if propositions cannot be explained in terms of natural objects such as sentences, they lack an adequate explanation themselves, they cannot be relied on in explanations of other phenomena. In general, Quine assumes that phenomena germane to meaning are to be explained, if at all, by reference to the stimulatory situations of utterances; and, such stimulatory situations are conceived by Quine as what we have called natural objects. I conclude, therefore, for these reasons, that Quine's theory of meaning is naturalistic in the relevant sense.

CHAPTER 3

MEANING: GRICE'S ANALYSIS

3.1 Meaning_{nn}

In this chapter, Grice's analysis of meaning is examined; the purpose of the chapter is to develop and elucidate the points concerning Grice's analysis of meaning which were introduced in Chapter 1. Specifically, an attempt will be made to show, first, what kind of meaning Grice is concerned to analyze, second, why Grice thinks any "causal" analysis of meaning must fail, and, third, why Grice thinks that any adequate analysis of meaning must involve reference to the notion of intention. To these ends, I discuss, in the first section of the chapter, Grice's distinction between the natural and non-natural senses of "meaning"; in the second section, I discuss Grice's arguments against "causal" accounts of meaning; and, finally, several generalizations concerning Grice's analyses of meaning in its non-natural sense are discussed.

As was noted in Chapter 1, Grice begins his original account of meaning by distinguishing between what he calls the "natural" and the "non-natural" senses of "meaning." Grice considers the following two pairs of sentences:

- (1) Those spots mean measles.

- (2) The recent budget means that we shall have a hard year.

and

- (3) Those three rings on the bell (of the bus) mean that the bus is full.
- (4) That remark, "Smith couldn't get on without his trouble and strife," meant that Smith found his wife indispensable.

Grice considers (1) and (2) to be examples of sentences in which "means" is used in its natural sense, and (3) and (4) to be examples of sentences in which "means" is used in its non-natural sense (Grice, 1967, pp. 39-40).

There is, according to Grice, a cluster of characteristics which attach to sentences in which "means" is used in its natural sense, but not to sentences in which "means" is used in its non-natural sense. Thus, for example, Grice says that in cases like (1) and (2), "x meant that p and x means that p entail p," while in cases like (3) and (4), "x means that p and x meant that p do not entail p"; also, for neither (1) nor (2) "can a restatement be found in which the verb 'mean' is followed by a sentence or phrase in inverted commas," while for cases like (3) and (4) restatements can be given "in which the verb 'mean' is followed by a phrase in inverted commas"; and there are other differences as well (Grice, 1967, pp. 39-40).¹ The general point is that

¹One further interesting difference deserves to be mentioned here. Stampe has argued that the ambiguity of "meaning" (i.e., that "meaning" has two senses, a natural and

in distinguishing between the natural sense of "means" and the non-natural sense (hereafter referred to as means_{nn}, following Grice), Grice has attempted to distinguish between the sense of "means" which is connected with communication from the sense of "means" which is not connected with communication (Grice, 1969, p. 147). It is meaning_{nn}--the sense of "means" which is connected with communication--which Grice is concerned to analyze. Of course, questions have been raised as to whether the distinction between the natural and non-natural senses of "meaning" is a genuine distinction. I do not question the distinction here, because I shall argue that, even granting the distinction between the natural and non-natural senses of "meaning," Grice's theory of meaning is untenable.

It should be obvious, from the examples given, that things besides verbal utterances are used in communication (e.g., bells may be used on buses to communicate something to an audience). Grice is not concerned with meaning as

a non-natural sense) can be explained by appealing to the transformational histories of sentences containing the word "meaning." His thesis is that "under their 'natural' interpretation, [sentences containing 'meaning'] are produced by a deletion transformation operating on strings the subject of which is factive nominalization. On the other hand, under its 'nonnatural' interpretation, the sentence is derived from another transformation or chain of transformations which operates on strings having an agent-denoting expression as subject and an adverbial prepositional phrase, deleting both Agent and Preposition. . . ." As Stampe acknowledges, what he has shown is that Grice's distinction between the natural and non-natural senses of "meaning" is supported by a grammatical investigation of English sentences containing "meaning"; he has not even considered the question of whether English grammar is in this case "philosophically perspicuous or philosophically misleading." V Stampe, 1968, pp. 157 and 166.

applied solely to sentences or other linguistic units (as is Quine, for example). According to Grice, other things besides sentences and words may mean_{nn} something: people, when they utter sentences or words, may mean_{nn} something; non-verbal gestures or signs may mean_{nn} something; and people, when they engage in non-verbal gesturing, may mean_{nn} something.

Grice takes as his basic analysandum the expression "By uttering x U meant something" (1967, p. 45; and 1969, p. 151), where x stands for any utterance, U is a human agent, and "meant" is taken in its non-natural sense. It should be obvious, in the light of the last paragraph, that Grice is using the term "utterance" in an extended sense; as he says, the terms "utterance" and "uttering" are intended "to apply to any act or performance which is or might be a candidate for nonnatural meaning" (1969, p. 151).² Both "acts" (such as speeches and gestures) and "objects" (such as words spoken and gestures gestured) are utterances, in this sense (Grice, 1967, p. 41); verbal as well as non-verbal acts, as well as verbal and non-verbal objects, may be utterances. Grice is interested in the general phenomenon of meaning_{nn}. His analysis is intended to be fully general, in covering all cases of meaning_{nn}; the case of a sentence

²While it is true that all utterances are candidates for meaning_{nn}, it is not true that all candidates for meaning_{nn} are utterances, even in Grice's extended sense of utterance. People are also candidates for meaning_{nn}; as we shall see, what utterances mean_{nn} is analyzed by Grice in terms of what people mean_{nn} by their utterances; what people mean_{nn} is the basic notion.

meaning_{nn} something is a sub-case of meaning_{nn}, and Grice's analysis is intended to apply to it as well as all other sub-cases (Grice, 1969, p. 161).

According to Grice, there are four important species of meaning_{nn}, each of which is in need of analysis. That is, Grice thinks that there are four types of "meaning-specification" which can be given, in response to questions concerning the meaning_{nn} of utterances.

First, consider the sentence

(5) If I shall then be helping the grass to grow, I shall have no time for reading.

According to Grice, we can say that this sentence means_{nn} both "If I shall then be assisting the kind of thing of which lawns are composed to mature, I shall have no time for reading," and "If I shall then be assisting the marijuana to mature, I shall have no time for reading." Grice says that such meaning-specifications may be called "specifications of the timeless meaning(s) of a 'complete' utterance-type (which may be a sentence, or may be a 'sentence-like' nonlinguistic utterance-type, such as a hand-signal)" (1969, p. 147). In addition (as a sub-species of this first kind of meaning-specification), we may give meaning-specifications of timeless meaning(s) of incomplete utterance-types, such as non-sentential words or phrases, or nonlinguistic utterance-types which are analogous to words or phrases (Grice, 1969, p. 147).

Second, we can give specifications of what Grice

calls the applied timeless meaning of either complete or incomplete utterance-types on particular occasions of utterance. Thus, we might say that when U (a speaker) uttered (5) at (time) t_1 , what (5) meant_{nn} was "If I shall then be assisting the kind of thing of which lawns are composed to mature, I shall have no time for reading," and not the second reading of (5) given in the last paragraph. Specification of the applied timeless meaning of complete or incomplete utterance-types attempts to give the correct reading or interpretation of ambiguous utterance-types (i.e., utterance-types which have more than one timeless meaning) on particular occasions of utterance (Grice, 1969, p. 148).

Third, Grice claims it is possible that when U uttered (5), he meant_{nn} by (5) [i.e., by the words of (5)] either

(i) If I am then dead, I shall not know what is going on in the world,

or

(ii) One advantage of being dead will be that I shall be protected from the horrors of the world.

Grice makes the assumption that "helping the grass to grow" is not (as is "pushing up the daisies") a recognized idiom for "being dead." On this assumption, says Grice, it would be incorrect to say either that the timeless meaning of "I shall be helping the grass to grow" is "I shall then be dead," or that the applied timeless meaning of the phrase

is "I shall then be dead." This is so, apparently, because specifications of timeless meaning and of applied timeless meanings of utterance-types must be, if true at all, true by virtue of recognized public convention and not simply by idiosyncratic personal habit. Grice says that because "helping the grass to grow" is not a recognized idiom, it is incorrect to say that "I shall be helping the grass to grow" has either the timeless meaning or the applied timeless meaning of "I shall be dead." Rather, in Grice's terms, we may say that for U the occasion-meaning of the utterance-type "I shall be helping the grass to grow" is "I shall be dead." That is, in saying that when U uttered (5) what he meant_{nn} was either (i) or (ii), what we have done is specified the occasion-meaning of an utterance-type for a speaker (Grice, 1969, pp. 148-149).

Fourth, we may say of U that when he uttered (5) he meant_{nn} by (5) either (i) or (ii), as well as that when he uttered (5) what he meant_{nn} by (5) was either that if he would then be dead he would not know what was going on in the world, or that one advantage of being dead would be that he would be protected from the horrors of the world. Such meaning-specification involves the use of indirect quotation rather than the use of direct quotation (as in the specification of timeless meaning, applied timeless meaning, and occasion-meaning of utterance types). Grice calls this kind of meaning-specification, which involves the use of indirect

quotation, specification of an utterer's occasion-meaning (1969, p. 149).

It may be helpful, at this point, to present Grice's summary of the four kinds of meaning-specification just introduced:

- (1) "x (utterance-type) means ' . . . '" [Specification of timeless meaning for an utterance-type which is either (1a) complete or (1b) incomplete]
- (2) "x (utterance-type) meant here ' . . . '" [Specification of applied timeless meaning for an utterance-type which is either (2a) complete or (2b) incomplete]
- (3) "U meant by x (utterance-type) ' . . . '" [Specification of utterance-type occasion-meaning]
- (4) "U meant by uttering x that" [Specification of utterer's occasion-meaning] (Grice, 1969, p. 149).

It needs to be emphasized that in distinguishing between the timeless meaning of utterance-types [(1) and (2) above] and occasion-meaning [(3) and (4) above], Grice is distinguishing between species of meaning which attach to utterance-types and species of meaning which attach to people: as we saw, both utterance-types and people can mean_{nn} things.

Grice argues in "Utterer's Meaning and Intentions" and "Utterer's Meaning, Sentence-Meaning, and Word-Meaning" that the notions of timeless meaning, applied timeless meaning, and utterance-type occasion-meaning may be analyzed in terms of the basic notion of utterer's occasion-meaning (1969, p. 150). Therefore, for the sake of brevity, in this thesis I will concentrate on Grice's analysis of utterer's occasion-meaning; if this analysis can be shown to be inadequate, then this inadequacy will infect Grice's analyses of each of the other kinds of meaning-specification just reviewed. Although I will not

directly examine Grice's analyses of timeless meaning, applied timeless meaning, and utterance-type occasion meaning, the notions themselves will figure in the discussion of Grice's analyses of meaning_{nn}.

3.2 Grice's Rejection of the Causal Theory of Meaning

Before reviewing Grice's analysis of the notion of utterer's occasion-meaning, it will be helpful to discuss his criticism of what he calls "causal" theories of meaning. It has been one of Grice's main contentions, in each of his papers on meaning, that any causal analysis of meaning must be inadequate, and that any adequate theory of meaning will analyze meaning in terms of utterer's intentions.³ Just as, in Chapter 2, Quine's criticisms of mentalistic theories of meaning provided an introduction to his naturalistic theory of meaning, so may Grice's criticisms of the causal theory of meaning provide an instructive introduction to his own analysis.

Grice interprets the causal theory of meaning as an attempt to answer the question "What is meaning_{nn}?" (1967, p. 41). As we have seen, meaning_{nn} is taken by Grice to be a property of utterances, in his extended sense;⁴ therefore,

³For example, in "Utterer's Meaning and Intentions" Grice says that his main purpose is to defend the thesis that utterer's occasion-meaning, and hence each of the other kinds of meaning, is explicable in terms of the notion of utterer's intention (1969, p. 150). Since Grice's analysis of meaning in "Meaning" is an analysis of utterer's occasion-meaning, his arguments there may also be taken as a defense of the basic thesis that meaning must be analyzed in terms of utterer's intention.

⁴People may also be vehicles of meaning_{nn}; this point is ignored in this section, mainly because Grice ignores it in his discussion of the causal theory of meaning.

Grice is considering the causal theory of meaning as a theory about the meaning of all kinds of acts or objects of communication. The causal type of answer to the question "What is meaning_{nn}?" is, according to Grice, "that for x to mean_{nn} something, x must have (roughly) a tendency to produce in an audience some attitude (cognitive or otherwise) and a tendency, in the case of the speaker, to be produced by that attitude, these tendencies being dependent on 'an elaborate process of conditioning attending to the use of the sign in communication'" (1967, p. 41, quoting C. L. Stevenson, 1944, p. 57).

Grice has three criticisms of this causal theory of meaning. The first criticism is developed in two stages. First, Grice says that the theory is too weak, that it allows things to mean_{nn} something which do not mean_{nn} anything. Second, Grice considers a reply to his claim that the theory is too weak, and argues that the reply reveals that the causal theory is "unhelpful," in Grice's terms (1967, p. 42). In what follows, I consider Grice's argument, and contend that the causal theory is neither too weak nor "unhelpful."

Grice supports his claim that the causal theory of meaning is too weak with the observation that "many people have a tendency to put on a tail coat when they think they are about to go to a dance, and it is no doubt also the case that many people, on seeing someone put on a tail coat, would conclude that the person in question was about to go to a dance" (1967, p. 41). Therefore, according to Grice's interpretation of the causal theory, putting on a tail coat must

mean_{nn} that one is going to a dance.

Grice says, however, that the existence of the above tendency to put on a tail coat ought not satisfy us that putting on a tail coat means_{nn} that one is going to a dance. Grice is not claiming that putting on a tail coat cannot mean_{nn} that one is going to a dance; for, consider the following example. Suppose a man puts on a tail coat in front of his wife, with the intention of inducing in her the belief that he is about to go to a dance, and with the intention that his putting on the tail coat will be recognized by the wife as being intended to induce in her the belief that he is about to go to a dance. Under these circumstances, Grice would say that the man's putting on a tail coat was an utterance which meant_{nn} that he was about to go to a dance. Therefore, according to Grice, the act of putting on a tail coat can be an utterance which means_{nn} something. Therefore, Grice is not claiming that the causal theory is inadequate simply because it allows the act of putting on a tail coat to be an utterance which means_{nn} something. Rather, his claim is that when the act of putting on a tail coat is construed solely as the manifestation of a behavioral tendency to put on a tail coat in certain circumstances (as against--as we shall see--construing the act as a manifestation of certain intentions on the part of the agent), then the act of putting on a tail coat cannot mean_{nn} anything.

Grice's first criticism of the causal theory is, then, that it cannot be right, because acts which are

as we saw, Grice takes the task at hand to be the characterization of what it is for something to be used in communication--i.e., for something to have meaning_{nn}. Therefore, if the causal theory is construed as a theory of meaning in Grice's terms, then it is, as Grice claims, unhelpful.

My reply to this second stage of Grice's first criticism of the causal theory is that the causal theory being considered is not in fact an attempt to answer the question "What is meaning_{nn}?" Since it is only causal theories which are attempts to answer this question which are subject to Grice's criticism of "unhelpfulness," my claim is that Grice's criticism is undamaging to the causal theory.

If the causal theory were an attempt to answer the question "What is meaning_{nn}?" then according to the theory any action or object would be a candidate for meaning_{nn}; this is a consequence of Grice's definitions of "utterance" and "meaning_{nn}." But, it is not the case that the causal theory being considered is an attempt to answer the question "What is meaning_{nn}?" because it is not the case that according to the theory any action or object is a candidate for meaning_{nn}.

The causal theory considered by Grice is that developed by C. L. Stevenson. Stevenson's theory is not an attempt to answer the question "What is meaning_{nn}?" because the theory does not assume that any action or object is a candidate for meaning in the relevant sense. Rather,

as we saw, Grice takes the task at hand to be the characterization of what it is for something to be used in communication--i.e., for something to have meaning_{nn}. Therefore, if the causal theory is construed as a theory of meaning in Grice's terms, then it is, as Grice claims, unhelpful.

My reply to this second stage of Grice's first criticism of the causal theory is that the causal theory being considered is not in fact an attempt to answer the question "What is meaning_{nn}?" Since it is only causal theories which are attempts to answer this question which are subject to Grice's criticism of "unhelpfulness," my claim is that Grice's criticism is undamaging to the causal theory.

If the causal theory were an attempt to answer the question "What is meaning_{nn}?" then according to the theory any action or object would be a candidate for meaning_{nn}; this is a consequence of Grice's definitions of "utterance" and "meaning_{nn}." But, it is not the case that the causal theory being considered is an attempt to answer the question "What is meaning_{nn}?" because it is not the case that according to the theory any action or object is a candidate for meaning_{nn}.

The causal theory considered by Grice is that developed by C. L. Stevenson. Stevenson's theory is not an attempt to answer the question "What is meaning_{nn}?" because the theory does not assume that any action or object is a candidate for meaning in the relevant sense. Rather,

Stevenson's theory is a theory about the meaning of "signs," by which he means, primarily, sentences and other units of linguistic behavior. Stevenson says that his theory of meaning is part of "a general theory of signs" (1944, p. 36); he is interested, however, in signs "which make up a recognized grammatical part of speech," or "conventionalized signs," as opposed to "natural" manifestations of emotions such as laughs and groans (1944, pp. 38-39). For Stevenson (as for other "causal" theorists, such as Bloomfield and Quine), only sentences, words, and other units of specifically linguistic behavior are candidates for meaning; for Grice, any act or object is a candidate for meaning_{nn}.

In the light of these considerations, let us reconsider Stevenson's theory of meaning. According to Stevenson (rather than Grice), for x to mean something in the relevant sense [the "narrow" or linguistic sense, in Stevenson's terms (1944, p. 39)], x must be a "sign" which is a recognized grammatical part of speech, which must have a tendency to produce in an audience some attitude and a tendency, in the case of the speaker, to be produced by that attitude, these tendencies being dependent on an elaborate process of conditioning attending the use of the sign in communication. Grice claims that this theory is unhelpful, because it asserts that an act or object being used in communication is sufficient for its meaning_{nn} something; it should be clear, however, that it is simply not the case that according to Stevenson's theory it is sufficient for an act or object to

mean_{nn} something that it be used in communication. It is a necessary condition, according to Stevenson, that a sign be used in communication, for the sign to mean something; but the sign will only mean something if it satisfies certain other conditions as well (notably, that it be a sign which makes up a recognized grammatical part of speech). These conditions specified by Stevenson may not be jointly necessary and sufficient for a sign to have meaning_{nn} for Grice, but they certainly do not constitute an "unhelpful" analysis, in Grice's sense.

Grice may reply, at this point, that the burden has simply been shifted from the notion of "being used in communication" to being a "sign," or a "unit of linguistic behavior." This reply is true, but undamaging. While a reliance in the analysis on the notion of "being used in communication" in an analysis of meaning_{nn} does render the analysis circular, a reliance on the notion of "being a unit of linguistic behavior" in an analysis of meaning does not render the analysis circular. For, units of linguistic behavior can be identified as such on grounds independent of their use in communication. (For example, signs, in Stevenson's sense, may be identified on syntactic, phonological, or etymological grounds.) And, once identified, units of linguistic behavior have meaning if they are manifestations of certain behavioral tendencies--namely, the ones specified in Stevenson's analysis. There is nothing "unhelpful" in this analysis.

The upshot of the above discussion is that Grice's

criticism of the causal theory of meaning has been successfully rebutted. The causal theory is not "unhelpful," because being used in communication is not, according to the causal theory, a sufficient condition for an act or object having meaning. And, the causal theory has not been shown to be too weak, by the "tail coat" and similar examples, because putting on a tail coat is not a sign which is a recognized grammatical part of speech. Putting on a tail coat is not a candidate for meaning, for Stevenson's theory, because it fails to satisfy a condition necessary for a sign to have meaning--viz., that the sign be a sign, in Stevenson's sense. Since putting on a tail coat does not mean anything, according to Stevenson's theory, the causal theory considered by Grice has not been shown to be too weak by Grice's argument.

Grice says that the second main difficulty he sees in the causal theory is "really the same difficulty" as the first--by which he is referring, I take it, to his claim that the theory is too weak (Grice, 1967, p. 42); if the first difficulty has been removed, then so must have been the second. Whether this is so or not, it will be worthwhile to examine Grice's argument against the causal theory. Grice's second complaint is that it is difficult, according to the causal theory, to distinguish between what is meant_{nn} and what is merely suggested by an utterance. For example, Grice claims that it is difficult, according to the causal theory, to avoid saying "that 'Jones is tall' is part of

what is meant by 'Jones is an athlete,' since to tell someone that Jones is an athlete would tend to make him believe that Jones is tall" (Grice, 1967, p. 42). Grice goes on to argue that this difficulty in distinguishing between what is meant_{nn} and what is suggested by an utterance constitutes a difficulty for the causal theory.

That the difficulty in distinguishing what is meant_{nn} from what is suggested constitutes a reason for revising or abandoning the causal theory is argued by Grice in the following way. Suppose that agent U at time t₁, on hearing the utterance

(6) Jones is an athlete,

is in fact disposed to believe

(7) "Jones is tall" is true.

According to the causal theory, it follows that at t₁ U must believe

(8) "Jones is tall" is part of the meaning of "Jones is an athlete."

We may suppose further that U is queried at time t₂, and it is found that he believes

(9) "There are non-tall athletes" is true.

Given this example, Grice argues that the inconsistency between (8) and (9) constitutes prima facie evidence for abandoning the causal theory of meaning.

The reply to Grice's objection is simply that there is no obvious need to distinguish between what is meant and what is suggested by verbal utterances. The question is,

Why not say that for U at t_1 part of the meaning of "Jones is an athlete" is "Jones is tall"? It will not help to merely cite (9), and say that the conflict between (8) and (9) renders the causal theory inconsistent and untenable. For, the apparent inconsistency between (8) and (9) can be explained away in either or both of two ways, neither of which involves abandoning the causal theory.

First, it may be said that people simply are non-uniformitarian in the way they interpret words and sentences: they attach a semantic interpretation to a linguistic form at one time which is inconsistent with the semantic interpretation attached to another form at a different time. Thus, the inconsistency between U's partial interpretation of (6)-- i.e., (8)--and U's interpretation of the word "athlete"-- which is given, in part, in (9)--simply reflects the facts of verbal behavior. In other words, that (8) and (9) are inconsistent [and that (8) is implied by the causal theory] does not render untenable the causal theory; rather, it merely shows that the causal theory reflects the (often inconsistent) facts of verbal behavior.

Second, it may be replied that to conclude from the inconsistency of (8) and (9) that the causal theory is untenable begs certain crucial questions about meaning which cannot be presumed to have been resolved in a discussion of the adequacy of the causal theory. Specifically, Grice must want to be able to distinguish between what is meant_{nn} and what is suggested by an utterance because he thinks that

it is in some sense worse for a theory to imply an inconsistency like that between (8) and (9) than it is for a theory to imply an inconsistency like that between (9) and

(10) "Jones is tall" is suggested by "Jones is an athlete."

More perspicuously, Grice must think that there is a difference in principle between the kinds of inconsistency claimed in

(11) "'There are non-tall athletes' is true" is inconsistent with "'Part of the meaning_{nn} of 'Jones is an athlete' is 'Jones is tall'" is true,"

and that claimed in

(12) "'There are non-tall athletes' is true" is inconsistent with "'Jones is tall' is suggested (but not meant_{nn}) by 'Jones is an athlete' is true."

What is Grice's justification for his claim that the inconsistency in (11) is in some sense more serious than that in (12)? One thing that could be used as such a justification would be the further claim that there is a difference between the kinds of inconsistency in

(13) "'Jones is tall" is part of the meaning of "Jones is an athlete"' is an analytic truth" is inconsistent with "'There are non-tall athletes' is true"

and

(14) "'Jones is tall" is part of the meaning of "Jones is an athlete"' is a contingent truth" is inconsistent with "'There are non-tall athletes' is true."

According to this interpretation, the difference Grice sees between the inconsistencies between (a) what is meant and what is said and (b) what is suggested and what is said must be precisely that between inconsistencies between (c) analytic truths and contingent truths [e.g., (13)] and (d) contingent truths and contingent truths [e.g., (14)]. I can see no other plausible justification for Grice's claim that it is all right for a theory to imply that what is suggested by (6) is inconsistent with (9), while it is not all right for a theory to imply that what is meant_{nn} by (6) is inconsistent with (9).

