Outline of the Theory

This study will address itself to the question, How is a generative grammar organized? In particular, how do the rules that deal with meaning interact with those that determine syntactic form? Katz and Postal's *An Integrated Theory of Linguistic Descriptions* (1964) presents an extremely appealing approach to this question that has had far-reaching consequences for linguistic theory. It is the contention of the present study, however, that their approach is incorrect, and that the vast amount of linguistic research that has taken place since 1964 can be accommodated more adequately within a rather different view of the role of semantics in grammar, to be proposed here.

1.1 The Problem of Semantics

In the early days of generative grammar, the nature of the rules relating syntactic structures to meaning was not discussed. Chomsky's *Syntactic Structures* (1957) shows that a linguistic theory in which meaning is determined at least in part by a level of underlying structure can capture important generalizations. But Chomsky does not propose explicit mechanisms for representing or deriving meaning; his main concern is with the formal syntactic devices of the language