However, it has been argued (by, perhaps not incidentally, Quine and other holders of causal theories of meaning) that there is no distinction in principle between the kind of inconsistency exhibited in (13) and that in (14): or, more generally, that there is no distinction to be made between analytic truths--i.e., truths "by virtue of meaning"--and contingent truths--i.e., truths "by virtue of facts." And, if there is no difference in principle between analytic truths and contingent truths, then there is no distinction between the kinds of inconsistency in (13) and (14), as well as none between that in (11) and (12). And, if there is no difference between the kind of inconsistency in (11) and

(12), then there is no apparent need to distinguish between what is meant from what is suggested by an utterance. And, if there is no need to make this distinction, then the causal theory ought not be criticized for not being able to make it.

Of course, that there is no difference in principle between analytic truths and contingent truths is a highly contentious point; I do not need to establish here, however, either that there is a difference or that there is not. For, everyone will agree that whether there is a difference between "truth by virtue of meaning" and "truth by virtue of facts" depends on one's theory of meaning. Therefore, to rely on any answer to the question of whether there is a difference between the two kinds of truth in a criticism of a theory of meaning is to assume that one's own theory of meaning is correct. And, Grice's second criticism of the causal theory of meaning begs the question in precisely this way: the claim that the difficulty for the causal theory to distinguish between what is meant_{nn} and what is suggested by utterances constitutes a reason to abandon the causal theory presupposes that there is a difference between what is suggested and what is meant_{nn} by an utterance; the claim that there is a difference between what is suggested and what is meant_{nn} presupposes that there is a difference between analytic truths and contingent truths; and, the claim that there is a difference between analytic truths

and contingent truths presupposes that one's own theory of meaning is correct.

Grice's third and final criticism of the causal theory is that even if it is acceptable, what it provides is "an analysis only of statements about the standard meaning, or the meaning in general, of a sign": no analysis is given of what it is for a particular speaker to mean_{nn} something by an utterance on a particular occasion (Grice, 1967, p. 42). Adopting terminology introduced earlier, Grice is claiming that the causal theory provides an analysis of the timeless meaning of utterances, but ignores occasion-meaning completely. Grice thinks that the fact that the causal theory is a theory solely about the timeless meaning of utterances constitutes a serious enough criticism of the causal theory; Grice says, however, that one could go further, and "maintain that the causal theory ignores the fact that the meaning (in general) of a sign needs to be explained in terms of what users of the sign do (or should) mean by it on particular occasions" (1967, p. 42).

Grice claims that it is not obvious how the causal theory could be adapted to provide an account of the occasion-meaning of utterances. It would seem, however, in the light of such causal theories of meaning as Bloomfield's and Quine's, that it is obvious how a causal account could be given of the occasion-meaning of utterances. Let us say, for example, that the occasion-meaning of utterance x for speaker U at time t is constituted by (a) the stimulatory

situation which prompts \underline{U} to utter \underline{x} at \underline{t} , and (b) the "practical events" (following Bloomfield) which are elicited in the audience by the utterance. Accordingly, the timeless meaning of \underline{x} could be identified as the occasion-meaning which \underline{x} generally has in some group of speakers. What we have here is a rudimentary causal theory of meaning, in which occasion-meaning is the basic notion analyzed, and in which the timeless meaning of utterances is analyzed in terms of occasion-meaning. Whether this rudimentary theory shows any promise of being adequate is not important; the point is that there is no reason to agree with Grice that there is some difficulty of principle in adapting the causal theory of meaning to provide an account of timeless meaning in terms of occasion-meaning. I conclude, therefore, that Grice has provided no reason to believe that a causal analysis of meaning cannot, in principle, be adequate.

3.3 Meaning_{nn} and Speakers' Intentions

Grice goes on, however, to give his own analysis of meaning_{nn} in terms of speakers' intentions. Grice's original analysis of meaning_{nn} has been widely discussed in the literature, and Grice has revised the analysis several times in attempting to meet various criticisms. In this section, I shall review Grice's original analysis, and make several generalizations concerning certain important features of his latest analysis.

In "Meaning," Grice adopts as his basic analysandum

"U meant_{nn} something by uttering x,"⁶ where U stands for a person and x stands for an utterance. As we saw, Grice interprets his analysis as an attempted analysis of utterer's occasion-meaning. Grice claims that

- (I) "U meant something by uttering x" is true iff, for some audience A, U uttered x intending
- (1) A to produce a particular response r
 - (2) A to think (recognize) that U intends (1)
 - (3) A to fulfill (1) on the basis of his fulfillment of (2) (Grice, 1969, p. 151).⁷

According to (I), to ask what U meant by uttering x is to ask for a specification of the intended effect of uttering x (Grice, 1967, p. 46). Grice goes on to place a restriction on clause (3), such that "to suppose A to produce r 'on the basis of' his thinking that U intends him to produce r is at least part of his reason for producing r, and not merely the cause of his producing r" (1969, p. 151). This restriction is important (partly because it reveals Grice's position on the issue of the relation between reasons and causes), but it is not crucial to our discussion; therefore, it will be ignored in what follows.

In "Utterer's Meaning and Intentions," Grice considers a long series of revised versions of (I). There is no

⁶For the sake of consistency, I have adopted Grice's 1969 re-wording of his 1957 analysis, given in Grice, 1969, p. 151.

⁷In this and all subsequent Gricean analyses of utterer's occasion-meaning, "meant" is, of course, intended to be "meant_{nn}."

need to consider each revision here; several general points, however, deserve to be mentioned. First, in all subsequent analyses, Grice never abandons his general assumption that utterer's occasion-meaning is the basic notion requiring analysis, and that both species of timeless meaning as well as utterance-type occasion-meaning must be analyzed in terms of utterer's occasion-meaning. A corollary of this assumption is, as we have seen, that what utterances mean must be analyzed in terms of what people mean by their utterances; more specifically, sentence- and word-meaning must be analyzed in terms of utterer's meaning. Grice never abandons these claims.

Second, Grice never abandons his general claim that utterer's occasion-meaning (and hence all subsidiary notions) must be analyzed in terms of speakers' intentions--with certain qualifications on what the intentions can be. Let us say, for the sake of brevity, that Grice never abandons his claim that "U meant_{nn} something by uttering x" is true iff U uttered x with the "right Gricean intention." In order to be the "right Gricean intention," the intention must satisfy certain conditions; according to (I), the intention must satisfy the three conditions specified in (1)-(3).

In Grice's latest version of his analysis of meaning_{nn}, the conditions which must be satisfied in order for an intention to be the "right Gricean intention" are more numerous as well as more complicated than they are in

(I);⁸ nevertheless, the basic claim is unchanged, that "U meant by uttering x that $*\psi p$ "⁹ (which is supposed to be a more perspicuous version of "U meant_{nn} something by uttering x") is true iff U uttered x with the right Gricean intention.

The third general observation concerning Grice's revisions of his analysis is that he never abandons his basic concern for utterances in his extended sense. That is, Grice is primarily concerned to show what it is for people to mean_{nn} things by their utterances; people meaning_{nn} things by sentences is a special case of people meaning_{nn} things by utterances. And, since the timeless meaning of utterances must be analyzed in terms of what people mean_{nn} when they utter the utterances, the timeless meaning of sentences must also be analyzed in terms of what people mean_{nn} when they utter the sentences.

3.4 Objections to Grice's Analysis of Meaning_{nn}

With these general features of Grice's analyses of meaning_{nn} in mind, it will be possible to consider some objections to Grice's analyses. In this section, I will consider only objections which serve to bring out certain

⁸y Grice, 1969, pp. 175-176.

⁹Where " $*\psi$ " "is a mood marker," an "auxilliary" correlated with propositional attitudes, and "p" is a variable ranging over propositional attitudes. Grice's definition of " $*\psi p$ " is not clear: if my interpretation is correct (and I can imagine no plausible alternative) then there is a misprint in Grice's definition of " $*\psi p$," in which Grice has " ψ " rather than "p" ranging over propositional attitudes, y Grice, 1969, p. 171.

important contrasts between Grice's analysis of meaning and the other analyses being considered in this thesis.

The first objection to Grice's analysis is directed toward his extended usage of "utterance." As we saw, utterances for Grice need not be strictly verbal or linguistic: any act or object used in communication may be an utterance, and, as such, may have meaning_{nn}. My objection to Grice's notion of utterance is simply that in considering the meaning of utterances in general rather than the meaning of specifically linguistic units, Grice runs the risk of glossing over certain features of meaning which may attach to meaningful linguistic units, and not to meaningful non-linguistic utterances, in Grice's sense. Of course, one goal of philosophical and scientific theories in general is to provide general accounts of disparate phenomena: i.e., to show that apparently unrelated phenomena are in fact related, and that one theory is sufficient for their explanation. If Grice were to succeed in demonstrating that a single theory is sufficient in accounting for the meaning of both linguistic and nonlinguistic utterances, this would be an important result. I shall argue, however, that there is evidence which counts against such an assimilation of the meaning of linguistic and nonlinguistic utterances.

Grice's assumption that the kind of meaning possessed by both nonlinguistic utterances and linguistic units is the same is questionable for several reasons. The first way in which Grice's assumption leads to trouble is that it forces

him to adopt certain ad hoc distinctions in his analyses of meaning_{nn}, which would not need to be made if he were to give separate analyses of the meaning of nonlinguistic utterances and of linguistic units. Thus, for example, in replying to an objection made by Searle (which we will consider in greater detail in the next chapter), Grice says the following:

Of course, I would not want to deny that when the vehicle of meaning is a sentence (or the utterance of a sentence) the speaker's intentions are to be recognized, in the normal case, by virtue of a knowledge of the conventional use of the sentences. . . . But as I indicated earlier, I would like, if I can, to treat meaning something by the utterance of a sentence as being only a special case of meaning something by an utterance (in my extended sense of utterance), and to treat a conventional correlation between a sentence and a specific response as providing only one of the ways in which an utterance may be correlated with a response (Grice, 1969, p. 161).

In accordance with this stipulation, Grice is forced to introduce a variable ranging over "modes of correlation" in all subsequent analyses of meaning_{nn}.¹⁰ What particular mode of correlation is relevant to particular meaning-specifications depends on what type of utterance is being considered. Thus, if the meaning of linguistic utterances is being explored, the relevant mode of correlation is, Grice says, "conventional"; if the meaning of nonlinguistic utterances is being explored, the relevant mode of correlation is, presumably, "nonconventional."

Grice never gives an example of an analysis of the meaning of a nonlinguistic utterance in which the mode of

¹⁰v Grice, 1969, pp. 163ff.

correlation is non-conventional. He does say, however, that correlations between utterances and responses may be "iconic" and "associative" as well as "conventional" (1969, p. 163). It must be the case, then, that "iconic" and "associative" correlations are non-conventional; however, the sense in which "iconic" and "associative" correlations are non-conventional, and the sense of "conventional" itself, are never made clear. It must be the case, however, that Grice is using "conventional" in an odd way, since for other philosophers (for example, Stevenson) for x to be a sign used in communication, x must be, in some sense, conventionally correlated with audience responses. It is not clear (especially in the light of Grice's criticisms of the causal theory of meaning), how an utterance could be non-conventionally correlated with audience responses, and still be an utterance, in Grice's sense.

Grice would not have been faced with the need of distinguishing between "modes of correlation" had he provided separate analyses of (a) meaning with respect to nonlinguistic utterances, and (b) meaning with respect to linguistic units. As it is, Grice is forced to make ad hoc distinctions, which are never clearly spelled out, between conventional and non-conventional modes of correlation between utterances and audience responses.

The second way in which Grice's concern for the meaning of utterances rather than linguistic units leads to trouble is that it obscures features of meaning which are

peculiar to the kind of meaning attaching to linguistic units, and not to the kind attaching to nonlinguistic utterances.

In order to show precisely which features of the meaning of linguistic utterances are obscured by Grice, consider Ziff's remarks on Grice's theory of meaning:

Grice's analysis rings untrue. It was bound to; his alloy lacks the basic ingredient of meaning: a set of projective devices.¹¹ The syntactic and semantic structure of any natural language is essentially recursive in character. What any given sentence means depends on what (various) other sentences in the language mean (Ziff, 1967, p. 7).

As Patton and Stampe correctly observe, in their reply to Ziff, what Ziff is claiming is, in effect, that "the primitive locus of meaning is in language" (Patton and Stampe, 1969, p. 4). Patton and Stampe seem ready to acknowledge, as everyone does, that "the semantic system of a natural language involves such a projective capacity" (1969, p. 2), but claim that Ziff's criticism misses its mark, since Grice's analysis was supposed to be an analysis of the meaning of utterances in general, rather than the meaning of units of natural languages. Patton and Stampe appear to be claiming that Grice can be excused for not including "a set of projective devices" in his analysis of meaning, because such a set is essential to semantic systems of natural languages, but not to the meaning of nonlinguistic utterances. But surely, if Grice's analysis is to succeed in being fully

¹¹The semantic structure of language "contains a set of projective devices" just in case it has the capacity "to enable its speakers to know what a person would mean were he to say something that has never before been said" (Stampe, 1968, p. 172).

general (as Grice wants it to be), then the analysis of the meaning of "mean" must include an analysis of the meaning of "mean" in linguistic contexts, and not just in nonlinguistic contexts.

Patton and Stampe correctly observe that two legitimate concerns may be distinguished within the theory of meaning: (1) an interest in the projective character of language, and (2) an interest in the meaning of "mean." They claim that Grice is concerned with (2), while Ziff is concerned with (1). But, again, if Grice's analysis of the meaning of "mean" is to be fully general, then it must account for the meaning of "mean" in linguistic contexts. And, if (as Patton and Stampe acknowledge) one ingredient in the account of the meaning of "mean" in linguistic contexts is a set of projective devices, and if (as is the case) Grice's analysis of the meaning of "mean" makes no mention of projective devices, then his analysis is either inadequate or not fully general.

My second general criticism of Grice's analysis of meaning concerns Grice's claim that what words and sentences mean must be analyzed in terms of what people mean when they utter the words and sentences. This cannot be right.

That this is false may be shown in two ways. First, as Stampe observes, "a sentence may have meaning even if it has never before been uttered, in which case it is not true that 'people mean' (or that 'one means') anything by it" (1968, p. 171). Stampe's reply, in defense of Grice's theory,

is that the phrase "what people mean when they utter the words and sentences" does not refer to historical occurrences, but rather to what one would mean were one to utter the words and sentences. The person in question is "the ideal speaker," not necessarily any actual speaker (Stampe, 1968, pp. 171-172). Therefore, Grice must be claiming that the meanings of words and sentences must be analyzed in terms of what the ideal speaker means by the words and sentences uttered by him.

The trouble with this reply is that the ideal speaker is not, as Stampe acknowledges, to be met with in experience. Therefore, one of the apparent advantages of Grice's theory is lost as soon as it is recognized that it is the intentions of the ideal speaker which determine the meanings of words and sentences, instead of the intentions of actual speakers. The intentions of actual speakers are (according to Stampe) discoverable independently of any knowledge of the meanings of his utterances;¹² but there is no way to discover the intentions of the ideal speaker without knowing the meanings of his utterances. In Stampe's felicitous terms, the notion of the intentions of the ideal speaker, unlike that of the intentions of actual speakers, "swims the same orbit of conceptual space as does 'meaning' itself." If the intentions of the ideal speaker can only be discovered via an

¹²y Stampe, 1968: "The great gain of Grice's analysis is that, whatever its difficulties, the concept of intention, unlike the concepts of an idea, a concept, semantic marker, semantic regularity, and so forth, at least does not swim the same orbit of conceptual space as does 'meaning' itself" (pp. 167-168).

investigation of his utterances, there is little point in insisting that the meanings of utterances are analyzable in terms of the intentions of the ideal speaker. Since the only evidence we can possibly have of the intentions of the ideal speaker are based on the meanings of utterances which the theory says are his, why not concentrate on investigating the meanings of utterances instead of the intentions of the ideal speaker?

The second way in which the falsity of Grice's claim that what words and sentences mean must be analyzed in terms of what people mean when they utter the words and sentences may be shown is suggested by an example of Ziff's, introduced in Chapter 1. A man cries out "Gleeg, gleeg, gleeg," intending to induce in his audience the belief that it is snowing in Tibet, by means of the recognition of his intention. According to Grice's analysis, the man means_{nn} that it is snowing in Tibet, and, it seems, the utterance "Gleeg, gleeg, gleeg" must mean_{nn} (have as its "timeless meaning") "It is snowing in Tibet." Ziff claims that Grice's analysis must be wrong because "Gleeg, gleeg, gleeg" does not mean in any language "It is snowing in Tibet." As Patton and Stampe observe, Ziff identifies what a person's utterance means with what Grice calls the timeless meaning of the sentence he uttered (with his utterance meaning nothing at all if no such timeless meaning is to be found) (Patton and Stampe, 1969, p. 15). And, Ziff's "identification" here, Patton and Stampe claim, is what he must use to justify

his claim that Grice's analysis is wrong. That is, Ziff claims that since "Gleeg, gleeg, gleeg" means nothing, and since the timeless meaning of an utterance is the same as the meaning of the utterance, "Gleeg, gleeg, gleeg" must have no timeless meaning; and, since Grice's analysis accords "Gleeg, gleeg, gleeg" a timeless meaning, it must be wrong.

Patton and Stampe have replied that while it is true that "Gleeg, gleeg, gleeg" has no timeless meaning, it is also true that a person can mean_{nn} something by an utterance, even when the utterance uttered has no timeless meaning (1969, p. 11). Grice himself, however, could not agree on this point with Patton and Stampe: Grice's exegetes cannot extricate him from the inconsistency noted by Ziff. For, according to Grice, the timeless meaning of an utterance is determined by what the person meant_{nn} by the utterance. It is impossible, according to Grice's theory, for an utterance to have no timeless meaning while the speaker meant_{nn} something by the utterance. Therefore, either "Gleeg, gleeg, gleeg" has as its timeless meaning "It is snowing in Tibet," or Grice's analysis is inconsistent. Patton and Stampe cannot have their cake (that "Gleeg, gleeg, gleeg" means timelessly nothing) while swallowing Grice's theory whole.

Stampe has also suggested, however, that we may say that "Gleeg, gleeg, gleeg" does have a timeless meaning, and that its timeless meaning is in fact "It is snowing in Tibet." Thus, he says that when "we suspend our preoccupation with the nonsensicality" of the madman's utterance, and "become

interested in cataloguing and studying the vocal behavior of such patients," then we "might well record" that "Gleeg, gleeg, gleeg" means "It is snowing in Tibet." Stampe says that the claim "'Gleeg, gleeg, gleeg' means 'It is snowing in Tibet'" may be paraphrased as "By 'Gleeg, gleeg, gleeg' the madman meant 'It is snowing in Tibet'" (Stampe, 1968, p. 174).

This paraphrase reveals the fact, however, that Stampe has not shown that "Gleeg, gleeg, gleeg" has any timeless meaning, much less that its timeless meaning is "It is snowing in Tibet." That Stampe thinks his reply to Ziff's example is adequate reveals the fact that he has simply forgotten the claim of Grice's which is in question. The claim in question is that the timeless meaning of utterances may be analyzed in terms of utterer's occasion-meaning. Stampe has merely shown (as his paraphrase reveals) that utterance-type occasion-meaning may be analyzed in terms of utterer's occasion-meaning--a claim which would be hard, indeed, to deny. Stampe's account of Ziff's example does not begin to show that the timeless meaning of utterances must be analyzed in terms of utterer's occasion-meaning.

Stampe has, however, supported in an ingenious way Grice's claim "that it is basically persons that mean things, and not expressions" (Stampe, 1968, p. 166), and that therefore the meaning of expressions must be analyzed in terms of what people mean when they utter the expressions. Stampe has shown, in brief, that a grammatical analysis of English sentences containing the word "meaning" shows that in

ordinary language the thesis that what an expression means may be identified with what a person means by that expression is, in fact, true. Stampe's argument is a highly sophisticated example of ordinary-language analysis: what it claims is that if one is interested in saving the ordinary meaning of "meaning" in a philosophical analysis of "meaning," then what an expression means must be identified with what a person means by that expression. "The ordinary meaning of 'meaning' is revealed, not by appeal to the analyst's intuitions, but by appeal to the principles of transformational grammar.

As Stampe acknowledges, however, "the murky question remains": granting that Grice's theory of meaning is "an expression of grammatical intuition, is the grammar--is English--in this regard philosophically perspicuous or philosophically misleading?" (1968, p. 166). It is, however, this "murky question" with which we are concerned here; since Stampe nowhere attempts to answer it, his argument, however ingenious, is irrelevant to our concerns.

The upshot of the above discussion is, first, that it is wrong to analyze the timeless meaning of utterances in terms of what people mean_{nn} by their utterances. For, as Ziff's example shows, what people mean_{nn} by their utterances (i.e., what utterers' occasion-meanings are) may be completely unrelated to the timeless meanings of their utterances. The timeless meanings of utterances are determined by the semantic structure of the languages in which the utterances occur; and, however the semantic structure of language is analyzed,

it is clear that it cannot be done solely in terms of individual speakers' intentions.

As we have seen, it is doubly wrong to analyze the timeless meaning of utterances in terms of what people mean by their utterances, if what people mean by their utterances is analyzed in terms of their intentions: as Ziff notes, such a program attempts to show that "good intentions [i.e., the right Gricean intentions] suffice to convert nonsense to sense" (1967, p. 5). As we shall see more clearly in a later chapter, what the timeless meaning of a linguistic unit is depends on factors other than speakers' intentions.

There have been, of course, various answers proposed to the question as to what are the independent variables on which the timeless meaning of linguistic units depend. Proponents of particular answers to this question have criticized Grice for ignoring the variables which they think are relevant. For example, it might be contended that the timeless meaning of linguistic units depends on certain naturalistic objects and/or processes, or on relations between naturalistic objects and/or processes (e.g., stimuli and dispositions to behave in certain ways). Quine's analysis (or an extension thereof, to be considered later) may be construed as an attempt to justify such a contention. It might also be contended that the timeless meaning of linguistic units depends on the linguistic conventions which define the semantic structure of the language to which the linguistic units belong. We shall have occasion, in the

next chapter, to explore this possibility. For now, it is only necessary to point out that according to Grice neither naturalistic objects and/or processes, nor linguistic conventions, are directly relevant to the meaning_{nn} of utterances; only speakers' intentions are relevant. As we shall see, however, in order to give any cash value to the notion of speakers' intention, it would seem that Grice must admit that either naturalistic acts and/or processes or linguistic conventions are relevant to the timeless meaning of linguistic units.

The argument for the claim that the notion of speakers' intention must, if it is to be of any help to the theory of linguistic meaning, be cashed out in terms of linguistic convention, will be examined in the next chapter. In the succeeding paragraphs, the argument for the claim that the notion of speakers' intention must be cashed out in terms of naturalistic objects and/or processes is briefly presented.

Grice claims that "U meant_{nn} something by uttering x" is true iff U uttered x with the right Gricean intention. The question immediately arises as to how we are to find out whether U uttered x with the right Gricean intention. Grice recognizes this problem, and says that, in general, "the criteria for judging [i.e., identifying] linguistic intentions are very like the criteria for judging nonlinguistic intentions" (1967, p. 48). As we shall see, even according to Grice, the criteria for identifying linguistic and

nonlinguistic intentions are grounded in observation of the overt behavior (or dispositions to overt behavior)--which are naturalistic objects and processes--of the agent and audience.

Thus, Grice says that in some cases utterances are accompanied by explicit "conscious plans" or "formulations of intentions," in which case utterer U's intention in uttering x is simply what he says it is. In such cases, if we consider only statements by a speaker of his intentions, and accept these statements at face value, then hypotheses about the meanings of utterances cannot be false (so long as we correctly transcribe the speaker's statements). Since meanings are, by definition, intentions, there cannot be any disagreement between candid statements of intentions and meanings. But if statements of intentions are taken to be the only evidence for meanings, and if there can be no discrepancy between statement of intention and meaning, it is perverse to insist that meanings are anything other than overt statements of intentions.

Of course, most utterances are unaccompanied by explicit declaration of intention. In these cases, Grice says, three methods are available for identifying the speakers' intentions: either we conclude that what the speaker intends by utterance x is what is "normally" intended by speakers uttering x (where what is "normally" intended has already been determined somehow), or we ask the speaker what he intended, or we attempt to determine by the

(verbal and non-verbal) context of the utterance what the speaker's intention is (Grice, 1967, p. 48). In other words, what Grice is saying is that we use the speaker's overt behavior (in conjunction with other relevant information about the overt behavior of the audience and about the context of the utterance) as a guide to his intentions, and that we use regularities of speakers' behavior to establish what is "normally" intended by utterances.

Grice has proposed, then, a mentalistic analysis of meaning, while admitting that the speakers' intentions, in terms of which meaning is analyzed, must be identified by observation of overt behavior of speaker and audience. In the face of this admission, however, it seems perverse to insist that meaning must be analyzed in terms of mentalistic concepts, such as that of speakers' intention, as Grice continues to do. It would seem simpler and more accurate to dispense with reference to the mentalistic intentions, and admit that the meanings of utterances are directly analyzable in terms of naturalistic objects and processes, such as the overt behavior of speakers.

To this proposal to dispense with reference to intentions, and to treat as meanings the naturalistic objects and processes which are taken by Grice to be mere evidence for the existence of underlying intentions, it will be objected that we have confused the evidence for the theory of meaning with the proper subject matter of the theory. The reply is that since the only evidence we can have--and

the only evidence adduced by Grice--for the existence of speakers' intentions (and hence for meanings in his mentalistic sense) is naturalistic objects and processes--such as speakers' overt behavior--it is incorrect to suppose that meanings are anything other than the naturalistic objects and processes. Until we have independent evidence for the existence of mentalistic speakers' intentions, we have no choice but to analyze meanings in terms of naturalistic objects and processes.

Grice has said that he is "not sympathetic towards any methodological policy which would restrict one from the start in an attempt to formulate a theory of meaning in extensional terms" (1968, p. 242). And, neither is Grice sympathetic toward a policy which holds that intensional concepts such as that of intending and believing are irreducibly embedded in the theory of meaning. He admits that even if it is true that intensionality is embedded in his theory of meaning, "one is not . . . precluded from being, in at least one important sense, an extensionalist" (1968, p. 242). One could be an extensionalist in the sense that one could claim (a) that the (intensional) psychological concepts which are needed, according to Grice, in the formulation of an adequate theory of meaning are not "among the most primitive or fundamental psychological concepts," and (b) that it is possible to derive the "intensional concepts" from certain other more primitive extensional concepts (Grice, 1968, p. 242).

If it is conceivable that concepts such as speaker's intention can be derived from extensional concepts such as speaker's and hearer's overt verbal behavior, then it is perverse to insist that any adequate analysis of meaning must rely on a notion of speaker's intention, and not on concepts such as speaker's and hearer's overt behavior. But, if Grice's insistence on the necessity of reference to speaker's intention in analyses of meaning_{nn} is abandoned, then the essential element which distinguishes Grice's theory of meaning from others is obscured, and Grice's theory collapses into something approximating what we have called a naturalistic theory of meaning. This point will be argued in greater detail in Chapter 6.

CHAPTER 4

MEANING: SEARLE'S ANALYSIS

4.1 Language as Rule Governed Behavior

In this chapter, Searle's analysis of meaning is examined. In the first section, we deal with Searle's distinctions between (a) institutional and brute facts, and (b) rule governed and non-rule governed behavior, and with his conception of language as a form of rule governed behavior. My arguments concerning this conception of language are of considerable importance, as they will be used, in Chapter 6, to show that Searle's analysis of meaning in terms of rules does not represent an advance over the naturalistic theory of meaning. In the second section of this chapter, Searle's distinctions between several kinds of speech acts will be discussed. Finally, in the third section, Searle's analysis of meaning will be examined, and contrasted with Grice's analysis.

Searle acknowledges that his theory of meaning presupposes a general "picture . . . of what constitutes the world and consequently what constitutes knowledge about the world" (1969, p. 38), which is not shared by all or even most philosophers. Searle calls the "picture" of what constitutes the world which he rejects the "classical

picture": the classical picture pictures "the world as consisting of brute facts, and of knowledge as really knowledge of brute facts" (1969, p. 50). Searle's own picture of the world is described only in negative terms, as claiming that "there are many kinds of facts . . . which are hard, if not impossible, to assimilate" to the classical picture (Searle, 1969, p. 51). That is, Searle's picture is simply the denial of the classical picture: i.e., that it is not the case that the world consists solely of brute facts.

Searle's distinction between his own picture of the world and the classical picture of the world rests on G. E. M. Anscombe's distinction between brute and institutional facts (Anscombe, 1958). According to Searle, brute facts are simply "physical movements, states, [or] raw feels" (1969, p. 51). They are facts which are expressed by "statements about physical or psychological properties of states of affairs": statements (or sets of statements) which either are or are reducible to (sets of) statements about such properties of states of affairs are statements of brute facts (1969, p. 51). The facts expressed by "This stone is next to that stone," "Bodies attract with a force inversely proportional to the square of the distance between them and directly proportional to the produce of their mass," and "I have a pain" are examples, according to Searle, of what the proponents of the classical picture call brute facts.

What Searle calls "institutional facts" are facts which are not expressible by statements (or sets of

statements . . ., or statements which are reducible to simple (sets of) statements . . .) about physical or psychological properties of states of affairs (Searle, 1969, p. 51).

Facts such as Mr. Smith marrying Miss Jones, the Dodgers beating the Giants three to two in eleven innings, Green being convicted of larceny, and Congress passing the Appropriations Bill are examples of what Searle calls institutional facts.

The difference between brute and institutional facts is obviously not that the former are but the latter are not constituted by "physical movements, states, and raw feels." On the contrary, institutional facts do involve "a variety of physical movements, states, and raw feels" (Searle, 1969, p. 51). The point is that institutional facts, but not brute facts, are "underlain" by human institutions. That is, according to Searle, the existence of certain institutions makes possible, and explains, the existence of institutional facts: e.g., the institution of baseball makes it possible that the Dodgers could beat the Giants three to two in eleven innings (Searle, 1969, p. 51).

For reasons which will be discussed shortly, Searle says that the institutions which underlie institutional facts are "systems of constitutive rules" (1969, p. 51). Thus, it is the existence of such systems of constitutive rules which is held to make possible the existence of institutional facts, and it is by reference to such systems of constitutive rules that institutional facts must be explained.

In order to see why this is so, Searle asks us to consider a description of a football game "in statements only of brute facts." According to Searle, the description could include a large corpus of brute facts, as well as statements of what he calls "brute regularities" (e.g., that "at statistically regular intervals organisms in like colored shirts cluster together in a roughly circular fashion (the huddle)" (1969, p. 52)). But, according to Searle, the description could not be a description of the game of football, because "all those concepts which are backed by constitutive rules, such as touchdown, offside, game, points, first down, time out, etc." cannot be included in the description (1969, p. 52). And, Searle says that in order to explain the brute facts and regularities of football, reference must be made to certain institutional facts; and these institutional facts "can only be explained in terms of the constitutive rules which underlie them" (1969, p. 52).

It should be clear that there are regularities which may be observed in the playing of football, even though football is what Searle wants to call a form of rule governed behavior. The difference between non-rule governed behavior and rule governed behavior is not that the former is, but the latter is not, regular; as the football example shows, rule governed behavior is regular. In fact, all rule governed behavior is regular, although not all regular behavior is rule governed.¹ There is some extra ingredient present

¹This thesis is also suggested by Jonathan Bennett's

in rule governed behavior, which is not present in "merely regular" behavior: as we have seen, Searle claims that the extra ingredient is that the regularities of rule governed behavior are underlain by systems of constitutive rules, while the regularities of "merely regular" behavior are not so underlain. I shall argue shortly that the differences alleged by Searle to obtain between rule governed behavior and regular behavior are not, in fact, real differences, and that therefore Searle has not shown there to be a distinction between the two kinds of behavior.

Thus, according to Searle, there are two kinds of facts: some facts are underlain by human institutions and some are not. Those facts which are underlain by human institutions are made possible by the institutions, and those facts which are not so underlain are independent of human institutions. Facts which are underlain by institutions can only be explained in terms of the underlying institutions; facts which are not so underlain are explicable without reference to any institution. Facts which are underlain by human institutions are institutional facts, and those which are not so underlain are brute facts.

Searle believes that the facts of language and linguistic behavior with which he is concerned are institutional facts: this is the best way to understand his claim

discussion of his well-known example of the dance of the bees (1964, pp. 15-16). Searle's distinction between rule governed and non-rule governed behavior owes much to Bennett's distinction between regular and rule-guided behavior (Bennett, 1964, pp. 8-21).

that "speaking a language is engaging in a rule-governed form of behavior" (Searle, 1969, p. 22). Linguistic behavior is rule governed in the sense that the facts and regularities of linguistic behavior are underlain by systems of constitutive rules (i.e., institutions), which both make possible and explain the facts and regularities.

Constitutive rules are distinguished by Searle from what he calls "regulative rules." Roughly, "regulative rules regulate antecedently or independently existing forms of behavior," while constitutive rules "constitute (and also regulate) an activity the existence of which is logically dependent on the rules" (Searle, 1969, pp. 33-34). The main difference, therefore, between regulative and constitutive rules is that the latter, but not the former, "create the possibility of new forms of behavior."²

When Searle says that speaking a language is engaging in a form of rule governed behavior, what he means, then, is that linguistic behavior is created and regulated by a system of constitutive rules. Generally, all human institutions, in Searle's sense (such as various games, marriage, court proceedings, political behavior, etc.), are created by systems of constitutive rules: that is, it is possible to say that one is engaging in institutional behavior (such as playing a game, getting married, etc.) only if there exist constitutive rules governing the behavior. And, since speaking a language is, according to Searle, engaging in

²For other differences, see Searle, 1969, pp. 33-42.

institutional behavior, therefore there must exist constitutive rules governing linguistic behavior.

The sense in which speaking a language is to engage in rule governed behavior explains the sense in which language is, for Searle, a matter of convention. In order to clarify this issue, Searle compares linguistic behavior with fishing, which he says is not conventional. Searle claims that the crucial difference between fishing and speaking a language is that the "means-ends relations" between my utterances (e.g., "I promise") and my goals (e.g., my desire to make a promise) are matters of convention, while the means-ends relations between my actions while fishing (e.g., using barbed hooks) and my goals (e.g., my desire to catch fish) are matters of natural physical facts, techniques, strategies, and procedures, but not of convention (1969, p. 37).

Thus, Searle claims that although speaking and fishing are similar in the sense that they are both purposive or goal directed behavior, they are dissimilar in the sense that the means-ends relations involved in the two cases are dissimilar. For example, in fishing I may use a particular kind of bait because it is a "natural physical fact" that the kind of fish I want to catch are often attracted to it; I may use a particular kind of hook because it is a "natural physical fact" that the kind of fish I want to catch are often secured on these hooks; and I may adopt a particular strategy in working the stream, because such a strategy has worked before. This is to say that the relations between

my means and my goals in fishing are matters of "natural physical facts," strategies, etc.; the relations are not conventional. In the case of speaking, however, the relations between my means and my ends are, according to Searle, matters of convention. Because one of the constitutive rules underlying language is that "the utterance of such and such expressions under certain conditions counts as the making of a promise," if I want to make a promise then I will (or should) utter the relevant expressions in the relevant circumstances. The means-ends relation in this case is conventional, because it is based on the underlying constitutive rule defining and regulating the speech act of promising. There are no analogous rules underlying the facts, strategies, and techniques of fishing.

Therefore, language is conventional for Searle in the sense that speaking a language is engaging in rule governed behavior. True, Searle admits there are regularities of linguistic behavior, but these regularities are held to be based on linguistic conventions--i.e., the constitutive rules underlying language. The regularities of language are regularities of institutional facts rather than brute facts, and, as such, are created by and are explicable solely in terms of the underlying rules. There are regularities of fishing as well; but, since these regularities, unlike the regularities of language, are regularities of brute facts, these regularities are explicable in terms of generalizations governing the facts, strategies, and techniques of fishing. No appeal to rules is necessary in explaining the fishing behavior of

fishermen, while such an appeal is necessary in explaining the linguistic behavior of speakers.

It so happens, however, that wherever Searle sees a rule which he says governs a bit of linguistic behavior, it is also possible to say that the bit of behavior is regular, and that it fits in with other observed regularities of linguistic behavior. And, where there are regularities, generalizations may be constructed which "cover" and, if the generalizations are lawlike, explain the behavior in question. Searle is committed, however, to the position that rules rather than regularities must be used in explaining linguistic behavior--even though he admits that linguistic behavior is regular. The importance of Searle's main hypothesis, that speaking a language is engaging in rule governed behavior, lies in the corollary of this hypothesis, according to which linguistic behavior can only be explained in terms of the rules underlying the behavior. Therefore, if it were true that the regularities of language could be explained without recourse to rules, then the importance of Searle's main hypothesis would be considerably diminished. Since it is true that wherever Searle sees a rule governing linguistic behavior there is also a regularity covering that behavior, Searle's argument for the claim that rules rather than regularities explain linguistic behavior is crucial to his thesis.

As was pointed out in Chapter 1, one difference between Searle's theory of meaning and naturalistic theories of

meaning is that Searle thinks that linguistic behavior is fundamentally different from other forms of what may be called "natural" behavior (e.g., fishing), while according to naturalistic theories (such as Quine's) linguistic behavior is not fundamentally different from other forms of behavior: all behavior is "natural." The point is that Searle thinks that linguistic behavior is rule governed behavior, and that natural behavior is not rule governed; the naturalistic theory of language asserts that linguistic behavior is similar to all natural behavior, being non-rule governed in Searle's special sense.

A second difference between Searle's theory and naturalistic theories is that Searle thinks that linguistic behavior can only be explained in terms of the rules underlying the behavior, while forms of natural behavior can and must be explained without recourse to such underlying rules (because there are no such underlying rules, in the case of natural behavior). Searle claims, in other words, that what were called in Chapter 1 "naturalistic explanations" are not sufficient for the explanation of linguistic behavior, even though they may be sufficient for the explanation of forms of natural behavior. Naturalistic theories assert, on the contrary, that naturalistic explanation is sufficient in explaining all forms of behavior, linguistic and otherwise.

Obviously, one factor behind the appeal of the naturalistic theory is its simplicity and generality: it not only posits fewer kinds of behavior and fewer kinds of

explanation than does Searle's theory, but it also applies to a wider range of behavior than does Searle's theory of language. And, since wherever Searle sees rules governing linguistic behavior there are also regularities covering that behavior, then the charge gains plausibility that Searle has gratuitously multiplied kinds of behavior and kinds of explanation. Since it is true that wherever Searle sees rules a naturalist would see only regularities, then the only justification open to Searle for positing the rules would be the claim that the regularities do not explain the behavior, while the rules do. Searle, in fact, makes this claim, in his justification for positing rules. Let us examine this claim, as well as the underlying challenge, that wherever Searle sees rules a naturalist sees regularities.

Searle asks us to consider the following example, taken from phonology [presumably, similar examples are available in semantics, since, as we shall see, according to Searle "the meaning of a sentence is determined by rules . . ." (1969, p. 48)].

In my dialect "linger" does not rhyme with "singer," nor "anger" with "hanger," though from the spelling it looks as though these pairs ought to rhyme. But "linger" and "anger" have a /g/ phoneme following the /ŋ/ phoneme, "singer" and "hanger" have only the /ŋ/ phoneme, thus /sɪŋgər/ but /lɪŋgər/. If you get a list of examples like this, you will see that there is a rule: Wherever the word is formed from a verb the /g/ phoneme does not occur; where it is not so formed the /g/ is separately pronounced (Searle, 1969, p. 41).

The obvious objection to this description of the case is that when one looks at the list of examples Searle uses, what one

sees is not a rule, but a regularity, to the effect that "Wherever the word is formed from a verb" ³ Searle has a reply to this objection; he says

I want to claim that this is a rule and not just a regularity, as can be seen from the fact that we recognize departures as "mispronunciations" and from the fact that the rule covers new cases, from its projective character (1969, p. 42).

More generally,

Two marks of rule-governed as opposed to merely regular behavior are that we generally recognize deviations from the pattern as somehow wrong or defective and that the rule covers new cases (Searle, 1969, p. 42).

Thus, Searle argues that linguistic behavior must be rule governed, and not merely regular, because linguistic behavior exhibits the above two marks. We cannot explain linguistic behavior solely by appeal to regularities, because reliance on regularities cannot account for either the fact that we recognize deviations from the rule as wrong or defective or the fact that the rule covers new cases.

The basic difference between rule governed and merely regular behavior is alleged to be, of course, that there is a normative element in rule governed behavior which is not present in merely regular behavior. I shall now argue, however, that regularities are sufficient to explain both the fact that we recognize deviations from linguistic rules/

³The same objection may be raised against Searle's description of the "institution" of promising, noted above (pp. 147-148). Searle says that one of the constitutive rules of language is that "the utterance of such and such expressions under certain conditions counts as the making of a promise"; but, this "rule" may be thought of, more simply, as a generalization concerning linguistic behavior.

regularities as defective cases and the fact that the rules/regularities cover new cases; it will be contended that no appeal to normative concepts such as rule-obeying is necessary in explaining these facts.

The first point in the argument is that it is true that wherever Searle sees a rule governing linguistic behavior there exists (if Searle's identification of the rule is correct) a corresponding regularity covering that behavior. This point was made earlier, in generalized form, in discussion of Searle's football example (p. 145).

The second and crucial point to my reply to Searle is my claim that the regularities covering linguistic behavior, as those covering and explaining all natural behavior, must be construed as laws rather than what have been called accidental generalizations (Goodman, 1965, pp. 17ff). It is the case that wherever Searle sees a rule governing linguistic behavior there exists (if Searle's identification of the rule is correct) a corresponding law covering that behavior.

Once the regularities covering linguistic behavior are recognized as laws rather than accidental generalizations, we have no trouble in explaining the facts (a) that we recognize deviations from the law as defective cases, and (b) that the laws are "projective" (i.e., cover new cases). Thus, to use Searle's example, if it has been established as being generally the case that

- (L1) Wherever the word is formed from a verb the /g/ phoneme does not occur; where it is not so formed the /g/ is separately pronounced,

then /fɪŋgər/ and /aɪŋgər/ are recognized as deviations from this general law. I see no reason why such deviations cannot be called "mispronunciations." Also, once (L1) had been drawn from observation of speech behavior, we could recognize new cases covered by it. Consider a foreign linguist who had come to construct (L1), without ever having come across the word "anger." He then observes the inscription "anger" (and, let us assume, is able to determine that "anger" is not a verb). Because he knows the general law governing the use of /g/ in this case, and because laws are, as is well-recognized (Goodman, 1965, pp. 84-122; Hempel, 1965, p. 342), projective, he will be justified in hypothesizing that "anger" is pronounced /aɪŋgər/. The only explanation which I can think of for Searle's claim that (L1) must be a rule and not a regularity is that he must be construing all regularities as accidental generalizations, which are, admittedly, not projective. But, since general laws are projective, it is unnecessary to construe (L1) as a rule rather than a lawlike regularity. Construing (L1) as a lawlike regularity has the advantages of simplicity and generality, as previously mentioned, although it has the disadvantage (from Searle's point of view) of rendering gratuitous and misleading Searle's positing of rule governed behavior and explanations in terms of rules as fundamentally distinct from natural behavior and naturalistic explanations.

The same conclusion, that Searle's talk of rule governed behavior and explanation in terms of rules is

gratuitous, may be reached via a different argument, which has been proposed by Wilfrid Sellars. Consider Searle's hypothesis that speaking a language is engaging in rule governed behavior. It seems plausible to infer from this that according to Searle learning to use a language (L) is learning to obey the rules of L. Sellars argues, however, that this claim cannot be correct; the argument is a reductio, and runs as follows:

Assume the thesis, that learning to use L is learning to obey the rules of L (which is held to follow from the more general thesis that speaking L is engaging in rule governed behavior).

But, a rule which enjoins the doing of an action (A) is a sentence in a language which contains an expression for A.

Hence, a rule which enjoins the using of a linguistic expression (E) is a sentence in a language which contains an expression for E--in other words, a sentence in a metalanguage.

Consequently, learning to obey the rules for L presupposes the ability to use the metalanguage (ML) in which the rules for L are formulated.

So that learning to use a language (L) presupposes having learned to use a metalanguage (ML). And by the same token, having learned to use ML presupposes having learned to use a metametalanguage (MML) and so on.

But this is impossible (a vicious regress).

Therefore, the thesis is absurd and must be rejected (Sellars, 1954, p. 321).

As Sellars recognizes, there is a ready reply to this argument. The reply is that "learning to obey the rules of L" must be construed as "learning to conform to the rules of L." This reply escapes the regress, while attempting to preserve the essential claim of the general thesis that speaking a language is engaging in rule governed behavior. The

reply does escape the regress, because "'conforming to a rule enjoining the doing of A in circumstances C' is to be equated simply with 'doing A when the circumstances are C'--regardless of how one comes to do it" (Sellars, 1954, p. 322). The reply does not preserve the essential claim of Searle's general thesis, however, because, as Sellars observes, conforming to rules, in the sense defined, cannot be admitted by the defender of the distinction between rule governed and merely regular behavior to be an adequate account of engaging in rule governed behavior (Sellars, 1954, p. 323). The defender of the distinction cannot agree that it is an adequate account because the putative normative aspect of engaging in rule governed behavior is not present in the notion of conforming to rules. For, all it is to conform to a rule to do A in C, in the requisite sense, is to do A in C.⁴ Therefore, one engages in this kind of rule governed behavior simply if one generally or habitually does A in C. But there is no need to appeal to rules here.

For the above reply to Sellars' argument to be valid, "conforming to rules" must not mean doing certain actions in certain circumstances because the rules enjoin them: it must mean nothing more than habitually or generally doing certain actions in certain circumstances. But precisely because "conforming to rules" means, in the reply to Sellars'

⁴This is Sellars' account of what it is to conform to a rule, and, of course, there are others. Some non-Sellarsian account may turn out to be more adequate than this one; such an outcome would only damage the present argument if the account of what it is to conform to a rule were a non-naturalistic account. My point here is that Searle has provided no evidence in support of the claim that any adequate account of what it is to conform to a rule must be a non-naturalistic account.

argument, simply habitually or generally doing certain actions in certain circumstances, the reply does not save Searle's hypothesis that speaking a language is engaging in rule governed behavior. The point is that actions conforming to rules in the requisite sense can be explained in terms of lawlike regularities concerning the agent's habits or dispositions to perform certain actions in certain circumstances; no appeal to rules is needed in explaining rule-conforming actions, and therefore any claim which asserted the need for an appeal to rules would be mistaken. Searle is caught, therefore, in the following dilemma: either his main thesis is absurd, as shown by Sellars' argument quoted above, or he is committed to construing (1) "obeying a rule of L to do A in C" as (2) "conforming to a rule of L to do A in C," and (2) as (3) "habitually doing A when in C."⁵ If Searle chooses the second horn of the dilemma, which it seems he must, then he is faced with the charge of gratuitously positing rules, when statements of lawlike regularities would have sufficed to explain linguistic behavior. I conclude that Searle has no justification for claiming that rules, and not statements of lawlike

⁵Sellars claims to have avoided this dilemma, by offering a rebuttal to our claim that the notion of conforming to rules is not an adequate account of rule governed behavior. If Sellars' rebuttal is correct, then there is a sense in which speaking a language can be construed as engaging in a rule governed form of behavior. However, even if Sellars' rebuttal is correct, it is not clear that it may be used by Searle, because the rebuttal depends on certain metaphysical assumptions concerning the notions of rules and meaning which are not obviously consistent with Searle's position.

regularities, are needed in explaining linguistic behavior.

4.2 Speech Acts

We have seen how Searle conceives of language and speech acts as analogous to games and moves in the games. Both games and language, as institutions, in Searle's sense, are created and governed by sets of constitutive rules. On this analogy, to speak a language is to perform linguistic acts: i.e., to make moves in a game which is constituted and governed by sets of constitutive rules.

Sets of constitutive rules define the institutional behavior of speaking a language; specific constitutive rules of a language define forms of behavior called speech acts. There are different kinds of speech acts, depending on the degree of generality of the rules defining the acts. We shall be concerned here with those speech acts defined by the most general kinds of rules discussed by Searle: we will be concerned with utterance acts, propositional acts, illocutionary acts, and perlocutionary acts.⁶

Searle's theory is a refinement of Austin's original theory of speech acts (v Austin, 1962; Searle, 1968). According to Searle, the four kinds of speech acts just mentioned are supposed to be abstractions from actual utterances, such as

⁶Examples of speech acts defined by more specific kinds of rules are such acts as promising, requesting, asserting, questioning, etc. (Searle, 1969, pp. 66-67); examples of speech acts defined by even more specific kinds of rules are such acts as advising Sam to go west, advising Sam to go east, asking Sam to pass the salt, asking Sam to pass the pepper, etc. (Warnock, 1971, pp. 80-81).

the following:

- (1) Sam smokes habitually.
- (2) Does Sam smoke habitually?
- (3) Sam, smoke habitually!

According to Searle, in uttering each of these sentences the agent S performed at least an utterance act, a propositional act, and an illocutionary act; he may have performed a perlocutionary act as well. An utterance act is performed whenever anyone utters words. But in uttering words, we may do certain things: for example, in uttering (1) S made an assertion; in uttering (2) S asked a question; and in uttering (3) S gave a command. Assertions, questions, and commands are all examples of illocutionary acts. In addition, Searle says that in uttering (1), (2), and (3) S, in each case, refers to Sam, and predicates the expression "smokes habitually" ["or one of its inflections" (Searle, 1969, p. 23)] of him. In doing so, S performs a particular propositional act: the same propositional act is performed in (1)-(3), because S refers to Sam in each of the sentences, and predicates the same expression of him in each of the sentences, even though S performed different illocutionary acts in (1), (2), and (3).

Searle adds to the above three kinds of acts Austin's notion of the perlocutionary act. Searle claims that the "identity criteria" [which term he encloses in "scare quotes" (1969, pp. 2 and 24)] of illocutionary and perlocutionary acts are different; he nowhere states the identity criteria for perlocutionary acts, however. Therefore, the precise

difference between the two kinds of acts is never made clear. While it is true that the central part of Speech Acts is devoted to specifying the identity criteria for various illocutionary acts, the closest Searle comes to specifying such criteria for perlocutionary acts is to give a list of examples of such acts (Searle, 1969, p. 25). Searle says he is relying on Austin's notion of perlocutionary acts; but Austin does not give a specification of the identity criteria for perlocutionary acts either.

Nevertheless, one is able to get an intuitive feeling for the difference between illocutionary and perlocutionary acts by studying the contrasts drawn by Austin between the two kinds of acts. Illocutionary acts are originally described by Austin as "ways of using locutions": for example, in saying something [i.e., in performing a locutionary act, in Austin's technical sense (1962, p. 94)], we may ask a question, answer a question, issue a command, make a promise, etc. In doing so, we perform illocutionary acts. The performance of an illocutionary act has, in standard cases, certain effects: for example, by making a promise the speaker becomes committed by his promise. In Austin's terminology, this is to say that illocutionary acts have a certain force; and, the force or effect of illocutionary acts is conventional (1962, p. 108). Thus, making a promise commits the speaker in a particular way, because of the conventions governing promising.

Now, doing something in uttering certain words (e.g.,

in uttering "I promise" Sam made a promise) may be contrasted with doing something by uttering certain words (e.g., by giving Sam the command "Stop smoking!" I got Sam to stop smoking). Doing something in uttering words is to perform an illocutionary act; doing something by uttering words is to perform a perlocutionary act.

Obviously, both illocutionary and perlocutionary acts have consequences; the difference is that while the consequences of illocutionary acts are conventional--in the sense that the consequences come about because of the conventions governing the speech act--the consequences of perlocutionary acts are not conventional. For example, although one consequence--my having commanded Sam to stop smoking--of my uttering the sentence "Stop smoking!" to Sam is conventionally tied to my utterance, another consequence--my getting Sam to stop smoking--is not conventionally tied to my utterance. Once I have uttered the sentence, I have, eo ipso, commanded Sam to stop smoking; but I have not, thereby, gotten Sam to stop smoking. He may not stop smoking, and I may try to get him to stop smoking by means other than commanding him to stop smoking. Although there is only one way to command Sam to stop smoking--viz., to utter the sentence "Stop smoking!" or its equivalent to Sam--there are many ways to try to get Sam to stop smoking--one of them being to utter the sentence "Stop smoking!"

Thus, utterances which are illocutionary acts may be perlocutionary acts as well: they will be perlocutionary

acts if they have non-conventional consequences which are distinct from the illocutionary effects of the illocutionary acts performed. In crying "Wolf!" I perform, eo ipso, the illocutionary act of warning my hearers that there is a wolf nearby; by crying "Wolf!" I may (if my hearers take me seriously) perform the perlocutionary act of scaring or alarming my hearers.

Thus, there are two main differences between illocutionary and perlocutionary acts. As we have seen, both kinds of acts have consequences; one difference between the acts lies in the difference between how the effects of the acts are related to the acts. The effects of illocutionary acts follow conventionally from the acts, while the effects of perlocutionary acts follow non-conventionally from the acts. The second difference between illocutionary and perlocutionary acts lies in the media in which such acts can be performed. Illocutionary acts are essentially linguistic acts, and are performed, according to Searle, only in the uttering of sentences:⁷ no other kind of behavior can issue in illocutionary acts. For example, the only way to perform the illocutionary act of promising is to utter a sentence of the form "I promise X." Perlocutionary acts, however, are not essentially linguistic acts, and can be performed

⁷As shown by Warnock (1971, p. 81). In taking sentences as the units for performing illocutionary acts, Searle is disassociating his work on illocutionary acts from the work of Hare and Strawson, who take words rather than sentences as the units for performing illocutionary acts. V Hare, 1970, p. 3.

in both linguistic and nonlinguistic media. For example, I can perform the perlocutionary act of threatening a hearer either by saying "Your money or your life!" or by holding a knife at his throat; the point is that the same perlocutionary effect follows (nonconventionally) from each action. We will return to Searle's distinction between illocutionary and perlocutionary acts shortly; they have been introduced here because the notions are used by Searle in his discussion of meaning.

4.3 Meaning

The first way Searle uses the notions of illocutionary and perlocutionary acts in his discussion of meaning is to claim that illocutionary acts are the vehicles of meaning. One may perform utterance acts which are nothing more than utterances of nonsense strings of words; and, one may perform perlocutionary acts in other than linguistic media, and in ways in which it makes no sense to say that the act has meaning. Illocutionary acts, however, are always meaningful; they are the units of meaning for Searle.

Two observations may be made concerning Searle's adoption of the illocutionary act as the vehicle of meaning. First, illocutionary acts are linguistic acts: in order to perform an illocutionary act one must utter (or inscribe) strings of words. Therefore, in taking the illocutionary act as the vehicle of meaning, Searle is committing himself to restricting his attention to the theory of meaning of linguistic units. Although Searle goes on to develop a Gricean theory

of meaning, one important difference between his theory and Grice's should already be apparent: namely, while Grice was concerned to develop a general theory of meaning, intended to apply, as a sub-case of the general theory, to linguistic units having meanings, Searle is concerned to develop a theory of meaning covering only the cases of linguistic units having meanings. By restricting his attention to the theory of the meaning of linguistic units, Searle escapes the methodological criticisms leveled against Grice's theory in Chapter 3.

The second observation concerning Searle's adoption of the illocutionary act as the vehicle of meaning is that in doing so Searle is diverging from Grice's theory of meaning in a second way, namely, in not taking some version of what was called "utterer's meaning" as the basic unit of meaning. It will be recalled that Grice argued that what words and sentences mean is best analyzed in terms of what people mean when they utter the words or sentences. Searle is not, however, interested in the phenomenon of people meaning things; he is interested in the meaning of illocutionary acts. And, since in order to perform an illocutionary act one must utter a string of words, what Searle is interested in is the meaning of strings of words. In taking strings of words rather than people as the basic unit of meaning, Searle avoids those criticisms of Grice's theory developed in Chapter 3 which gave reasons to believe that Grice was mistaken in thinking that what words and sentences mean must

be analyzed in terms of what people mean when they utter the words and sentences.

It may be noted in passing that Searle's adoption of sentences (since illocutionary acts are the vehicles of meaning, and sentences are the vehicles of illocutionary acts) as the basic unit of meaning is consistent with Quine's views on the issue. It remains to be seen, however, whether Searle's labeling of the sentences which are the vehicles of meaning as "illocutionary acts" represents a worthwhile advance beyond Quine's theory of meaning.

Searle goes on, after adopting the illocutionary act as the basic unit of meaning, to claim that Grice's theory of meaning may be used as a basis for developing an adequate theory of the meaning of illocutionary acts. Although Searle thinks Grice's theory is a useful foundation for developing an adequate theory, he also thinks it to be inadequate as it stands. Searle's own theory of meaning is developed out of criticisms of Grice's theory.

The analysis of meaning of Grice's which Searle criticizes is the following:

- (1) Speaker S means_{nn} something by X =
 - (a) S intends (i-1) the utterance U of X to produce a certain perlocutionary effect PE in hearer H.
 - (b) S intends U to produce PE by means of the recognition of i-1 (Searle, 1969, p. 49).

The above is claimed by Searle to be a reconstruction of Grice's original 1957 analysis of meaning; there is some question, however, whether it is an accurate reconstruction. The

source of the doubt over whether (1) is an accurate reconstruction of Grice's analysis lies in Searle's use of the notion of perlocutionary effects; Searle's use of this notion is important, because he goes on to criticize Grice's analysis for including it. However, Grice's analysis does not include a reference to perlocutionary effects; only Searle's reconstruction of Grice's analysis includes such a reference. The notion of perlocutionary effect is foreign to Grice's 1957 analysis (as well as to his later revisions, some of which were made with full knowledge of Searle's criticisms of the 1957 analysis); the notion of illocutionary effects is equally foreign to Grice's analyses. Searle seems to sense that his reconstruction of Grice's analysis is unfair, because he admits that Grice's account claims "in effect . . . , that saying something and meaning it is a matter of intending to perform a perlocutionary act" (1969, p. 46, emphasis added).⁸ Rather than imputing this view directly to Grice, Searle says merely that "in the examples Grice gives, the effects cited are invariably perlocutionary" (1969, p. 46). As we shall see, Searle's misrepresentation of Grice's analysis has important bearing on the validity of Searle's criticisms of Grice's analysis.

Searle has two criticisms of Grice's theory of

⁸As we shall see, Grice not only does not say that "saying something and meaning it is a matter of intending to perform a perlocutionary act," but he also is not even primarily concerned to analyze Searle's notion of "saying something and meaning it." Searle not only misrepresents Grice's analysis, but he also (often) misidentifies the analysandum adopted by Grice.

meaning; the first is that Grice's analysis "does not show the connection between one's meaning something by what one says, and what that which one says actually means in the language" (Searle, 1969, p. 43). According to Searle, meaning is a matter of both speaker's intention and the rules or conventions of language; Grice's analysis is "useful" because it shows the connection between meaning and speaker's intention; it fails, however, because "it fails to account for the extent to which meaning can be a matter of rules or conventions" (Searle, 1969, p. 43). Because Grice's analysis fails in this way, it is, according to Searle, too weak: it accords meaning to things which we do not want to say mean anything.

In order to establish this thesis, Searle offers the following example. U, an American soldier in the Second World War, is captured by A, some Italian troops. U wishes to get A to believe that he is a German soldier, in order to get them to release him. What U would like to do would be to tell A in German or Italian that he is a German soldier, but U does not know enough German or Italian to do that. So U attempts "to put on a show" of telling them that he is a German soldier by reciting the only line of German he knows (which is a line of a poem he had to memorize in school): "Kennst du das Land wo die Zitronen blühen?" Searle claims the above situation may be described in Gricean terms as follows: U intends₁ to produce a certain effect in A, namely, the effect of believing he is a German soldier, and

U intends₂ to produce this effect by means of their recognition of his intention₁. What U intends₁ is, more explicitly, that A should think that what he is trying to tell them is that he is a German soldier (Searle, 1969, p. 44).

The issue posed by Searle's counterexample is related to an issue raised in the last chapter, concerning the adequacy of any theory of meaning which analyzes word- and sentence-meaning in terms of utterer's meaning (and this in terms of speaker's intention). It was argued in the last chapter that it cannot be right that the meaning of words and sentences is determined by what people mean when they utter the words and sentences. Searle is arguing here that not only is it wrong to think that the meaning of words and sentences is determined by what people mean when they utter the words and sentences, but, in addition, the fact is that "what we can mean is at least sometimes a function of what we are saying" (Searle, 1969, p. 45). The "American soldier" counterexample is supposed to establish this strong claim.

It is wrong to think that the meaning of words and sentences is determined by what people mean when they utter the words and sentences because, to use Searle's example, no matter what U means by his utterance of the sentence "Kennst du das Land wo die Zitronen blühen?" what the sentence means is "Knowest thou the land where the lemon trees bloom?" And, what people mean is at least sometimes a function of the meanings of the sentences they utter, because, no matter what U's intentions are, U cannot mean by his utterance of

"Kennst du das Land wo die Zitronen blühen?" "I am a German soldier."

Searle claims that Grice's analysis of meaning can be reformulated, in order to account for the above difficulties. According to Searle, the reformulation must include the stipulations that

In the performance of an illocutionary act in the literal utterance of a sentence, the speaker intends to produce a certain effect by means of getting the hearer to recognize his intention to produce that effect; and furthermore, if he is using words literally, he intends this recognition to be achieved in virtue of the fact that the rules for using the expressions he utters associate the expression with the production of that effect (1969, p. 45).

That is, Searle is claiming that Grice is only half right: in addition to what Grice says, we must add that the intended effects of literal utterances must be conventionally associated (associated "in virtue of the rules for using the expressions") with the utterances. What effects speakers intend are determined, in the case of literal utterances, not only by speakers' intentions but by the conventional effects of the utterances. Speakers cannot intend just any effect by their utterances; they must intend, in the case of literal utterances, just those effects which are conventionally linked (by the rules governing the expressions uttered) with the expressions uttered.

Grice has responded to Searle's "American soldier" counterexample with two arguments. He has argued, first, that the example is underdescribed, and that when the details are filled in (in any one of three ways) it becomes obvious

that the example is not really a counterexample--i.e., that the case is not really one in which Grice's theory and our intuitions differ on the issue of to which utterances ought we accord meaning (Grice, 1969, pp. 160-163). Grice argues, secondly, that even if the example is a counterexample, it is only a counterexample to a part of Grice's theory. Assuming that some legitimate counterexample of the "American soldier" type can be constructed (Grice offers no reason to suppose that one could not be constructed), let us examine Grice's second argument.

In this argument in reply to Searle, Grice does not deny that "when the vehicle of meaning is a sentence (or the utterance of a sentence) the speaker's intentions are to be recognized, in the normal [i.e., literal] case, by virtue of a knowledge of the conventional use of the sentence" (1969, pp. 160-161). In other words, Grice agrees with Searle's principle that in the case of literal linguistic utterances, utterer's meaning must be understood in terms of both speaker's intention and the conventional use of the sentences uttered. But, Grice goes on to caution against drawing the conclusion that the meaning of all utterances (in his extended sense of the term) must be understood in terms of both the speaker's intention and the conventional use of the utterance uttered. Of course, there must be some mode of correlation between the utterance and its effects: in the case of literal utterances of sentences, the mode of correlation is indeed conventional, as Searle would have it;

but in other cases, the mode of correlation may be nonconventional. As Grice says,

I would like, if I can, to treat meaning something by the utterance of a sentence as being only a special case of meaning something by an utterance (in my extended sense of utterance), and to treat a conventional correlation between a sentence and a specific response as providing only one of the ways in which an utterance may be correlated with a response (1969, p. 161).

Grice admits, therefore, that his original analyses of utterer's meaning in terms of utterer's intention, and of word- and sentence-meaning in terms of utterer's meaning, were defective. He admits that reference must be made in the analyses to "modes of correlation" between utterance and effect.⁹ But, he thinks that Searle is wrong to insist that the only mode of correlation between meaningful utterances and effects is the conventional mode. It is true that the mode of correlation between literal utterances of sentences and intended effects must be conventional; but such is not true of correlations between other kinds of utterances and intended effects.

The upshot of the above discussion is, therefore, that Searle's criticism of Grice's analysis is admitted by Grice to be a valid criticism of his analysis as it applies to the special case of meaning something by the literal utterance of a sentence. Searle's criticism is not valid in general, because not all modes of correlation are conventional; but

⁹V Grice's revised analysis, made in the light of Searle's criticisms, 1969, pp. 163-164.

Grice accepts the point that, along with a reference to speaker's intention, some reference must be made in the analysis of utterer's meaning to the correlation obtaining between utterance and intended effect.¹⁰

Before considering Searle's second criticism of Grice's analysis, mention should be made of Searle's rejoinder to Grice's reply to the "American soldier" counterexample. As we have seen, intended effects are correlated in various ways, according to Grice, with utterances uttered. The intended effects of meaningful utterances are always propositional attitudes, such as belief and intention. Thus, for example, the effect intended by speaker U of his utterance to A of "The engine has stopped" is that A have the belief that the engine has stopped; the effect intended by U of his utterance to A of "Stop the engine!" is that A have the intention to stop the engine (Grice, 1969, p. 166). In the case of literal utterances of sentences, the mode of correlation between utterance and intended effect is conventional; therefore, whether U's utterance of "The engine has stopped" will have its intended effect depends on whether A knows the relevant conventions governing the use of the sentence.

Searle has replied, however, apparently on the basis of Grice's claim that not all intended effects are conventionally correlated with utterances, that the relation between intended effects and utterances of literal sentences

¹⁰Grice's reply to Searle, just reviewed, is subject to the criticisms of Grice's distinction between conventional and nonconventional modes of correlation, developed in 3.4.

is not "essentially conventional" for Grice (Searle, 1971, p. 9). According to Searle, however, the relation between intended effects and utterances of literal sentences is "essentially conventional." Searle admits that for Grice the relation between intended effects and literal utterances of sentences "just so happens" to be conventional; therefore, the issue Searle sees between himself and Grice seems to be whether the relation between intended effects and literal utterances of sentences is "essentially conventional" or merely "accidentally conventional." Searle goes on to say that this issue between he and Grice "is one of the most important unresolved controversies in contemporary philosophy of language" (1971, p. 9). I think that it should be clear, however, that the issue remaining between Grice and Searle can be explained away in terms of Grice's interest in utterances in general, and Searle's interest in specifically speech acts; therefore, it seems extravagant to insist that whatever issue still remains between Grice and Searle is "one of the most important unresolved controversies in contemporary philosophy of language."

Searle's second criticism of Grice's analysis of meaning is that while according to Grice's theory "saying something and meaning it is a matter of intending to perform a perlocutionary act," the truth is that "saying something and meaning it is a matter of intending to perform an illocutionary act" (1969, p. 46). It has already been noted that (a) Grice never says that saying something and meaning

it is a matter of intending to perform a perlocutionary act, and (b) Searle never makes clear the precise differences between perlocutionary and illocutionary acts. In the light of these two points, it is difficult to evaluate Searle's second criticism of Grice's analysis.

Searle rests much of his case for the claim that for Grice saying something and meaning it is a matter of intending to perform a perlocutionary act on the kinds of effects Grice says are intended by meaningful utterances. As we have seen, Grice thinks that the intended effects of utterances are propositional attitudes (such as belief) in the hearer. And, Searle claims that since the production of such propositional attitudes is a perlocutionary rather than an illocutionary effect of utterances, Grice is committed to the view that the utterances (qua acts) are perlocutionary rather than illocutionary acts.

It is difficult for me to criticize Searle's claim that the production of propositional attitudes in hearers is a perlocutionary rather than an illocutionary effect, since I do not know what the difference between the two is. Searle does say that he wishes to claim that the intended effect of meaning something is that the hearer should "understand" something (the "illocutionary force" and "propositional content" of the utterance); and, he claims that such understanding "is an illocutionary not a perlocutionary effect" (Searle, 1971, pp. 8-10). Searle never explains why the production of Grice's "propositional attitudes" in hearers

is a perlocutionary effect, while the production of Searle's "understanding" (which is surely a propositional attitude) is an illocutionary effect. Indeed, the difference between Grice's "propositional attitudes" and Searle's "understanding" is not at all clear to me; neither is Searle's second criticism of Grice's theory, which rests on the distinction. If there is, as Searle claims, a difference between Searle's and Grice's theory on whether the intended effects of utterances are illocutionary or perlocutionary, then it is up to Searle to show wherein the difference lies.

Searle might object, of course, that the differences between illocutionary and perlocutionary acts (and effects) have been specified as clearly as they can be. The differences are that (a) illocutionary but not perlocutionary acts are conventionally correlated with their effects, and (b) illocutionary but not perlocutionary acts are essentially linguistic acts. However, as we have seen, Grice agrees that utterances in his sense can be both linguistic and conventionally correlated with their effects. Therefore, it would appear that Gricean utterances (qua acts) can be what Searle would call performances of illocutionary acts. (It is important to remember, of course, that not all Gricean utterances are performances of illocutionary acts: some utterances are performed in nonlinguistic media, and some utterances are nonconventionally correlated with their effects.) It would seem, therefore, that Searle has failed to find a difference between his theory and Grice's concerning

the issue of whether the intended effects of utterances are illocutionary or perlocutionary.

Searle uses his two criticisms of Grice's analysis in developing his own analysis of meaning. His analysis is as follows:

- (2) Speaker S utters sentence T and means it (i.e., means literally what he says) =
S utters T and
- (a) S intends ($\underline{i-1}$) the utterance U of T produce in H the knowledge (recognition, awareness) that the states of affairs specified by (certain of) the rules of T obtain. (Call this effect the illocutionary effect, IE)
 - (b) S intends U to produce IE by means of the recognition of $\underline{i-1}$
 - (c) S intends that $\underline{i-1}$ will be recognized in virtue of (by means of) H's knowledge of (certain of) the rules governing (the elements of) T (Searle, 1969, pp. 49-50).

One explanatory comment seems called for, concerning clause (a). Searle says, just before introducing his analysis, that "the meaning of a sentence is determined by rules, and those rules specify both conditions of utterance of the sentence and also what the utterance counts as" (1963, p. 48). This is a crucial claim for Searle, but, unfortunately, it is all he says directly about the meaning of sentences simpliciter.¹¹ Therefore, (2) is not an analysis of meaning simpliciter; it is, as it says, an analysis of what it is for a speaker to utter a sentence and mean it. The meaning of a literal utterance of a sentence is determined by the constitutive rules defining and governing the use of the utterance.

¹¹Although the claim might be inferred from Searle's views that illocutionary acts are the vehicles of meaning, that illocutionary acts are performed in uttering sentences, and that illocutionary acts are constituted and governed by certain rules.

As just noted, the constitutive rules of utterances specify the conditions under which the utterances may be made (and under which they may be counted as speech acts ϕ , where ϕ ranges over such illocutionary acts as assertions, questions, requests, etc.). And, clause (a) of (2) claims that the primary intention of the speaker must be to produce in his hearers the knowledge that the conditions under which his utterance may be made actually do obtain.

I shall assume, on the basis of the earlier discussion of Searle's criticisms of Grice's analysis of meaning, that the differences between Searle's and Grice's analyses are fairly clear. It will be helpful, however, to take note of two differences obtaining between the analysanda chosen by Searle and Grice. First, Searle's choice of analysandum reveals his commitment to the priority of sentence-meaning, while Grice's choice reveals his commitment to the priority of utterer's meaning. We have already seen how for Grice sentence-meaning is to be analyzed in terms of utterer's meaning. As revealed by (2), Searle is not interested primarily in speakers meaning something by their utterances (as is Grice), but rather in speakers literally meaning what they say. Sentence-meaning has priority, in the sense that the analysis of sentence-meaning in terms of the rules governing the use of sentences must be used in analysis of what it is for a speaker to utter a sentence and mean it.

The second difference between Searle's and Grice's analysanda has been introduced earlier. It is that while

Searle is concerned to analyze the notion of sentence-meaning (since illocutionary acts are the vehicles of meaning, and the unit for the performance of an illocutionary act is the sentence), and the derivative notion of speakers uttering sentences and meaning them, Grice is concerned to analyze the notion of utterer's meaning, and the derivative notion of utterance-meaning, where "utterance" is taken in Grice's extended sense. If Searle's analysis is taken as a revision of Grice's analysis, therefore, it must be taken as a revision of only part of his analysis: as a revision of that part meant to apply to the special case of utterer's meaning something by the literal utterance of sentences.

It has already been noted that by adopting a restricted and revised analysandum, Searle avoids the two main criticisms which were leveled against Grice's theory in Chapter 3. Specifically, Searle's concern for speech acts rather than utterances in Grice's extended sense exempts Searle's Gricean analysis of meaning from my methodological criticisms of Grice's adoption of the utterance as the vehicle of meaning. And, Searle's analysis of speaker-meaning (specifically, of the notion of a speaker saying something and meaning it) in terms of sentence-meaning exempts Searle's analysis from my criticism of Grice's theory for attempting to analyze utterance-meaning in terms of utterer's meaning.

It was also contended in Chapter 3 that the timeless meanings of linguistic units must depend on something other

than speakers' intentions. We explored the possibility that the timeless meanings of linguistic units depend on certain regularities in the verbal behavior of members of groups of speakers. In this chapter, we have explored the possibility that the timeless meanings of linguistic units depends on the linguistic conventions (i.e., the constitutive rules) defining and governing the semantic structure of the language to which the linguistic units belong.¹² I hope that the arguments presented in the first section of the present chapter, concerning Searle's claim that speaking a language is engaging in a form of rule governed behavior, suffice to indicate my feeling that an analysis of meaning in terms of underlying rules does not represent an advance over a naturalistic analysis of meaning in terms of regularities of verbal behavior. This "feeling" will be defended with arguments in Chapter 6, in which the arguments developed here concerning Searle's general notion of rule governed behavior will be applied to his analysis of meaning.

¹²This is how I interpret Searle's claim that "the meaning of a sentence is determined by rules. . . ."

CHAPTER 5

"LEVELS OF MEANING" AND COMPETING THEORIES OF MEANING

5.1 Harman's Characterizations of Grice's, Searle's, and Quine's Theories of Meaning

In this chapter I discuss the ways in which the theories of meaning outlined in the earlier chapters of the thesis are competing theories of meaning, and then proceed to a discussion of the criteria to be used, in Chapter 6, in evaluating the theories of meaning. To these ends, I shall first return to Harman's discussion of the three "levels of meaning," introduced in Chapter 1; I argue that Harman's characterizations of Quine's, Grice's, and Searle's theories of meaning are incorrect. In the second section of this chapter, I argue that the above three theories of meaning are in fact competing theories of meaning, and that what I have called Harman's "reconciliation thesis" concerning the theories is false. In the third section I discuss the criteria to be used in evaluating the conflicting theories of meaning. The criteria adopted were originally proposed as criteria for evaluating competing scientific theories, and, in the final section of this chapter, use of these criteria in evaluation of competing philosophical theories of meaning

is defended; in addition, my choice of criteria is defended against an alternative method of evaluating and justifying rival grammars, proposed by Chomsky.

First, then, Harman's characterizations of the three theories of meaning will be reviewed and criticized. Two points will be argued for in this section: (a) that even on assumptions made by Harman about theories of his three levels of meaning, it is not the case that Quine's, Grice's, and Searle's theories of meaning are adequately described as approaches to different levels of meaning, and (b) that Quine's, Grice's, and Searle's theories are each (at least in part) an attempt to provide a theory of what I will call the meaning of linguistic units. This second point will be made independently of any assumptions made by Harman concerning theories of his three levels of meaning.

Harman says that theory of meaning of level 1, of which Quine's theory is alleged to be a case, "attempts to explain what it is to think that p, what it is to believe that p, to desire that p, etc." (Harman, 1968, p. 593). The supposition is made by Harman that thinking is silent speech, or, in Harman's terms, that thinking is done in language (1968, pp. 591, 593). According to this supposition, to think some thought is simply to think some sentence (or other linguistic unit); in particular, for agent A to think that p is for A to think some sentence (or other linguistic unit) by which A means that p (Harman, 1968, p. 593).

Harman contends that theory of meaning of level 1 is

really a theory about the meaning of thoughts; theories about the meaning of thoughts are contrasted by Harman with theories about (a) the meaning of messages, and (b) the meaning of speech acts. However, according to Harman's supposition that thought is silent speech, the meaning of thoughts must be explained in terms of the meaning of sentences or other linguistic units. Therefore, on Harman's assumptions, theories of meaning of level 1 presupposes a theory of the meaning of linguistic units. This conclusion should be obvious: Harman says that according to level 1 theories "the problem of saying what it is to think, believe, desire, etc. that p can be reduced to the problem of saying what it is to mean that p by certain words used in thinking" (1968, p. 594); therefore, on Harman's assumptions, no theory of meaning of level 1 can be adequate unless it rests on an adequate theory of the meaning of linguistic units.

According to Harman's taxonomy, therefore, we may distinguish four different kinds of theories of meaning:

- (1) theories about the meaning of thoughts;
- (2) theories about the meaning of messages;
- (3) theories about the meaning of speech acts;
- (4) theories about the meaning of linguistic units.

I have already shown how, on Harman's supposition that thought is silent speech, theories about the meaning of thoughts presupposes theories about the meaning of linguistic units; it will also be shown that on Harman's assumptions theories about the meaning of both messages (e.g., Grice's theory)

and speech acts (e.g., Searle's theory) must presuppose theories about the meaning of linguistic units. If my argument is correct, it follows that even on Harman's assumptions it is not the case that Quine's, Grice's, and Searle's theories of meaning are adequately described as approaches to different levels of meaning; on the contrary, it is the case that on Harman's assumptions each of the three theories incorporates or presupposes an approach to one and the same level of meaning--namely, a theory of the meaning of linguistic units.

It is also possible to show, independently of Harman's assumption about the nature of thought, that Quine's theory of meaning is (or is in part) a theory of the meaning of linguistic units. Appealing to the principle that a philosopher is not always his own best interpreter ("proponents of the various theories [of meaning] have occasionally been confused about their objectives"--Harman, 1968, p. 593), Harman suggests that we disregard many of Quine's own claims about what kind of theory of meaning his is, and interpret it as a theory about the meaning of thoughts. Thus, Harman says that "in Chapter II of Word and Object Quine presents considerations mainly relevant to level 1 theories of meaning," even though "by describing language as a set of dispositions to verbal behavior, he [Quine] suggests wrongly that he is concerned with communication or more sophisticated speech acts" (Harman, 1968, p. 598). But I hope that it emerged clearly from my earlier chapter on Quine's theory of meaning that

Quine is primarily concerned with neither communication nor sophisticated speech acts nor the meaning of thoughts. Rather, Quine is, as he says himself, concerned with the stimulus meanings of sentences. And stimulus meanings are simply, according to Quine, what naturalism can salvage from the notion of meaning. In this sense, Quine is primarily concerned to develop a theory about the meaning (in his sense) of linguistic units.

Of course, as Harman would point out, sentences and other linguistic units may be used in a variety of ways. For example, it may be the case (if thought really is silent speech, as Harman assumes) that speakers use sentences in thinking (i.e., as Harman would describe it, we think sentences); it may also be the case that speakers use sentences in sending messages and communicating (i.e., as Harman would describe it, when we communicate we communicate a thought which is a thought sentence); it may also be the case that speakers use sentences in performing speech acts (i.e., as Harman would describe it, when we perform a speech act we communicate a thought which is a thought sentence). Therefore, it may be the case that a philosopher develops a theory of the meaning of sentences, in order simply to provide a base for a more general theory of either the meaning of thoughts, or of messages, or of speech acts. But the point is that according to Harman's assumptions about the nature of thoughts, communication, and speech acts, the theory of the meaning of linguistic units is basic: a theory of

meaning of thoughts, according to Harman's assumptions, can be adequate only if it rests on an adequate theory of the meaning of linguistic units; similarly, according to Harman's assumptions, theories of the meaning of messages and of the meaning of speech acts cannot be adequate unless they rest on an adequate theory of the meaning of linguistic units.

Therefore, Harman may be correct in his conjecture that in Chapter II of Word and Object Quine is interested, ultimately, in the meaning of thoughts. But it is also true that Quine's primary interest is in developing a theory of the meaning of linguistic units (i.e., in the stimulus meaning of sentences), even if this theory may be used as a foundation for a more general theory of the meaning of thoughts.

Harman says that theory of meaning of level 2, of which Grice's theory is alleged to be a case, "attempts to say what communication is and what is involved in a message's having a particular meaning" (Harman, 1968, p. 596). In his discussion of level 2 theories, Harman makes the supposition that communication is communication of thoughts (1968, p. 593). Therefore, on Harman's assumptions about thought and communication, theory of level 2 presupposes theory of level 1 in the sense that to explain what it is for A to have communicated the thought that p it is necessary to explain what it is for A to have thought that p. But we have already seen that on Harman's assumptions it is also the case that to explain the meaning of A's thought that p it is necessary

to explain the meaning of the sentence that A thought when he thought that p. Therefore, on Harman's assumptions any theory (such as Grice's) about the meaning of communicated messages presupposes a theory about the meaning of sentences or other linguistic units. And, it follows that even on Harman's assumptions Grice's theory of meaning is not adequately described as simply an approach to the meaning of communicated messages; rather, Grice's theory must incorporate, in addition to a theory about the meaning of communicated messages, some theory about the meaning of linguistic units.

As in the case of Quine's theory, it is also possible to show, independently of Harman's assumptions, that Grice's theory is (at least in part) a theory of the meaning of linguistic units. Harman says that Grice's theory is a theory about the meaning of communicated messages; he denies that it is a theory about the meaning of thoughts or of speech acts, and would presumably deny that it is a theory about the meaning of linguistic units (in order to save his thesis that Grice's and Quine's theories are approaches to different levels of meaning). But, again, I hope that it emerged clearly from my discussion of Grice's theory of meaning that his theory is (in part) a theory about the meaning of linguistic units. Thus, it will be remembered that Grice is primarily concerned to explicate the notion of an utterer meaning something by an utterance. He claims that the meanings of utterances are to be explicated in terms of what utterers mean when they utter the utterances; in addition,

Grice holds that sentences and other linguistic units are species of utterances. Therefore, it is correct to say that Grice's theory of meaning is, in part, an attempt to explain the meaning of sentences and other linguistic units.

Again, of course, sentences may be used in a variety of ways; for example, they may be used in communication. Therefore, if Harman's assumptions about the nature of communication and thought are correct, Grice's theory about the meaning of linguistic units may be used as a foundation for a more general theory about the meaning of communicated messages. Indeed, Grice is more interested in such a general theory than he is in the restricted theory about the meaning of linguistic units (even though there is evidence showing that he does not share Harman's assumptions about communication and thought). Nevertheless, it remains true--with or without Harman's assumptions--that only if Grice's theory about the meaning of linguistic units is true, can his general theory about the meaning of communicated messages be true.

Finally, Harman says that theory of meaning of level 3, of which Searle's theory is alleged to be a case, attempts to "explain how the existence of [social institutions, games, practices, etc.] can make certain acts possible" (Harman, 1968, p. 597). According to Harman, the institution in which an act is embedded confers meaning on the act, and it is because of this fact that theory of meaning of level 3 is indeed a theory of meaning (1968, p. 597). Although there

could be, presumably, theories of meaning of level 3 of non-linguistic acts, the acts with which Harman is concerned are specifically linguistic acts; this is why Harman considers only speech act theories under the heading of theories of meaning of level 3.

Now, Harman makes the assumption that in performing a speech act one must communicate something; for example, "in promising one must communicate what it is one has promised to do" (Harman, 1968, p. 593). Therefore, on this assumption, theory of meaning of level 3--of which Searle's is an example--presupposes a theory of meaning of level 2--namely, the theory of the meaning of communicated messages. But, we have already seen that according to Harman's assumptions about communication and thoughts, theory of level 2 presupposes a theory of the meaning of thoughts, and that this theory in turn presupposes a basic theory of the meaning of linguistic units. Therefore, according to Harman's assumptions, in order to explain what it is for A to have performed speech act x, it is necessary, ultimately, to explain what it is for the sentence uttered by A in his performance of x to have the meaning which it has. It follows that, on Harman's assumptions, Searle's theory of meaning is not adequately described as simply an approach to the meaning of speech acts; Searle's theory must incorporate some theory about the meaning of linguistic units.

It is also possible to show, independently of Harman's assumptions about speech acts, communication, and thought,

that Searle's theory is (in part) a theory about the meaning of linguistic units. It will be remembered from our discussion of Searle's theory that Searle explicitly acknowledges that his discussion of speech acts rests on a theory of meaning of linguistic units--namely, on his version of Grice's theory. Thus, Searle claims that "the meaning of a sentence is determined by rules . . . ," and goes on to analyze the notion of a speaker uttering a sentence and meaning it (partly) in terms of these rules (Searle, 1969, pp. 48-50).

Again, since sentences may be used in a variety of ways, and since one way sentences may be used is in the performance of speech acts, it is not unfair to say that Searle's theory of the meaning of sentences is but a part of a more general theory of the meaning of speech acts. But it cannot be denied that at least part of Searle's theory is a theory about the meaning of linguistic units; as we have just seen, Searle's general theory about the meaning of speech acts would be incomplete without the sub-theory about the meaning of linguistic units. And, only if Searle's theory about the meaning of linguistic units is true, can his general theory about the meaning of speech acts be true.

The upshot of the above discussion is, then, that both with and without Harman's assumptions about the nature of speech acts, communication, and thoughts, Harman's characterizations of Quine's, Grice's, and Searle's theories of meaning are incomplete. We have also seen that Harman's division of the theory of meaning into theories of three

levels is incomplete. A fourth kind of theory of meaning has been distinguished--the theory of the meaning of linguistic units--and it has been shown that (a) on Harman's assumptions, Quine's, Grice's, and Searle's theories are each, in part, theories of the meaning of linguistic units; and (b) on grounds independent of Harman's assumptions the theories are each, in part, theories of the meaning of linguistic units.

5.2 Three Competing Theories of Meaning

Let us turn, then, to the question of the sense in which Quine's, Grice's, and Searle's theories are competing theories of meaning. As was pointed out in Chapter 1, the main argument of Harman's paper on the "Three Levels of Meaning" is concerned to show that the common assumption, that the three approaches to the theory of meaning exemplified by Quine's, Grice's, and Searle's theories "are approaches to the same thing," is false. Harman's claim is, on the contrary, that Quine's, Grice's, and Searle's theories are theories of different levels of meaning, and are, therefore, not competing theories of meaning. This claim has been labeled Harman's "reconciliation thesis." I shall now argue that insofar as Quine, Grice, and Searle are each concerned to provide a theory of the meaning of linguistic units, and insofar as the three theories are incompatible with each other, the three theories are in fact competing theories. If my argument is sound, it will follow that Harman's reconciliation thesis is false.

The first part of the argument, that Quine, Grice, and Searle are each concerned to provide a theory of the meaning of linguistic units, has just been established. It remains, therefore, to show that the three theories are incompatible with one another. I shall argue, first, that Quine's theory is incompatible with Grice's.

The one basic conflict between Quine's theory of meaning and Grice's is ontological in nature; from this basic conflict arises a second conflict over the kind of explanation to which meaning is susceptible. The basic conflict between Quine and Grice is at bottom rather simple; baldly stated, it is that while Quine believes there to be just one kind of entity in the world--natural entities--Grice believes there to be two--let us call them entities located in the causal nexus and entities located outside of the causal nexus. And, while Quine believes that one kind of explanation suffices to explain all natural phenomena (and that there are no non-natural phenomena), Grice believes that at least two kinds of explanation are required--one to explain phenomena within the causal nexus and one to explain phenomena outside of the causal nexus.

I shall take it that Quine's naturalistic ontology is relatively clear. In summary, Quine believes that meanings (stimulus meanings) are natural entities, and that explanations of phenomena germane to meaning are of the same kind as explanations of all natural phenomena. Quine believes that issues in philosophy of language, and in the theory of

meaning in particular, are continuous with issues in what are usually called the natural sciences, and that the methods of investigation of phenomena germane to meaning do not differ in kind from the methods of investigation of all natural phenomena.

Grice's ontological position is more difficult to specify, at least partly because he has not described it, in general terms, himself. But on issues in philosophy of language, his position is clear. It is, basically, that meanings differ radically from other "natural" phenomena. Meanings must be explained in terms of speakers' intentions, while other phenomena can be explained without reference to agents' intentions. That there is an important difference in kind between explanations in terms of intentions and explanations in terms of things other than intentions emerges from Grice's original analysis of meaning. It will be remembered that Grice claims that for an agent to have meant something by an utterance it is necessary that (among other things) the agent intend (a) that his audience produce a particular response, and (b) that his utterance be a reason, and not a cause, of the audience's response. Grice believes that reasons for actions are different from causes of events, and that the difference lies in the fact that while a person can both have a reason for his action and have control over his action, a person cannot both be caused to do an action and have control over it (Grice, 1967, p. 46).

In drawing such a distinction between reasons and

causes, Grice is making explicit his assumption that there is a difference between what might be called rational action and causal phenomena. Grice is claiming that agents meaning things by their utterances is a form of rational behavior which cannot be explained by "merely" causal explanations-- i.e., explanations not couched in terms of agents' intentions. Grice's work on meaning is an explicit attempt to show that intensional concepts such as that of intending are "embedded in the very foundations of the theory of language" (Grice, 1968, pp. 241-242); and insofar as Grice denies that such intensional concepts are embedded in other theories (say, of the non-behavioral sciences), his theory of meaning conflicts with Quine's. Grice's theory conflicts with Quine's because Grice believes that rational behavior, of which meaningful speech is one form, is fundamentally different from other kinds of natural phenomena. The difference lies in the alleged fact that reference to agents' intentions is necessary in any explanation of rational action, while no such reference is required in causal explanations of natural phenomena. As we have seen, Quine denies that meaningful speech is different in kind from explanation of other natural phenomena; and, in particular, he denies that reference to agents' intentions is necessary in explanation of meaningful speech.

The way in which Quine's theory of meaning conflicts with Grice's should serve to indicate the way in which Quine's theory also conflicts with Searle's. Searle accepts

a division of the world into two parts, which is analogous to Grice's division just discussed; in Searle's terms, all phenomena are either brute or institutional facts, or, as he also says, non-rule governed or rule governed phenomena.

Meaningful speech is alleged to be a form of rule governed behavior, and, as such, is not susceptible to the same kind of explanation as are non-rule governed phenomena.

Explanation of rule governed behavior, and of meaningful speech in particular, must be in terms of rules underlying the behavior; such explanations are different from and irreducible to explanations--given in terms of regularities rather than rules--of non-rule governed behavior. Again, then, the conflict between Quine's and Searle's theories is that between a naturalistic monism and an ontological and methodological dualism.

Finally, although Grice's and Searle's theories of meaning agree on several major points, there are still several ways in which Grice's theory conflicts with Searle's. These conflicts were introduced in Chapter 4, and will be reviewed here. The first conflict between the theories is that while Grice thinks that people are the primary vehicles of meaning, and that sentence- and word-meaning must be explicated in terms of utterer's meaning, Searle thinks that illocutionary acts (and, therefore, ultimately, sentences) are the vehicles of meaning. As against Grice, Searle claims that what people mean by their utterances must be explained in terms of what their utterances mean.

The second conflict between Grice's and Searle's theories of meaning concerns the relation that holds between speakers' utterances and speakers' intended effects. Searle claims that the relation is "essentially conventional," and not, presumably, natural or causal. Grice agrees that the relation between utterance and intended effect cannot be natural or causal, but claims that not all utterances are conventionally related to intended effects. In Searle's terms, the difference between Grice and Searle is that the latter believes, while the former denies, that the relation between utterance and intended effect is "essentially conventional." As was noted in 4.3, Grice agrees with Searle that linguistic utterances are conventionally related to intended effects; the difference remaining between Grice and Searle can be traced to Grice's extended use of the term "utterance," to cover nonlinguistic as well as linguistic utterances. Nevertheless, if one admits the distinction between accidental and essential characteristics, as, apparently, does Searle, then the issue still remains whether the relation between utterance and intended effect is "essentially conventional."

Finally, Searle has noted a further conflict between his theory and Grice's which deserves mention. The conflict arises over Searle's characterization of the intended effects of utterances as illocutionary effects. Searle claims, first, that there is an important difference between illocutionary and perlocutionary effects, and, second, that according to Grice's theory, as against his own, the intended effects of

utterances are perlocutionary effects. I am not convinced that the conflict Searle sees over the characterization of the intended effects is a genuine conflict, for two reasons. First, as explained in Chapter 4, I am not convinced that the distinction between illocutionary and perlocutionary effects is a genuine distinction; and second, Grice gives no evidence that he has, as Searle alleges, characterized the intended effects of meaningful utterances as perlocutionary rather than illocutionary effects. Nevertheless, for those who (a) admit a distinction between illocutionary and perlocutionary effects, and (b) believe that Grice is committed to the position that the intended effects of meaningful utterances are perlocutionary effects, there exists a conflict between Searle's and Grice's theory concerning the correct characterization of the intended effects of meaningful utterances.

It follows from the above discussion that (a) any claim, such as that made in Harman's reconciliation thesis, that Quine's, Grice's, and Searle's theories of meaning are not theories about the same thing is false, and (b) Quine's, Grice's, and Searle's theories are genuinely competing theories of meaning. It remains, in this chapter, to discuss the criteria according to which competing theories of meaning may be evaluated.

5.3 Criteria for Evaluating Competing Theories of Meaning

The criteria to be applied, in Chapter 6, to a

comparative evaluation of (an extension of) Quine's, Grice's, and Searle's theories of meaning are those proposed by Shaffner in his paper "Outlines of a Logic of Comparative Theory Evaluation with Special Attention to Pre- and Post-Relativistic Electrodynamics" (1970). It should be obvious from the first two sections of this chapter that the assumption is made here that philosophical or scientific theories can be competing theories, and that theories which are in competition with one another can be evaluated on neutral grounds. In the light of recent challenges to such standard assumptions about the nature of scientific theories from what has become known as the historical school of philosophy of science, it seems prudent to briefly defend my assumptions, before discussing Shaffner's logic of comparative theory evaluation.

Philosophers of science such as Feyerabend and Kuhn have challenged the standard "logical" view of science with evidence concerning the nature of science taken from case studies of episodes in the history of science. According to such philosophers of science, what the case studies show is, generally, that the "cumulative view" of science, according to which science progresses, subject to the controls of observation and experiment, toward an ever more accurate objective picture of reality, is radically mistaken. It is denied that scientific theories are subject to the control of observation and experiment, because (a) there is no observation language common to a variety of scientific theories, and (b) there are no experiments (and a fortiori

no "crucial experiments") which are theory-neutral, and which might provide a common test for two or more theories. It is also claimed that recalcitrant observations or experiments do not necessarily bring about the rejection or even modification of theories (Shaffner, 1970, pp. 311-313).

As a prolegomenon to his logic of comparative theory evaluation, Shaffner has offered an analysis of the nature of scientific theories which, he thinks, both accommodates some of the claims made by philosophers of the historical school of philosophy of science, while admitting the possibility of genuinely competing scientific theories and of comparative evaluation of such theories. For present purposes, I shall adopt Shaffner's analysis of the nature of scientific theories--not because I think that Shaffner's analysis itself is ultimately adequate, but because (a) some analysis which admits the possibility of comparative theory evaluation seems most defensible, and (b) Shaffner's analysis satisfies this condition while incorporating elements of the historians' views of the nature of scientific theories. A brief and uncritical summary of Shaffner's analysis of the nature of scientific theories follows.

Shaffner begins by defining "scientific theory" as a set of sentences of universal form. This set is constituted by a subset of analytic sentences and a subset of synthetic sentences.¹ The analytic sentences are what Shaffner calls

¹No hard-and-fast distinction is made between analytic and synthetic sentences; Shaffner's distinction is both relativistic and gradualistic, and seems consistent with what

"meaning-creating sentences," and they create the meanings of entity-terms, such as "gas molecule," by "providing an appropriate analogical description, usually drawn from several diverse domains including branches of mathematics" (Shaffner, 1970, p. 314). Such entity-terms are examples of primitive nonlogical terms which have what Shaffner calls "antecedent theoretical meaning" (1970, p. 314).

In addition, Shaffner defines the "O-sentences" of a theory as "those sentences which describe an intersubjectively testable experience," or "name an entity that is observationally accessible without the immediate use of artificial instruments" (1970, pp. 315-316). Further, "C-sentences" are defined as sentences associated with the theory in question which "state how the entities and/or processes described by the theory's axioms ultimately affect our sense organs," or, in short, "connect theoretical terms with 'observational' terms (1970, pp. 315-316).

So far, it might be thought that Shaffner's analysis accords with the standard logical analysis of the nature of scientific theories. There are differences, however. First, Shaffner goes on to say that the differences between his C-sentences and what are labeled by the standard view "correspondence rules" is that C-sentences "are further analyzable into chains of sentences, each element of which characterizes a state of affairs, and which causally or nomically implies

critics such as Putnam have had to say about the analytic/synthetic distinction. See Putnam, 1965.

the next element in the chain" (1970, p. 316). For example,

in the C-sentence of the Bohr atomic theory connecting the "theoretical notion" of an "electron transition between orbits" with the "observational" notion of a "spectral line," the theory of physical optics is appealed to in order to account for the behavior of light in prisms and telescopes which makes the "spectral line" associable with the "electron transition" (1970, p. 316).

The causal or nomological implications occurring in the C-sentence of the Bohr atomic theory connecting the notion of an electron transition between orbits and the notion of a spectral line are warranted by the relatively well-corroborated auxiliary theory of physical optics. That Shaffner believes the theory of physical optics must be appealed to in support of a C-sentence of Bohr atomic theory reveals his commitment to a Duhemian or Quinian view of the interdependence of scientific theories (1970, p. 316).

Second, Shaffner's analysis of O-sentences differs from the standard analysis of observation sentences, insofar as O-sentences contain both a theory-neutral and a theoretical component. Shaffner says that an O-sentence may contain "(1) terms which are 'ostensively definable' . . . , (2) logical terms . . . , and (3) mathematical terms which are theory neutral" (1970, p. 323). The terms belonging to the first set are called "O-terms," and it is claimed that there are two types of senses ascribable to O-terms. "The first or primary sense is that associated with the referent [class]," and "the second sense is that which is associated with the O-term by virtue of the term's being embedded or hooked into a theory through . . . C-sentences" (1970, p. 323). The

primary sense of O-terms is theory neutral, and provides the common ground for evaluating the experimental adequacy of competing theories; the secondary sense of O-terms is not theory neutral. Although Shaffner's analysis of the kinds of O-terms occurring in O-sentences seems open to question, his conclusion, that O-sentences contain both a theory neutral and a theoretical component, seems plausible. For the sake of this thesis, I shall accept Shaffner's conclusion about O-sentences, without questioning his analysis of O-terms.

On such an understanding of the nature of scientific theories, it is possible for theories to be competing theories, and it is possible to evaluate such theories on common ground. Shaffner goes on to describe his three criteria on which competing theories may be evaluated, and it is this aspect of Shaffner's work which is crucial to this thesis.

Shaffner's logic of comparative theory evaluation is based on a generalization of the categories of comparative theory evaluation proposed by the physicist Hertz. The origin of Shaffner's logic, in the work of Hertz, is of more than historical interest, as it reveals Shaffner's aim of constructing a logic of comparative theory evaluation that is faithful to the actual practices of working scientists. As we shall see, the main justification of Shaffner's logic lies in its fidelity to scientific practice.

The first dimension within which competing theories may be evaluated is called "theoretical context sufficiency."

Roughly, theoretical context sufficiency is "concordance between a theory to be assessed and the corpus of accepted scientific knowledge" at a date (Shaffner, 1970, p. 321). Because the corpus of accepted scientific knowledge changes over time, judgments of theoretical context sufficiency are relative to the time at which they are made. Not only are judgments of theoretical context sufficiency temporally relative, but Shaffner also acknowledges that they may also vary from assessor to assessor--because of the difference in knowledge and perspective which exists in assessors.

The second dimension within which competing theories may be evaluated is what Shaffner calls "experimental adequacy." The category of experimental adequacy is most readily employed in cases in which there is overlap in experimental results and in observations judged relevant to each of the theories being compared. In such cases, the theory which agrees with the most experimental outcomes and observations is the most adequate (Shaffner, 1970, p. 322).

The third category for comparison of theories is simplicity. Shaffner distinguishes between three kinds of simplicity; the first kind is called "fitness." Roughly, the more fit a theory is, the fewer distinctions must it make which have no empirical consequences. For example, Lorentz's theory of the electro-dynamics of moving bodies is less fit than is Einstein's special theory of relativity, because Lorentz's theory "makes a distinction between the case of a magnet stationary in the ether and a circuit moving near it,

and the case of a circuit stationary in the ether with the magnet moving through the ether," while Einstein's theory rejects this distinction. Lorentz's theory is less fit than Einstein's because the theoretically distinct states of affairs just described have the same observable consequences (Shaffner, 1970, p. 344).

The second kind of simplicity described by Shaffner is called "terminological and/or ontological" simplicity. With certain restrictions,² Shaffner accepts the principle that the fewer kinds of entities posited by a theory, the simpler is the theory. Because Shaffner is a theoretical realist, the fewer entity-terms employed in a theory, the fewer entities will be posited by the theory, and the simpler will be the theory.

Finally, Shaffner defines a notion of "simplicity of system." The notion of simplicity of system assumes that the theories being compared have been axiomatized, and that "each axiom represents a further irreducible physical effect" (1970, p. 328). The simpler the axiom system of a theory, the simpler will be the theory within the dimension of simplicity of system.

Shaffner notes that "by making a virtue of simplicity we make a vice of complexity" (1970, pp. 328-329). In particular, theories which include ad hoc hypotheses are to be rated low on simplicity. Shaffner defines two kinds of ad hocness of hypotheses. First, "intrasystemic ad hocness"

²For which, see Shaffner, 1970, pp. 359 and 371.

is defined as the ad hocness attaching to "a hypothesis which is added to a theory solely to enable it to outflank some unexpected and embarrassing result, an E-result, and which hypothesis (plus the theory under test) has no further additional testable consequences which differ from the E-result in an interesting and significant way" (Shaffner, 1970, p. 329). "Intersystemic ad hocness" is the kind of ad hocness attaching to a hypothesis H if H "is conjoined to a theory T_1 , to obviate the necessity of accepting a new theory T_2 , which accounts for the E-result without H" (1970, p. 329).

Several explanatory notes must be added to this sketch of Shaffner's categories within which competing theories may be evaluated. First, we have noted that within at least one of the categories (theoretical context sufficiency) judgments about the comparative adequacy of theories are relative to both the times of the judgments and the people making the judgments. Shaffner acknowledges, more generally, that his "logic of comparative theory evaluation" does not employ what he calls "formal" or "effective" concepts (1970, p. 330). That is, in his words, his logic "does not constitute a type of easily applicable schema which can result in an automatic decision for the person thinking in terms of its categories" (1970, p. 330). This is to say that two rational men employing Shaffner's logic of comparative theory evaluation may give contradictory judgments of the comparative adequacy of two theories.

The second observation which needs to be made

concerning Shaffner's logic is that even though the logic does not provide an effective decision procedure for ranking scientific theories along some scale of adequacy, it does, as Shaffner claims, capture those categories actually considered by scientists engaged in "making the history of science" (1970, p. 330). The chief justification Shaffner gives for his logic is its fidelity to the practice of working scientists; this fidelity is also the chief recommendation for adoption of the logic for our present purposes.

The third observation which needs to be made concerning Shaffner's logic is that the three criteria proposed by Shaffner for evaluating theories are intended to be applied in a combinatorial way (Shaffner, 1970, p. 371). That is, each of the three criteria is given some particular weight in the total decision made by the assessor. For theory T_1 to be more adequate than theory T_2 , it is not necessary that T_1 rank above T_2 in each of the three categories; it is sufficient that T_1 rank above T_2 on a total score, arrived at by combining judgments made in each of the categories.

As to the question of what weight to give each of the three categories, Shaffner gives no final answer. He does not suggest what factors might enter into determination of the weighting of the three categories, even though he claims that the weight given to the individual categories may vary as different sets of theories are evaluated. In general, however, it is clear that Shaffner believes that consideration

of theoretical context sufficiency and experimental adequacy "have roughly the same weight" in all evaluations of competing theories, and that considerations of relative simplicity become "an important factor when the assessment made by the application of the previous categories is contradictory or indeterminate" (1970, p. 372).

As a final observation on Shaffner's logic of comparative theory evaluation, the common complaint needs to be acknowledged, that although philosophers of science make much of the simplicity of scientific theories, no theory of theory-simplicity is available. Thus, although Shaffner includes considerations of simplicity as one of the dimensions within which theories may be evaluated, and although he does distinguish between three kinds of simplicity of theories, the notion of theory-simplicity remains notoriously unclear. In defense of Shaffner's inclusion of such an unsystematized notion of simplicity in his logic, it need only be noted that working scientists do in fact seem to rely on some intuitive notion of theory-simplicity in evaluating theories. That this notion of simplicity has not as yet been systematized is no reason for Shaffner to exclude considerations of simplicity from his logic of comparative theory evaluation, especially since Shaffner's main justification for his logic is its fidelity to the practices of working scientists. If and when a theory of theory-simplicity is developed, it may be appended to Shaffner's logic.

In sum, then, Shaffner proposes his three criteria of

adequacy as a metascientific logic of comparative theory evaluation. The logic is supposed to be neutral as between any rival scientific theories, and is supposed to be an abstraction of the way in which practicing scientists actually go about the business of resolving conflicts between competing theories. In Chapter 6, Shaffner's logic will be applied to the problem of evaluating the three competing theories of meaning described in the earlier chapters of this thesis.

5.4 A Defense of the Criteria for Evaluating Competing Theories of Meaning

Before proceeding to the application of Shaffner's logic of comparative theory evaluation, it is necessary to defend my choice of criteria for evaluating competing theories against two objections. The first objection is that Shaffner's criteria for evaluating competing theories were originally proposed as criteria for evaluating scientific theories; application of the criteria to philosophical theories of meaning requires some justification.

Application of Shaffner's logic to philosophical theories of meaning is justified on at least two grounds. First, just as Shaffner's main justification for his criteria for evaluating scientific theories is that they are in fact the criteria employed by scientists in evaluating scientific theories, it is also true that in evaluating philosophical theories philosophers rely on consideration of analogous notions of theoretical context sufficiency, experimental

adequacy, and simplicity. That philosophers do appeal to such categories sanctions application of the categories to the problem of evaluating the three competing theories of meaning.

Shaffner's explication of theoretical context sufficiency, experimental adequacy, and simplicity are easily adapted to the task of evaluating philosophical theories. Theoretical context sufficiency is explicated, in the case of philosophical theories, as concordance between the theory being evaluated and other relevant and generally accepted philosophical and scientific theories. Experimental adequacy is explicated as agreement with the data--which is, in this case, speakers' beliefs about the meanings of linguistic units. As we shall see in Chapter 6, a philosophical theory of meaning is justified within the category of experimental adequacy to the extent that what it says are the meanings of linguistic units agrees with what native speakers say are the meanings of the linguistic units. Simplicity, as applied to philosophical theories, is explicated in the same way it is in the case of scientific theories.

The second reason why application of Shaffner's logic of comparative theory evaluation to philosophical theories is justified is that there is no property possessed by philosophical theories and not by scientific theories, which renders inapplicable Shaffner's logic of theory evaluation to philosophical theories. As suggested in Chapter 1, the structure of the theories of meaning being considered

here and certain scientific theories are identical: each is an attempt to provide a functional account of phenomena germane to meaning, and each has as its goal the explanation of these phenomena. There is no reason to suppose that the structure of philosophical theories of meaning and scientific theories of some other subject matter are different. The only important difference between philosophical theories of meaning and scientific theories is the obvious difference of subject matter; therefore, application of Shaffner's logic of comparative theory evaluation to philosophical theories of meaning is no more problematic than is its application to various scientific theories.

The second objection to my adoption of Shaffner's logic of comparative theory evaluation concerns my rejection of Chomsky's suggestions concerning the criteria according to which rival grammars may be evaluated; Chomsky's suggestions, and the reasons for my rejection of them, need to be explained.

According to Chomsky, there are two levels on which a grammar can be evaluated. At the first level, that of descriptive adequacy, a "grammar is justified to the extent that it correctly describes its object, namely the linguistic intuition--the tacit competence--of the native speaker" (Chomsky, 1965, p. 27). That is, to be justified at the level of descriptive adequacy a grammar must not only be consistent with all of what Chomsky calls the "primary linguistic data"--which are simply examples of actual linguistic performances--

but it must also be consistent with all of the linguistic intuitions of an idealized native speaker of the language in question (1965, pp. 24-25). Since all grammars of a particular language which are justified at the level of descriptive adequacy are consistent with the primary linguistic data, we may say that such theories are extensionally equivalent. They are extensionally equivalent in the sense that they each generate all and only the well formed sentences of the language in question.

At the second level, that of explanatory adequacy, "a grammar is justified to the extent that it is a principled descriptively adequate system, in that the linguistic theory with which it is associated selects this grammar over others, given primary linguistic data with which all are compatible" (Chomsky, 1965, p. 27). In other words, to be justified at the level of explanatory adequacy a grammar must (1) be justified at the level of descriptive adequacy, and (2) be the grammar selected by some decision procedure (provided by the general linguistic theory with which the grammar is associated) from all those grammars that meet the condition of descriptive adequacy. Such a decision procedure must itself incorporate, evidently, further conditions of adequacy to be applied to descriptively adequate grammars.

There are two reasons why Chomsky's discussion of the levels of adequacy of proposed grammars is not helpful in our task of evaluating rival theories of meaning. First, Chomsky's levels of adequacy are not, as are Shaffner's

three categories, truly metascientific or neutral as between the theories being evaluated. Shaffner's categories are neutral, because they do not involve any commitment to the kind of facts which are relevant to the particular theories being evaluated. Chomsky's level of descriptive adequacy is not theory-neutral, because it will be rejected by any theoretician not agreeing that the linguistic intuitions of an idealized native speaker constitute the subject matter of linguistics. Since a substantial number of linguists and philosophers of language do reject the thesis that such intuitions constitute the subject matter of linguistics, it would be fruitless to insist that consistency with such intuitions be one condition of adequacy of any proposed grammar; to so insist would be to load the dice in favor of those grammars which have as their objects the intuitions of idealized native speakers of the language in question.

The second reason why Chomsky's levels of adequacy are unhelpful to the task of evaluating rival theories of meaning is that, by themselves, they do not provide what is needed, namely, a way of evaluating rival theories. As we saw, Chomsky's level of explanatory adequacy does not provide a way of evaluating rival grammars; on the contrary, what it does is presuppose that some way of evaluating rival grammars is already available. For, to be justified at the level of explanatory adequacy, a grammar must, among other things, be the grammar--selected from among competing grammars, by the linguistic theory with which the selected grammar and all

the competing grammars are associated--which is the most adequate on some scale. The crucial point is that Chomsky's level of explanatory adequacy does not itself provide the procedure for selecting explanatorily adequate grammars; it merely assumes that such a procedure is available and has been applied. As Quine has complained, Chomsky has offered no methodology for selecting one "most adequate" grammar from a set of extensionally equivalent grammars (i.e., from all those grammars which satisfy the condition of descriptive adequacy) (Quine, 1972, p. 444). True, Chomsky has made some vague references to what he calls the "weighting function" which selects one grammar over other extensionally equivalent grammars. Chomsky says that his "weighting function" must be discovered "on empirical grounds, by considering the actual relations between input data and acquired grammars" (1965, p. 85). But until Chomsky further specifies the properties of this "weighting function," the point remains that he has proposed no methodology for selecting a grammar from among extensionally equivalent grammars.

Chomsky's notion of "weighting function" is only made all the more mysterious when he says that the general notion of theory-simplicity "in a general context of epistemology and philosophy of science" is "entirely irrelevant" to his notion of "weighting function" (1965, p. 85). Thus, Chomsky says that

when I speak of "simplicity of grammar," I am referring to a "weighting function," empirically determined, that selects a grammar of the form permitted by the universal schematism over others that

are also of the proper form and are compatible with the empirical data. I am not using the term "simplicity" to refer to that poorly understood property of theories that leads the scientist to select one rather than another (1965, p. 85).

The problem is, however, that Chomsky does not say exactly how his "weighting function" differs from the scientists' notion of simplicity; all he suggests on this matter is that while his "weighting function" is determined empirically, the scientists' notion is determined by "a priori insights" (1965, p. 85). However, that, e.g., Shaffner's notion of theory-simplicity is in fact determined by "a priori insights," rather than from historical case studies, is far from clear. And, furthermore, until Chomsky provides a specification of his "weighting function," which shows it to be significantly different from, e.g., Shaffner's notion of simplicity, there seems to be no reason not to rely on Shaffner's logic of comparative theory evaluation in our task of evaluating Grice's, Searle's, and (an extension of) Quine's theories of meaning.

CHAPTER 6

TOWARD A NATURALISTIC THEORY OF MEANING

6.1 The Naturalistic Theory of Meaning: Stimuli vs. Stimulations

Two tasks remain to be accomplished in this thesis. The first is to propose three revisions of Quine's naturalistic theory of meaning, to which task the first three sections of this chapter are devoted. The second task remaining is the application of Shaffner's logic of comparative theory evaluation to the problem of evaluating Grice's, Searle's, and the proposed naturalistic theory of meaning. It will be argued in the fourth section of this chapter that according to the criteria proposed by Shaffner, the naturalistic theory developed here is more adequate than either Grice's or Searle's theories of meaning.

The naturalistic theory to be defended here is, basically, Quine's theory of meaning, with three important differences. To begin with, Quine's definition of "stimulus meaning" is adopted, with a revised interpretation of the term "stimulation." Quine's definition of "stimulus meaning" is as follows:

a stimulation σ belongs to the affirmative stimulus meaning of a sentence S for a given speaker if and only if there is a stimulation σ' such that if the speaker were given σ' , then were asked S, then were given σ , and then were asked S again, he would dissent the first time and assent the second. We may define the negative stimulus meaning similarly with "assent" and "dissent" interchanged, and then define the stimulus meaning as the ordered pair of the two (Quine, 1960, pp. 32-33).

Quine stipulates that the stimulus situations which constitute the stimulus meanings of sentences be interpreted as what I shall call "stimulations" rather than "stimuli." Stimulations are firings of nerve receptors (e.g., ocular irradiations); stimuli are sensed spatio-temporal objects. It should be noted that both stimulations and stimuli are to be contrasted with neurophysiological events occurring "deep inside the body." Thus, Quine stipulates that, with reference to the problem of determining the stimulus meanings of sentences for speakers, "it is the stimulation at the bodily surface that counts, and not just the objective existence of objects of reference off in the distance, nor yet the events deep inside the body" (1969c, p. 158).

It was pointed out in Chapter 2 that Quine's interpretation of the stimulatory situations constituting stimulus meanings as stimulations rather than stimuli is inconsistent with other aspects of Quine's semantic theory--aspects which are important to retain. In what follows, three ways in which Quine's interpretation of "stimulatory situations" conflicts with important aspects of his semantic theory are reviewed, and a simple remedy is proposed, which, when adopted, renders consistent those parts of Quine's semantic

theory with which I am concerned.

First, it was shown in Chapter 2 that Quine's construal of stimulus situations as stimulations is inconsistent with his claim that it is possible for two or more speakers to speak the same language. It will be remembered that Quine defines the language of a speaker as the speaker's dispositions to respond verbally to stimulations within a certain modulus. It has been established that two speakers (or one speaker in two different moduli) are speaking the same Quinian language if and only if during particular moduli they have the same dispositions to verbal behavior to respond to the same stimulations.

Of course, Quine takes it for granted that according to his theory of meaning two or more speakers may speak the same language; this assumption is made explicit in, among other places, Word and Object, when Quine says, "we are concerned here with language as the complex of present dispositions to verbal behavior, in which speakers of the same language have perforce come to resemble one another" (1960, p. 27). A theory according to which no two speakers can speak the same language would hardly merit serious consideration; and yet, I shall argue that Quine's theory, unrevised, is in fact such a theory.

The difficulty with Quine's theory is located in its assertion of the following two theses:

- (T1) One language is being spoken by two people if and only if during particular moduli they have the same dispositions to verbal behavior to respond to the same stimulations.

(T2) Since the stimulus situations of uttered sentences are identified with firings of nerve ends, the notion of "same stimulation" presupposes, in the case of two speakers, complete intersubjective homology of nerve ends.

As Quine himself recognizes, the notion of homologous nerve ends is highly suspect. We have seen that Quine admits that "if we construe stimulation patterns my way, we cannot equate them without supposing homology of receptors; and this is absurd, not only because full homology is implausible, but because it surely ought not to matter" (1969c, p. 158). Because of the implausibility of the assumption of intersubjective homology of nerve receptors, it is clear that one would never be warranted in claiming that two persons have received the same stimulations. Because, according to Quine, languages are identical if and only if speakers have the same verbal dispositions to respond to the same stimulations, and because it is implausible to say that two speakers have received, in Quine's sense, the same stimulations, it follows that it is also implausible to say that two people can speak the same Quinian language.

Secondly, it has been shown that Quine's construal of stimulus situations as stimulations is inconsistent with the thesis that it is possible for one speaker to speak the same language twice (i.e., in two different moduli). Again, the thesis that it is possible for one speaker to speak the same language in two different moduli is more of an assumption of Quine's theory than an explicit claim; and yet, Quine's construal of stimulus situations as stimulations renders it

highly unlikely that one speaker can ever speak the same language twice. As we have seen, for one person to speak the same language twice, he must, among other things, receive the same set of stimulations twice; but, since stimulations are firings of nerve endings, it is highly unlikely that precisely the same set of nerve endings will be stimulated twice. And, if it is highly unlikely that one person will receive the same stimulation twice, then it is also unlikely that one person will speak the same Quinian language twice.

Finally, it was shown in Chapter 2 that Quine's interpretation of stimulus situations as stimulations is inconsistent with his claim that the stimulus meanings of sentences can be the same. There were two cases considered in this argument: first, it was shown that Quine's construal of stimulus situations as stimulations is inconsistent with his claim that in some cases one sentence can have the same (or a similar) stimulus meaning for two or more speakers; and, second, it was shown that Quine's construal of stimulus situations as stimulations is inconsistent with his claim that in some cases two (or more) sentences can have the same (or a similar) stimulus meaning for one or more speakers.

Case i) As we saw in Chapter 2, Quine claims that there are cases in which we can say, with justification, that one sentence has a similar stimulus meaning for two or more speakers. In fact, Quine's definition of "observation sentence" relies on just such a claim about intersubjective similarity of stimulus meanings. Since, as Quine says, "the

observation sentence is the cornerstone of semantics" (1969b, p. 89), it is of particular importance to Quine's semantic theory that stimulus meanings can be intersubjectively similar. I shall argue, however, that Quine's claim that there are cases of intersubjective similarity of stimulus meanings is inconsistent with his construal of stimulus situations as stimulations. If my argument is correct, it will follow that Quine's definition of "observation sentence" is incompatible with his construal of stimulus situations.

It will be remembered that according to Quine the distinguishing mark of observation sentences is constancy of stimulus meanings over a large number of speakers of one language (assuming, for the moment, that there can be a large number of speakers of one Quinian language). Obviously, this mark for distinguishing between observation and non-observation sentences presupposes the principle that stimulus meanings are intersubjectively comparable--as does any assumption that one sentence can have the same (or a similar) stimulus meaning for two or more speakers. Because stimulus meanings consist of sets of stimulations, Quine's "behavioral definition" of observation sentences presupposes the possibility of intersubjective comparison of stimulations. But, as we have seen, intersubjective comparison of stimulations presupposes full homology of nerve ends, which Quine admits is an "implausible myth" (1969a, pp. 158-159). Therefore, Quine's definition of observation sentence is no less mythical than is the presupposition of full homology

of nerve ends. Specifically, the implausibility of the presupposition of full homology of nerve ends renders implausible Quine's claim that one sentence can have the same (or similar) stimulus meaning for two or more speakers.

Case ii) Quine also claims that there are cases in which we can say, with justification, that two sentences have the same (or a similar) stimulus meaning for one or more speakers. Such cases are cases of what Quine calls stimulus synonymy.¹ Two sentences may be, according to Quine, either intrasubjectively stimulus synonymous or intersubjectively stimulus synonymous. I shall argue, however, that the notions of both intrasubjective stimulus synonymy and intersubjective stimulus synonymy are inconsistent with Quine's construal of stimulus situations as stimulations.

Intrasubjective stimulus synonymy is inconsistent with Quine's construal of stimulus situations because of the implausibility of one person ever receiving the same Quinian stimulation twice. Because stimulations are firings of nerve ends, it is highly unlikely that one person will ever receive the same stimulation twice (although, I admit, one person may receive two very similar stimulations). And because stimulus meanings are sets of stimulations, it follows that it is highly unlikely that two sentences will ever have the same stimulus meaning for one person (although they might have similar stimulus meanings).

Furthermore, intersubjective stimulus synonymy is

¹For discussion, *v* Quine, 1960, pp. 62ff.

also inconsistent with Quine's construal of stimulus situations, because of the by now familiar implausibility of the presupposition of full intersubjective homology of nerve ends. The notion of intersubjective stimulus synonymy (and even the weaker notion of stimulus meaning similarity) depends on the presupposition of intersubjective homology of nerve ends; and, the implausibility of this presupposition carries over to the subordinate notion of intersubjective stimulus synonymy. So long as Quine retains his identification of stimulus situations with stimulations, it will be implausible to suppose that any two sentences are intersubjectively stimulus synonymous.

The obvious way to resolve the above inconsistencies in Quine's theory of meaning is to construe the stimulus situations of uttered sentences as stimuli rather than stimulations. To interpret stimulus situations in this way is the first of two revisions of Quine's theory to be proposed here. It seems perverse that Quine has not adopted this measure himself, realizing, as he does, at least some of the consequences of his interpretation of stimulus situations.

It might be supposed that Quine has overwhelming reasons for rejecting a construal of stimulus situations as stimuli, and for maintaining his construal of stimulus situations as stimulations. However, this does not seem to be the case. Quine gives two reasons for rejecting an identification of stimulus situations with stimuli, neither of which are persuasive. The first reason is taken from

Quine's theory of linguistic acquisition. Quine claims that "even a primitive mother, in encouraging or discouraging a child's use of a word on a given occasion, will consider whether the relevant object is visible from where the child sits" (1969c, p. 158). Quine's second reason for rejecting an identification of stimulus situations with stimuli rather than stimulations is that "stimulation can remain the same even though the [stimulus] be supplanted by a counterfeit," and that "conversely, stimulation can vary in its power to prompt assent to [e.g.] 'Gavagai' because of variations in angle, lighting, and color contrast, though the rabbit remain the same" (1960, p. 31).

In reply, we may note, first, that Quine's observation that a language teacher will consider whether the relevant stimulus is visible to the language learner is compatible with the proposed construal of stimulus situations as stimuli. Of course, if I want to teach you, by ostension, the meaning of "rabbit," I will be sure that there is a rabbit within your view; but, I cannot see that Quine's theory of linguistic acquisition (much less his theory of meaning) forces us to identify the meaning of "rabbit" with your view of the rabbit.

Second, Quine's observations concerning (a) the constancy of stimulation in the face of changing ("counterfeit") stimuli, and (b) the converse change in stimulations in the face of constant stimulus (presented in varying perceptual fields), are also compatible with the proposed

construal of stimulus situations as stimuli. Presumably, Quine wants to be able to preserve the claim that, for example, if a speaker of English assents to the query "Duck?" each time when prompted by, first, a view of a duck and, second, a view of a duck decoy, then the explanation of this error rests on sameness of the speaker's stimulation by both duck and duck decoy, rather than on sameness of stimulus-- which would obviously be incorrect. It would seem, however, that we may say that what prompts the speaker's assent in this case is the seen duck and duck decoy, rather than his seeings of the duck and duck decoy. To say that a confused speaker assents to the query "Duck?" when prompted by both a duck and a duck decoy is consistent with the observation that his confusion stems from the fact that the speaker's stimulations of the duck and duck decoy may be similar, though the duck and duck decoy are undeniably different.

An analogous argument may be constructed with respect to Quine's observation concerning changing stimulations in the face of constant stimulus. Thus, presumably Quine wants to preserve the claim that, for example, if a speaker assents to the query "Duck?" when prompted by a view of a duck in a well-lit visual field, but does not assent to "Duck?" when prompted by a view of an (imperceptible) duck in a totally dark visual field, the explanation rests on difference of stimulation rather than on a difference of stimuli; because, in both cases a duck is present, but in the first it is visible while in the second it is not.

Again, however, we may say that the stimulus meaning of "Duck" for our speaker is constituted by, to put it simply, the class of all ducks viewed by him, instead of the class of all his views of ducks. Accordingly, we may explain a speaker's dissent from the query "Duck?" when prompted by a dark field containing a duck, with the observation that the duck was invisible to the speaker. We need not fly to the conclusion that the stimulus meaning of "Duck" consists of stimulations of ducks; that it consists of the (viewed) ducks themselves is compatible with the observation that stimulations change as constant stimuli are presented in varying perceptual fields.

Our solution to Quine's problems is, then, to interpret the stimulus situations of uttered sentences as stimuli rather than stimulations. By adopting this measure, we avoid the absurd consequence of basing judgments of stimulus situation identity on a presupposition of complete intersubjective homology of nerve ends. This presupposition of intersubjective homology of nerve ends has been shown to be the locus of several of the inconsistencies noted in Quine's semantic theory; by ruling the presupposition to be irrelevant to our concerns, these inconsistent parts of Quine's theory are rendered consistent.

One problem remains, however. Not all of the inconsistencies in Quine's theory were traceable to the suspect notion of intersubjective homology of nerve ends. In particular, it may seem as if the conclusions that (a) one

speaker cannot speak the same language twice, and (b) two sentences cannot have the same stimulus meaning for one speaker, are both preserved in a semantic theory according to which stimulus situations are construed as stimuli. The case for this supposition may be made as follows.

In identifying the stimulus situations of utterances with stimuli we are following closely claims concerning meaning which were made by Bloomfield in his Language, as noted in Chapter 2. But, as Bloomfield recognized, "the situations which prompt us to utter any one linguistic form, are quite varied; philosophers tell us, in fact, that no two situations are ever alike" (1935, p. 140). Because of this fact, and because of the fact that we have identified stimulus meanings with the situations which prompt speech utterances, it follows that no two utterances (including no two utterances of the same linguistic form by one speaker) ever have the same stimulus meaning. And, because of our definition of language, as present disposition to respond verbally to present stimuli, it follows that one speaker cannot speak the same language twice.

Before replying to this objection to interpreting stimulus situations as stimuli, it will be helpful to review Bloomfield's original reply. Bloomfield's reply was, not to give up either (a) the identification of stimulatory situations of utterances with the meanings of the utterances, or (b) the interpretation of stimulatory situations as stimuli, but to distinguish between the semantically relevant (the

"distinctive") and the semantically irrelevant (the "non-distinctive") features of situations (Bloomfield, 1935, p. 141). According to Bloomfield, not all features of a situation are relevant to the meaning of the utterance which is prompted by that situation; only the semantic features are relevant. And, even though no two situations are identical, Bloomfield claims that the semantic features of different situations may be identical. Therefore, in cases in which two utterances have been prompted by situations whose semantic features are identical, the utterances may be said to have the same meaning.

Bloomfield's reply faces the obvious difficulty of having to distinguish between semantically relevant and semantically irrelevant features of situations. This is a difficulty, because the distinction cannot be made on grounds referring to the meanings of the utterances in question, because the meanings are, ex hypothesi, unknown. Presumably, if the meanings were known, it would be easy to distinguish between semantically relevant and semantically irrelevant features of situations--but the meanings are precisely what we are trying to determine.

At this point, an advantage of Quine's theory of meaning over Bloomfield's should be apparent. The advantage is that Quine is not forced to make the impossible distinction between semantically irrelevant features of stimulus situations, simply because he identifies the (affirmative) stimulus meaning of sentence S as the totality of all the

features of the class of stimulations that prompt assent to S when the speaker is queried. Quine rejects any distinction between semantically relevant and semantically irrelevant features of stimulations at the price of admitting that no two utterances ever have precisely the same meaning (because no two stimulations are precisely identical). If it were not for Quine's interpretation of stimulus situations as stimulations, this price would not be too high, however; for, we could talk of resemblance of stimulus situations instead of identities, and degrees of synonymities instead of precise synonymities.²

My own resolution of the problem at hand, then, is to take what I think is true from both Quine's and Bloomfield's theories. From Quine, I adopt the definition of stimulus meaning, given earlier. From Bloomfield, I adopt the interpretation of stimulus situations as stimuli rather than stimulations. From Quine, I reject any distinction between semantically relevant and semantically irrelevant features of stimulus situations, and treat the (affirmative) stimulus meaning of sentence S for speaker A as the class of all the stimuli that would prompt A's assent to S when queried, including all features of the stimuli within the stimulus meaning of S. This definition of stimulus meaning yields a graded notion of stimulus synonymy, because no two stimulus situations are precisely identical. Nevertheless, two stimulus situations may approximate one another to

²V Quine, 1969c, p. 159.

varying degrees, and insofar as the classes of stimuli that prompt assent to sentences S_1 and S_2 resemble one another for a given speaker, the stimulus meanings of S_1 and S_2 will also resemble one another for that speaker.

Quine considers stimulations of spatio-temporal objects and events to be constitutive of his stimulus meanings; I take the objects and processes themselves to constitute stimulus meanings. That is, instead of treating the stimulatory situations of utterances as perceptions (views, hearings, etc.) of objects and events, I treat the stimulatory situations as the perceived objects and events themselves. Thus, not necessarily all of the objects and events in a speaker's field of view get counted as part of the stimulus situation of his utterance; only the objects and events actually perceived by him get counted. There should be no difficulty in determining what objects and events are perceived by speakers, beyond any difficulties inherent in Quine's program of determining the perceptions (i.e., "stimulations") of the speakers. That is, if stimulations are intersubjectively checkable (by which I take it, Quine means that they are in principle publicly assessable), as Quine says they are (1960, p. 31), then, assuming what might be called a "realist" theory of perception, so should we be able to determine what objects are perceived (i.e., what objects give rise to the stimulations) by the speaker. And, the class of objects and events actually perceived by speaker A, which would prompt his assent to sentence S when queried,

constitutes the affirmative stimulus meaning of S for A.

6.2 The Naturalistic Theory of Meaning: Speakers' Internal States

The second important revision of Quine's theory to be proposed is an enrichment of his notion of stimulus meaning to include speakers' internal states (in the sense defined in Chapter 2) as well as stimulations of nerve receptors. The need for such an enrichment was argued in Chapter 2, and little needs to be added here.

Inclusion of speakers' internal states in the notion of stimulus meaning represents a recognition of the fact that internal states as well as sensory stimulations are determinants of verbal behavior. Internal states and sensory stimulations are each partial causes of verbal behavior, and, in particular, are partial causes of speakers assenting or dissenting to sentences when queried. Quine's notion of stimulus meaning is inadequate without reference to speakers' internal states. Once speakers' internal states are included in the notion of stimulus meaning stimulus meanings may be said to have two parts: the "inner," constituted by the speakers' internal states, which are partial determinants of verbal behavior, and the "outer," constituted by the stimuli proper (i.e., sensed spatio-temporal objects), which are also partial determinants of verbal behavior.

It was argued in Chapter 2 that speakers' internal states are naturalistic objects or processes--that is, that they are subject to naturalistic explanations. One

consequence of this fact is that internal states are not mental states; a second consequence is that knowledge of the internal states is not private. A corollary of this second point (which is of some importance to the arguments of the preceding section) is that although speakers' internal states may be idiosyncratic, they are still intersubjectively knowable to be so. There is no reason to suppose that a speaker's internal state, though idiosyncratic, may not be accessible to others, and may not be known by others to be idiosyncratic.

The differences that may obtain between (a) the internal states of different speakers, and (b) the interactions of the internal states with the sensory stimulations of different speakers enormously complicate the task of intersubjective comparison of stimulus meanings, but they do not make it impossible. Of course, in practice speakers usually get along without making precise intersubjective comparisons of the meanings of sentences; but such comparisons could be carried out if necessary. Internal states of speakers of the sort being included in the notion of stimulus meaning are neurophysiological states, with genetic, chemical, anatomic, and other properties of the speakers constituting the states. What differences there are between the internal states of speakers are differences in purely physical properties of the speakers. Therefore, intersubjective comparison of the internal states of speakers is no different, in principle, from comparison of the "outer" parts of stimulus

meanings. The internal states of two speakers are similar if the states are similar in ways discoverable by the natural sciences: e.g., appropriate genetic, chemical, and anatomic similarities in speakers give rise to corresponding similarities in the speakers' internal states.

The situation with respect to the "internal" part of the stimulus meanings of sentences for speakers is parallel to that of the "outer" part, vis à vis the question of intra- and intersubjective similarity of stimulus meanings. The "outer" parts of the stimulus meanings of one or more sentences are rarely identical, because the "outer" parts of the stimulus prompting assent to the sentences are rarely identical; the "outer" parts of the stimuli may be more or less similar, however, and we can therefore speak of resemblances of stimulus meanings instead of identities of stimulus meanings. And, the "inner" parts of the stimulus meanings of sentences are perhaps even more rarely identical (and are never, in fact, numerically identical for two or more speakers), but, since they may also be more or less similar, we may still speak of resemblances of stimulus meanings, if not identities. There is no extra difficulty in comparing the "inner" parts of stimulus meanings that is not present in comparing the "outer" parts; therefore, including internal states in the notion of stimulus meaning does not jeopardize the comparability of stimulus meanings.

By including speakers' internal states in the notion of stimulus meaning, Quine's original notion is enriched in

important ways. Stimulus meanings are seen to contain two components, one being constituted by the speakers' internal states, and the other being constituted by the stimuli proper. Enriched in this way (and in the way to be described in the next section), the notion of stimulus meaning seems to capture all the variables of which verbal behavior is a function.

6.3 The Naturalistic Theory of Meaning: Verbal Stimuli

The third important revision of Quine's theory to be proposed is an extension of his notion of stimulus meaning to include linguistic components of stimulatory situations of utterances (and, simply, linguistic stimulatory situations) as well as what Quine calls "non-verbal" stimulatory situations. As we saw in Chapter 2, Quine draws a sharp distinction between verbal and non-verbal stimulatory situations, and denies that verbal stimulations are to be included in the notion of stimulus meaning. I shall argue, however, that in nearly every case verbal stimuli (speech qua act and object-- i.e., sentences, etc.) are relevant to the stimulus meanings of sentences.

There is one clear sense in which verbal stimuli are relevant to the stimulus meanings of all sentences for all speakers, which Quine overlooks because of an over-strict distinction between what "prompts" utterances and what "elicits" assent or dissent to queries of particular situations. What prompts an utterance, according to Quine, is

some stimulation; but a speaker's assent or dissent to a sentence when queried in some situation is elicited by the combination of the query plus the speaker's stimulation of the stimulus situation (Quine, 1960, p. 30). Affirmative stimulus meaning is defined as the class of all stimulations that prompt assent to sentences; but what prompts the (utterances of) assent are situations made up of verbal and non-verbal components--the verbal component consisting of the query. Since a knowledge of the stimulus meanings of assent and dissent is essential in the determination of the stimulus meanings of all sentences, and since the verbal components of situations are relevant to the stimulus meaning of sentences signifying assent and dissent, verbal components of situations are relevant to the stimulus meanings of all sentences.

It might be objected that the above is a rather indirect and weak sense in which verbal stimuli are parts of the stimulus meanings of all sentences. There is a much stronger sense in which verbal stimuli are parts of the stimulus meanings of most sentences, however. This sense is suggested by the observation, noted in Chapter 2, that what we mean in a particular situation is at least partly a function of what is said in that situation--including both the utterance whose meaning is in question plus the other utterances occurring in the stimulatory situation of that utterance. More precisely, I wish to argue that the stimulus meaning of all sentences not uttered in complete isolation from other utterances are functions not only of the sentences

uttered themselves and their non-verbal stimulatory situations, but also of the other sentences uttered in the stimulatory situations.

The reason why the stimulus meanings of sentences are functions, in part, of the other sentences uttered in the stimulatory situations of the sentences is that the stimulus meaning of a sentence is simply the set of stimuli which would prompt assent to and dissent from the sentence when queried. Since such sets of stimuli include verbal as well as non-verbal stimuli, as was argued in Chapter 2, and since there is no reason to exclude verbal stimuli from stimulus meanings, we may conclude that the stimulus meanings of sentences are functions, in part, of the other sentences uttered in the stimulatory situations of the sentences.

The stimulatory situations which prompt all utterances except for those that are completely isolated from other utterances have linguistic components; that is, the stimulatory situations include verbal stimuli. One way in which different stimulatory situations may differ, then, is in their linguistic components. Therefore, an analysis of the differences in linguistic components of the stimulatory situations of two utterances will be an important part of the analysis of the differences in the meanings of the two utterances. A difference in linguistic component of the stimulatory situations of two utterances is a reliable indication of a difference in meaning of the two utterances.

It is important to note that it is not being claimed

that sameness of the linguistic components of the stimulatory situations of two utterances is a reliable indication of sameness of meaning of the two utterances. Stimulatory situations of utterances may differ, even though their linguistic components are the same. But, since meaning is taken to be a matter of the stimulatory situations of utterances, whenever the linguistic components of the stimulatory situations of two utterances are different, then the meanings of the utterances will also be different. The meanings of utterances are different if the linguistic components of their stimulatory situations are different; but the meanings of utterances need not be the same if the linguistic components of their stimulatory situations are the same.

It also needs to be noticed that the relative importance of the linguistic components of the stimulatory situations of utterances differs from utterance to utterance. That is, in some cases, such as that in which a Quinian native spontaneously (i.e., in the absence of queries, etc.) reports his observation of a rabbit, the linguistic components of the stimulatory situations will be non-existent or negligible. In other cases, such as that in which a speaker utters a sentence in the course of a conversation with another speaker, the linguistic components will assume some importance. The linguistic components will be of paramount importance if, for example, the utterance in question is uttered in response to a request for a definition. Without further clarifying the notion of the relative importance of different components of

stimulatory situations, it seems safe to say that in some cases the importance of the linguistic component of a stimulatory situation will be relatively great, while in other cases it will be relatively small. To say that the linguistic component of the stimulatory situation of an utterance is relatively important is to say that as a matter of fact the stimulatory situation of the utterance contains few nonlinguistic parts; to say that the linguistic component is relatively unimportant is to say that as a matter of fact the stimulatory situation contains few linguistic parts.

In those cases in which the linguistic components of the stimulatory situations of utterances are relatively important, we may say that similarity of the linguistic component of the stimulatory situations of utterances is a reliable indication of similarity of meaning of the utterances. This is so because in these cases there are few nonlinguistic components of the situations to contribute to the meanings of the utterances. It is not the case, therefore, that in all cases similarity of the linguistic components of the stimulatory situations of utterances indicates similarity of meaning of the utterances; similarity of linguistic components only indicates similarity of meaning in those cases in which the linguistic components are relatively important.

Our position, that as the linguistic components of the stimulatory situations of two utterances differ, so will the meanings of the two utterances differ, may be contrasted with Quine's position. In Quine's terms, our thesis asserts

that as the verbal aspects of the stimuli of two utterances differ, so will the stimulus meanings of the utterances differ. It follows that in every case of utterances not uttered in complete isolation from other utterances, the verbal aspects of the stimuli of an utterance are part of the stimulus meaning of the utterance. This claim is to be contrasted with Quine's claim that the verbal aspects of the stimuli of an utterance are not part of the stimulus meaning of the utterance.

The theory of meaning which results from revising Quine's theory in the three ways outlined above is a naturalistic theory of meaning which is considerably richer than Quine's original theory. The proposed theory differs from Quine's in three ways: it interprets the stimulatory situations of utterances as stimuli rather than stimulations, it includes reference to speakers' internal states in the description of stimulus meaning, and it includes in an analysis of the stimulus meanings of utterances an analysis of the correlations that obtain between the utterances and the linguistic components of the stimulatory situations of the utterances. In most other respects--e.g., in our notions of stimulus meaning, stimulus synonymy, stimulus analyticity, and observation sentence--the proposed theory is in substantial agreement with Quine's.³

³There are issues over which my proposed naturalistic theory of meaning would, if it were fully developed, conflict with Quine's. For example, Quine claims that his thesis of the indeterminacy of radical translation is a consequence of any naturalistic theory of meaning; I am prepared to argue,

6.4 The Naturalistic Theory of Meaning Defended

It remains, then, to compare the proposed theory of meaning with that of Grice and Searle. The comparison will be carried out within the categories proposed by Shaffner in his logic of comparative theory evaluation.

1) Experimental Adequacy

The first dimension within which we will compare Grice's, Searle's, and the proposed naturalistic theories of meaning is that of experimental adequacy. Experimental adequacy is, essentially, agreement with the data, obtained by experiment and observation. In our case, the relevant data are, as indicated in Chapter 1, the preanalytic or extra-theoretical beliefs of native speakers of particular languages concerning the meanings of utterances in those languages. A theory of meaning is justified within the category of experimental adequacy to the extent that what it says are the meanings of particular utterances in particular languages agrees with what native speakers of the languages in question say are the meanings of the utterances. One theory of meaning is to be ranked as more adequate than another theory, within the category of experimental adequacy, if the first theory

however, that the thesis of the indeterminacy of translation is a consequence of neither Quine's nor the proposed theory of meaning. Presentation of this argument lies beyond the scope of this thesis, however. An argument for a similar conclusion, that the thesis of the indeterminacy of translation is not a consequence of Quine's naturalistic theory of meaning, is presented in Hockney, forthcoming.

diverges less than the second from pronouncements of native speakers concerning the meanings of utterances.

In what follows, I will not argue that any one theory is more adequate than the others in the category of experimental adequacy in precisely the sense just defined. Rather, I will argue that it is impossible to even test the experimental adequacy of Grice's and Searle's theories (insofar as these theories are distinguishable from the proposed naturalistic theory), and that therefore, even if the proposed naturalistic theory is finally shown to be inadequate within the category of experimental adequacy, it is at least more adequate than Grice's and Searle's, in the sense that it is, but theirs is not, open to tests of experimental adequacy.

Naturally, the three theories under consideration differ as to precisely what is supposed to agree with native speakers' beliefs about the meanings of utterances, if the theories are to be experimentally adequate; if the theories did not disagree on this issue, they would not be competing theories of meaning. Both Grice and Searle take meanings to be certain speakers' intentions,⁴ and claim that their theories will be experimentally adequate to the extent that these speakers' intentions are found to match the beliefs of native speakers about the meanings

⁴They take these intentions to be related in different ways to utterances, but this difference may be ignored here.

of utterances. If the theories are to be tested for experimental adequacy by determining whether the intention said by the theories to be the meaning of utterance x really is the meaning of utterance x, then the obvious question is, "How do we find out what the intention is that the theories say is the meaning of x?" Since meanings are taken to be certain speakers' intentions, what is needed is a method of investigating speakers' intentions. In Chapter 3, we saw that Grice takes speakers' intentions to be mentalistic entities, and that he admits that these entities are not directly observable. Grice says that we must take speakers' overt physical behavior as evidence of the underlying intentions; and obviously the relevant overt behavior is, when available, statements by speakers of what their intentions are. When such direct statements are not available, we must, according to Grice, use other overt behavior as a guide to the existence of intentions.⁵ The basic point is, however, that the intentions themselves are mentalistic entities which are not directly observable; according to Grice, we must infer their existence from overt behavior. I shall argue, however, that since there is no way to identify speakers' intentions save by investigation of certain

⁵There may be neurophysiological evidence for the existence of appropriate intentions also, but Grice and Searle ignore this possibility. Neurophysiological evidence will also be ignored, for the most part, in what follows: it is safe to ignore it here, because what neurophysiological evidence there is for intentions is, like speakers' overt behavior, naturalistic evidence.

naturalistic objects and processes (such as the overt behavior of speakers, and certain neurophysiological data concerning speakers' internal states), there can be no question of testing the experimental adequacy of Grice's and Searle's theories, insofar as these theories are distinguishable from the naturalistic theory.

There are two ways of investigating speakers' intentions, then; one is to accept the speakers' statements of the intentions, and the other is to use other aspects of the speakers' overt behavior as a guide to the intentions. Let us consider the first way. Obviously, if we consider only statements by a speaker as to his intentions, and accept these statements at face value, what Grice's theory says are the meanings of utterances for a speaker must be the meanings of the utterances for the speaker; the theory cannot be wrong. What must agree in this case for the theory to be experimentally adequate are statements of the speaker as to his intentions and statements of the speaker as to what he believes the meanings of his utterances to be. But since meanings are, by definition, intentions, there cannot be any disagreement between candid statements of intentions and meanings. In other words, if statements of intentions are taken to be the only evidence for intentions, and if these statements are taken at face value, Grice's theory of meaning cannot be experimentally inadequate--because there can be, by definition, no disagreement between

statement of intention and statement of meaning. But if a theory cannot, by definition, be experimentally inadequate, we may fairly characterize the theory as not being open to tests of experimental adequacy.

Of course, it will be replied that a linguist will not take statements of intention as the only evidence for the existence of the intentions--he will also consider other kinds of overt behavior. As Grice would admit, we must usually use the overt behavior of speakers to indirectly determine the timeless meaning of utterances--i.e., what is "normally" or "in the standard cases" intended by speakers' utterances.

There is a sense, then, in which Grice's and Searle's theories of meaning (insofar as they are open to tests of experimental adequacy) presuppose a method for investigating certain naturalistic objects and processes--namely, speakers' overt behavior. Meanings, or intentions, are inscrutable for Grice and Searle in the sense that they are not open to direct empirical investigation; what is scrutable is overt behavior. Therefore, Grice's and Searle's theories are only open to confirmation and disconfirmation (i.e., tests of experimental adequacy) insofar as the theory in which is embedded the method of investigation of overt verbal behavior presupposed by the theories themselves is open to confirmation and disconfirmation. Although we cannot directly determine the experimental adequacy of Grice's

and Searle's theories, we can test the theory in which is embedded the method of investigating the overt behavior of speakers presupposed by the theories themselves.

It should be clear, however, that part of the proposed naturalistic theory of meaning provides something very much like what is presupposed by Grice's and Searle's theories of meaning: namely, a method for investigation of speakers' overt behavior and of the connections obtaining between verbal behavior and objects and events in the world. Of course, the difference between the proposed naturalistic theory of meaning and the theory presupposed by Grice's and Searle's theories is that while the former takes meaning to be a matter of regular connections between behavior and certain naturalistic objects and/or processes (i.e., spatio-temporal objects and events sensed by speakers and speakers' internal states), the latter take these regularities to be merely evidence for the existence of underlying intentions or meanings. But, as was argued in Chapter 2, if meanings are supposed to be speakers' intentions, and if speakers' intentions must be investigated by investigating certain naturalistic objects and/or processes, then it is perverse to insist that meanings are hidden intentions, and not simply the naturalistic objects and/or processes which are alleged to be evidence for the intentions. Since we have, in the proposed naturalistic theory of meaning, a framework

for investigating the relevant naturalistic objects and processes, it would be simpler and more in keeping with the spirit of the scientific method to dispense with the use of intentions, and admit that meanings simply are naturalistic objects and/or processes.

To our proposal to dispense with the intentions, and to treat as meanings the naturalistic objects and/or processes which were heretofore treated as evidence for underlying meanings, it will be objected that we have confused the evidence for the theory of meaning with the subject matter of the theory. This is the standard objection to reductionist interpretations of mentalistic theories; the standard reply-- which I take to be valid in this case--is that since the only evidence we can have for the existence of what Grice and Searle call speakers' intentions are naturalistic objects and/or processes, it is incorrect to suppose that the speakers' intentions are anything other than the naturalistic objects and/or processes. Taking a realist approach to the ontological status of theoretical entities (as the mentalists in linguistics would have us do) we may say that until we have independent evidence for the existence of speakers' intentions, we have no choice (if we are to remain realists) but to identify meanings with naturalistic objects and/or processes.

We saw before, however, that the only significant difference between (a) the theory in which is embedded the method of investigation of connections between instances of verbal behavior and other naturalistic objects and processes

(including other instances of verbal behavior)--which is presupposed by Grice's and Searle's theories--and (b) the proposed naturalistic theory is that while the latter takes meaning to be a matter of certain naturalistic objects and/or processes, the former takes the naturalistic objects and/or processes to be merely evidence for the existence of underlying meanings. But in the preceding argument it was suggested that this difference between the two theories is illusory: it was suggested that naturalistic objects and/or processes cannot be merely evidence for underlying mentalistic meanings, for which there can be no other evidence, but are best construed as the meanings themselves. It follows that the theory in which is embedded the method of investigating naturalistic objects and processes which is presupposed by Grice's and Searle's theories of meaning is indistinguishable from the proposed naturalistic theory of meaning.

But we have also shown that the only part of Grice's and Searle's theories of meaning which is open to tests of experimental adequacy is the theory embedded in their method of investigating speakers' overt behavior--viz., the proposed naturalistic theory of meaning. It follows, therefore, that insofar as Grice's and Searle's theories of meaning are open to confirmation and disconfirmation, their theories are indistinguishable from the proposed naturalistic theory. This is to say that insofar as Grice's and Searle's theories are open to tests of experimental adequacy, the empirical

consequences of the theories are identical to the empirical consequences of the proposed naturalistic theory of meaning.

The proposed naturalistic theory of meaning is open to direct empirical confirmation and disconfirmation. The way to check the experimental adequacy of the theory is to determine whether the meaning of utterance *x* is in fact the objects and/or processes deemed by the theory to be the meaning of utterance *x*. The meaning of an utterance is to be discovered, extra-theoretically, by querying native speakers of the language. Since the responses of such speakers, and the objects and/or processes which are deemed by the theory to be meanings, are all open to direct observation, it is possible to test the experimental adequacy of the theory.

In sum, I have argued that insofar as Grice's and Searle's theories of meaning are open to empirical confirmation and disconfirmation, they are indistinguishable from the proposed naturalistic theory. The proposed naturalistic theory may be ultimately inadequate, but at least it is, as opposed to those parts of Grice's and Searle's theories which are distinguishable from it, open to empirical confirmation and disconfirmation. It is true that the proposed naturalistic theory has not been shown to be fully experimentally adequate, in Shaffner's sense. But at least it is possible to determine the experimental adequacy of the proposed naturalistic theory; since this is not the case for those parts of either Grice's or Searle's theories which are

distinguishable from the naturalistic theory, I conclude that the proposed naturalistic theory of meaning is more adequate within the dimension of experimental adequacy than is either Grice's or Searle's theories.

2) Theoretical Context Sufficiency

The second category within which our three theories may be compared is that of theoretical context sufficiency. Theoretical context sufficiency is a matter of the degree of concordance obtaining between the theory being evaluated and the corpus of accepted scientific knowledge. In what follows, I argue for two points: (a) that both Grice's and Searle's theories of meaning (insofar as they are distinguished from the naturalistic theory) conflict with a basic methodological principle accepted by virtually all of the theories within the corpus of accepted scientific knowledge, and (b) that there is some degree of concordance between the proposed naturalistic theory of meaning and the corpus of accepted scientific knowledge.

The argument for the first point is simple. One of the methodological principles of accepted scientific knowledge is that a proposed theory must be open to empirical confirmation and disconfirmation. Shaffner suggests that all acceptable scientific theories must be founded on this methodological principle, when the first category he proposes within which theories are to be evaluated is that of experimental adequacy: no theory which is not open to empirical confirmation and disconfirmation can be experimentally

adequate to the slightest degree.

It has already been shown, however, that insofar as Grice's and Searle's theories of meaning are distinguishable from the proposed naturalistic theory of meaning, they are not open to empirical confirmation and disconfirmation. There cannot be, therefore, a very high degree of concordance between those parts of either Grice's or Searle's theories which are distinguishable from the naturalistic theory and the corpus of accepted scientific knowledge. Any concordance there appears to be between those parts of either Grice's or Searle's theories which are distinguishable from the naturalistic theory and accepted theories in, say, linguistics, will be overruled by the lack of concordance between those parts of Grice's and Searle's theories and the basic methodological principle according to which scientific theories must be open to confirmation and disconfirmation. It has not been shown that there is any lack of concordance between Grice's and Searle's theories and any other particular scientific laws; rather, what has been argued is that crucial parts of Grice's and Searle's theories conflict with a basic methodological principle of virtually all accepted scientific knowledge.

That there is some concordance between the proposed naturalistic theory of meaning and accepted scientific knowledge is more difficult to establish, and, because of the present state of the science of linguistics, any evaluation of the theoretical context sufficiency of the proposed naturalistic theory of meaning will be open to question.

Several points seem clear, however. First, as shown above, the proposed naturalistic theory of meaning does not conflict with the same methodological principle with which crucial parts of Grice's and Searle's theories conflict. Because of the importance of the principle of openness to confirmation and disconfirmation, this first point grants a prima facie higher degree of theoretical context sufficiency to the proposed naturalistic theory of meaning than was granted to Grice's and Searle's theories, as these theories are distinguishable from the naturalistic theory.

The second point that seems clear is that insofar as the theories of structural linguistics are included in the corpus of presently accepted scientific knowledge, the proposed naturalistic theory of meaning may be granted a high score within the category of theoretical context sufficiency. This is so, because, as quotations from Bloomfield in Chapter 2 and elsewhere show, the proposed naturalistic theory of meaning is firmly rooted in the traditions of structural linguistics. Structuralists have generally agreed with the position advocated here, that the meanings of sentences are functions of the stimulatory situations of the sentences.⁶ However, that the relevant theories of structural linguistics are a part of the corpus of presently accepted scientific knowledge is a matter of some debate. I shall argue, however, that no amount of evidence accumulated on the lack of concordance between the currently ascendant linguistic theory--transformational

⁶See Harris, 1951, p. 187.

grammar--and the proposed naturalistic theory of meaning will serve to render the proposed theory inadequate (or, more importantly, less adequate than either Grice's or Searle's theories) within the dimension of theoretical context sufficiency. The main challenge from linguistics to the proposed naturalistic theory of meaning may be expected to come from transformational grammarians (since the proposed theory will naturally be identified with the theories of structural linguistics, which have been challenged by transformationalists). This is not the place to pose and reply to all the objections to a naturalistic theory of meaning from transformational grammar; however, my reasons for saying that any lack of concordance that exists between transformational grammar and the naturalistic theory of meaning will not diminish the adequacy of the naturalistic theory of meaning may be sketched.

In the first place, it is not obvious that there is in fact a high degree of direct conflict between transformational grammar and the proposed theory of meaning. This is so because most of the work of transformational grammarians has been devoted to syntax and phonology; almost no work has been done in semantics, and that which has been done has been called "trivial" by a sympathetic critic (Searle, 1972, p. 20). Although I must leave this argument incomplete, insofar as the work of transformational grammarians on issues in semantics cannot be dealt with here, I think it is fair to say that until transformational grammar yields a fully articulated theory of meaning which is clearly competitive with those

being considered here, there will be minimal conflict between the proposed naturalistic theory and transformational grammar.

In the second place, although Chomsky and others have shown there to be certain weaknesses in parts of the behaviorist account of language (with which the proposed naturalistic theory of language will naturally be associated), it is not clear that they have shown there to be anything wrong with the proposed naturalistic theory of meaning. Most of Chomsky's attacks against behaviorism have been attacks against the behaviorist account of linguistic acquisition; but, as a second sympathetic critic has noted, Chomsky has never claimed that it follows from his criticism of the behavioristic account of linguistic acquisition that no aspect of language can be given an adequate behavioristic or naturalistic account (Lyons, 1970, pp. 93-94). In particular, so far as I know, Chomsky has never argued that a naturalistic theory of meaning such as that proposed here cannot be adequate.

It may be replied that Chomsky's criticisms of the behaviorist account of linguistic acquisition must be taken as criticisms of a behaviorist theory of meaning, since he believes that the theory of linguistic acquisition is inextricably linked with accounts of all other aspects of language, in the sense that a grammar which is justified at the level of explanatory adequacy must be an "acquisition model" for the language in question.⁷ This view of Chomsky's may or may

⁷ Chomsky, 1965, pp. 30-37. The feeling that Chomsky's criticisms of the behaviorist account of linguistic acquisition can be taken as criticisms of the proposed naturalistic theory

not be correct; this is not the place to examine it in detail. Suffice it to say that neither Grice nor Searle attempt to build into their theories of meaning a theory of linguistic acquisition; therefore, if the proposed naturalistic theory is to be faulted for not incorporating the correct theory of linguistic acquisition, then so must Grice's and Searle's theories be faulted for failing to incorporate any theories of linguistic acquisition. That the theory of linguistic acquisition that Chomsky takes to be the correct theory is the only theory consistent with either Grice's or Searle's theories of meaning is far from clear. Therefore, it is premature to judge either (a) that either Grice's or Searle's theory of meaning is supported by Chomsky's theory of linguistic acquisition, or (b) that the proposed naturalistic theory of meaning is to be faulted for not incorporating the correct theory of linguistic acquisition.

In addition, concerning the challenge to the theoretical context sufficiency of the proposed naturalistic theory of meaning posed by transformational grammar, we may add that it is, in the opinion of many,⁸ too early to assess the importance of a challenge from transformational grammar. There is much disagreement among linguists as to the fruitfulness of

of meaning may be strengthened by the fact that Quine mixes his discussion of the theory of meaning--on which the proposed theory of meaning is based--with a discussion of a behaviorist theory of linguistic acquisition. I believe Quine's theory of meaning to be independent of his theory of linguistic acquisition, and I have attempted to present the naturalistic theory of meaning being considered here in complete isolation from any theory of linguistic acquisition.

⁸To cite only one important example, among many, see Hockett, 1968.

Chomsky's approach in linguistics, and until this disagreement is resolved it is difficult to say with which linguistic theories it is important for a theory of meaning to concord. A judgment of theoretical context sufficiency depends on a judgment of which theories are part of the accepted scientific corpus. Because of the insecure status of transformational grammar, it may not be prudent at this point to discontinue investigation of any hypothesis which appears to be incompatible with transformational grammar. For this reason, and for the reasons cited above, I conclude that the proposed naturalistic theory of meaning is more adequate than either Grice's or Searle's theories, within the dimension of theoretical context sufficiency.

3) Simplicity

Shaffner's category of simplicity is the final category within which the three theories of meaning are to be evaluated. There are three kinds of simplicity to be considered; the first is fitness. It will be remembered that the more fit a theory is, the fewer distinctions must it make which have no empirical consequences. While I shall leave it to critics to show that the proposed naturalistic theory of meaning is fit or unfit (or, indeed, simple or complex, in any of the three senses considered here) in the sense that, in Quine's terms, it spurns "distinctions which make no difference," it has already been shown that both Grice's and Searle's theories are based on distinctions which have no empirical consequences.

The distinction made by Grice which we have repeatedly shown to have no empirical consequences is his distinction between speakers' intentions and naturalistic objects and/or processes. Again, Grice says that the naturalistic objects and/or processes are merely evidence for the existence of the underlying intentions; but it has been replied that in the absence of any other evidence, we may simply dispense with the intentions and investigate the naturalistic objects and/or processes directly. In the absence of any independent evidence for the existence of speakers' intentions, the distinction between intentions and naturalistic objects and/or processes is a distinction without a difference.

Because Searle also claims that meanings are certain speakers' intentions, and not naturalistic objects and/or processes, his theory is based on the same distinction-without-a-difference that is Grice's. A second distinction of Searle's which has been shown to have no empirical consequences is the distinction drawn by Searle when he says that speaking a language is rule governed behavior rather than merely regular behavior. Of course I do not claim that the general distinction between rule governed and regular behavior is bankrupt: there is a difference between behavior which fits a rule (in the sense that the rule is a true description of the behavior) and behavior which is guided by a rule (in the sense that the rule causes the behavior). But a rule cannot guide behavior unless it is known and can be stated by the

agent;⁹ and Searle has in mind rules which need to be neither known nor stateable by the speaker (Searle, 1969, p. 42). It is Searle's distinction between rules of language and regularities of language which has been shown to be bankrupt.

Two arguments were used to show Searle's distinction to be bankrupt; one was taken from Wilfrid Sellars' discussion of language games, and will not be repeated here; the other is similar to our reductionist argument just reviewed against Grice's distinction between speakers' intentions and naturalistic objects and/or processes, and will be given. In the present case, we may note (a) that Searle agrees that all rule governed behavior is regular behavior, in the sense that all behavior that is governed by a rule is also covered by a generalization concerning regularities of verbal behavior, and (b) that Searle has no evidence for the (unknown and unstated) rules of language other than the regularities whose existence he uses to infer the existence of the rules. But in the absence of any independent evidence for the existence of the rules, we may say that Searle's distinction between verbal behavior which is rule governed and verbal behavior which is merely regular is a distinction which diminishes the fitness of his theory of meaning.

The second kind of simplicity to be considered is ontological simplicity. Here, the general rule is that the fewer kinds of entities posited by a theory, the simpler the theory--so long as a reduction in the number of entities

⁹V Quine, 1972, p. 442, on the distinction between fitting and guiding.

posited by a theory does not drastically complicate what Quine calls the "ideology" of the theory.¹⁰ The arguments just cited against Grice's and Searle's theories may now be seen to show that the two theories are each guilty of unnecessarily multiplying entities, and thereby decreasing their simplicity relative to the proposed naturalistic theory. Since Grice's distinction between speakers' intentions and certain naturalistic entities and Searle's distinction between rule governed and regular behavior are both purported to be ontological distinctions, and since we have shown the distinctions to have no significant empirical consequences, we may conclude that both Grice's and Searle's theories score low on ontological simplicity.

The third kind of simplicity considered by Shaffner is that of simplicity of system. Since this kind of simplicity requires that the theories being evaluated have been axiomatized, and since none of the theories being evaluated here have been axiomatized, we cannot evaluate the theories of meaning with Shaffner's category of simplicity of system.

Before concluding, a note may be added concerning an application of Shaffner's observation that "by making a virtue of simplicity we make a vice of complexity." One particularly odious kind of complexity is ad hocness of hypotheses. Theories which include ad hoc hypotheses are to be rated low on simplicity. Grice's theory of meaning (and, by extension, Searle's) has been demonstrated to be constituted almost solely by intra-systemic ad hoc hypotheses, and

¹⁰V Koslow, 1970, p. 359.

this demonstration deserves to be reviewed. MacKay has argued in "Professor Grice's Theory of Meaning" that Grice's theory of meaning is constructed in the following manner:

We begin by accepting [some] clause (1) as an account of meaning. A counterexample is suggested to clause (1), which is met by adding the qualifying clause (2). A counterexample is suggested to the resultant "theory" (i.e., clause (1) as qualified by clause (2)), which is then met by adding the qualifying clause (3). And so on. In short, Professor Grice introduces and justifies each successive addition, qualification, or in general each change N+1 in his theory, exclusively in terms of its success in defeating a counterexample which threatens the theory as qualified up to and including the latest addition N (MacKay, 1972, pp. 57-58).

MacKay shows that each addition N to Grice's theory has as its sole justification its ability to defeat a counterexample. He considers, for each successive addition N, whether it has any possible justification for either the meaning theorist (i.e., Grice) or the speaker whose meaningful utterances are being analyzed, and concludes that (a) it is implausible in all cases to say that any addition N has a justification from the point of view of the speaker, and (b) it is impossible in several cases for there to be any justification from the point of view of the speaker for certain additions. Therefore, most of the clauses of Grice's theory of meaning are perfect examples of Shaffner's intra-systemic ad hoc hypotheses--i.e., hypotheses which are added to a theory solely to enable it to outflank unexpected and embarrassing results--E-results--and which hypotheses (plus the theory under test) have no further testable consequences which differ from the E-results in an interesting way.

I shall omit the details of MacKay's demonstration; the only flaw in it which is apparent to me is that he has shown one of the intermediary versions of Grice's theory, rather than its final version, to be ad hoc. Nevertheless, I think that each of the versions of Grice's theory can be shown to be ad hoc, by arguments similar to MacKay's; even a cursory examination of the pattern of development of each version of Grice's theory of meaning (1967 and 1969) should indicate the plausibility of this view. The upshot is, then, that Grice's theory of meaning is to be rated low on simplicity, because of its ad hoc character; and, to the extent that Searle's Gricean theory of meaning incorporates the ad hoc hypotheses of Grice's original theory, Searle's theory is to be rated low on simplicity also.

The conclusion to be drawn from the above remarks on the adequacy of the three theories of meaning under review is that the proposed naturalistic theory of meaning is to be preferred over both Grice's and Searle's theories. Application of Shaffner's logic of comparative theory evaluation shows that the proposed naturalistic theory rates higher than does either Grice's or Searle's theories within each of the categories of experimental adequacy, theoretical context sufficiency, and simplicity. A cautionary note must be added, however, against interpreting our argument as yielding a final and absolute verdict in favor of the proposed naturalistic theory of meaning. The verdict given is relative to the time of the evaluation and to the person doing

the evaluating. What we have argued is not that the proposed naturalistic theory of meaning is the correct theory of meaning, but that at this time the proposed naturalistic theory of meaning is to be preferred, relative to the knowledge of this evaluator, over Grice's and Searle's theories of meaning. None of the three theories has as yet been stated in a rigorous and complete form, and until this is done any judgment of the adequacy of the theories must be recognized as tentative at best. In addition, some theory of meaning different from any of the theories considered in this thesis may emerge as more adequate than the proposed naturalistic theory of meaning. The argument in this thesis has been that the three currently dominant theories of meaning in philosophy of language are genuinely competing theories, and that an extension of one of them--Quine's--is more adequate than either of the other two.

REFERENCES

- Anscombe, G. E. M. On brute facts. Analysis, 1958, 18(3), 69-72.
- Austin, J. L. How to do things with words. London: Oxford University Press, 1962.
- Bennett, J. Rationality. London: Routledge & Kegan Paul, 1964.
- Bloomfield, L. Language. London: Allen & Unwin, 1935.
- Chomsky, N. Aspects of the theory of syntax. Cambridge, Mass.: M. I. T. Press, 1965.
- Chomsky, N. Linguistics and philosophy. In S. Hook (Ed.), Language and philosophy. New York: New York University Press, 1969.
- Cohen, L. J. The diversity of meaning. (2nd ed.) London: Methuen, 1962.
- Danto, A. C. Naturalism. In P. Edwards (Ed.), Encyclopedia of philosophy. New York: Macmillan, 1967.
- Davidson, D., & Hintikka, J. (Eds.) Words and objections, essays on the work of W. V. Quine. Dordrecht: Reidel, 1969.
- Davidson, D., & Harman, G. (Eds.) Semantics of natural language. Dordrecht: Reidel, 1972.
- Edwards, P. (Ed.) Encyclopedia of philosophy. New York: Macmillan, 1967.
- Feigl, H., & Maxwell, G. (Eds.) Minnesota studies in the philosophy of science. Vol. III. Minneapolis: University of Minnesota Press, 1962.
- Goodman, N. Fact, fiction, and forecast. (2nd ed.) Indianapolis: Bobbs-Merrill, 1965.
- Grice, H. P. Meaning. In P. F. Strawson (Ed.), Philosophical logic. London: Oxford University Press, 1967. Original publication in Philosophical Review, 1957, 66, 377-388.

- Grice, H. P. Utterer's meaning, sentence-meaning, and word-meaning. Foundations of Language, 1968, 4, 225-242.
- Grice, H. P. Utterer's meaning and intentions. Philosophical Review, 1969, 78, 147-177.
- Hare, R. M. Meaning and speech acts. Philosophical Review, 1970, 79, 3-24.
- Harman, G. H. Three levels of meaning. Journal of Philosophy, 1968, 65, 590-601.
- Harris, Z. S. Structural linguistics. Chicago: University of Chicago Press, 1951.
- Hempel, C. G. Aspects of scientific explanation and other essays. New York: Free Press, 1965.
- Hintikka, J. Behavioral criteria of radical translation. In D. Davidson and J. Hintikka (Eds.), Words and objections, essays on the work of W. V. Quine. Dordrecht: Reidel, 1969.
- Hockett, C. F. The state of the art. The Hague: Mouton, 1968.
- Hockney, D. J. Conceptual structures. In P. Maynard and G. Pearce (Eds.), Conceptual change. Dordrecht: Reidel, forthcoming.
- Hook, S. (Ed.) Language and philosophy. New York: New York University Press, 1969.
- Kiefer, H. E., & Munitz, M. (Eds.) Language, belief, and metaphysics. Albany: SUNY Press, 1970.
- Koslow, A. Comment by Arnold Koslow. In R. H. Stuewer (Ed.), Minnesota studies in the philosophy of science. Vol. V. Minneapolis: University of Minnesota Press, 1970.
- Lyons, J. Noam Chomsky. New York: Viking Press, 1970.
- MacKay, A. F. Professor Grice's theory of meaning. Mind, 1972, January, 57-66.
- Maynard, P., & Pearce, G. (Eds.) Conceptual change. Dordrecht: Reidel, forthcoming.
- Patton, T. E., & Stampe, D. W. The rudiments of meaning: On Ziff on Grice. Foundations of Language, 1969, 5, 2-16.

- Putnam, H. The analytic and the synthetic. In H. Feigl and G. Maxwell (Eds.), Minnesota studies in the philosophy of science. Vol. III. Minneapolis: University of Minnesota Press, 1965.
- Quine, W. V. O. Word and object. Cambridge, Mass.: M. I. T. Press, 1960.
- Quine, W. V. O. From a logical point of view. (2nd ed.) New York: Harper Torchbook, 1961.
- Quine, W. V. O. Linguistics and philosophy. In S. Hook (Ed.), Language and philosophy. New York: New York University Press, 1969. (a)
- Quine, W. V. O. Ontological relativity and other essays. New York: Columbia University Press, 1969. (a)
- Quine, W. V. O. To Chomsky. In D. Davidson and J. Hintikka (Eds.), Words and objections, essays on the work of W. V. Quine. Dordrecht: Reidel, 1969. (a)
- Quine, W. V. O. To Hintikka. In D. Davidson and J. Hintikka (Eds.), Words and objections, essays on the work of W. V. Quine. Dordrecht: Reidel, 1969. (a)
- Quine, W. V. O. Philosophy of logic. Englewood Cliffs: Prentice-Hall, 1970. (a)
- Quine, W. V. O. Philosophical progress in language theory. In H. E. Kiefer and M. Munitz (Eds.), Language, belief, and metaphysics. Albany: SUNY Press, 1970. (b)
- Quine, W. V. O. On the reasons for indeterminacy of translation. Journal of Philosophy, 1970, 67, 178-183. (c)
- Quine, W. V. O. Methodological reflections on current linguistic theory. In D. Davidson and G. Harman (Eds.), Semantics of natural language. Dordrecht: Reidel, 1972.
- Schick, K. Indeterminacy of translation. Journal of Philosophy, 1972, 69(22), 818-832.
- Schuldenfrei, R. Quine in perspective. Journal of Philosophy, 1972, 69(1), 5-16.
- Searle, J. R. Austin on locutionary and illocutionary acts. Philosophical Review, 1968, 77, 405-424.
- Searle, J. R. Speech acts. Cambridge: Cambridge University Press, 1969.
- Searle, J. R. (Ed.) The philosophy of language. London: Oxford University Press, 1971.

- Searle, J. R. Chomsky's revolution in linguistics. New York Review of Books, 1972, 18(12), 16-24.
- Sellars, W. Science, perception, and reality. London: Routledge & Kegan Paul, 1963.
- Sellars, W. Some reflections on language games. In W. Sellars, Science, perception, and reality. London: Routledge & Kegan Paul, 1963. Original publication in Philosophy of Science, 1954, 21, 204-228.
- Shaffner, K. E. Outlines of a logic of comparative theory evaluation with special attention to pre- and post-relativistic electrodynamics. In R. H. Stuewer (Ed.), Minnesota studies in the philosophy of science. Vol. V. Minneapolis: University of Minnesota Press, 1970.
- Skinner, B. F. Verbal behavior. New York: Appleton-Century-Crofts, 1957.
- Stampe, D. W. Toward a grammar of meaning. Philosophical Review, 1968, 77, 137-174.
- Stevenson, C. L. Ethics and language. New Haven: Yale University Press, 1944.
- Strawson, P. F. (Ed.) Philosophical logic. London: Oxford University Press, 1967.
- Stuewer, R. H. (Ed.) Minnesota studies in the philosophy of science. Vol. V. Minneapolis: University of Minnesota Press, 1970.
- Wallace, J. A query on radical translation. Journal of Philosophy, 1971, 68, 143-151.
- Warnock, G. J. Hare on meaning and speech acts. Philosophical Review, 1971, 80, 80-84.
- Wilder, H. T. Lewis and Quine on private meanings and subjectivism. Canadian Journal of Philosophy, 1971, I(1), 25-44.
- Winch, P. The idea of a social science. London: Routledge and Kegan Paul, 1958.
- Ziff, P. On H. P. Grice's account of meaning. Analysis, 1967, 28, 1-8.
- Ziff, P. A response to stimulus meaning. Philosophical Review, 1970, 79, 63-74.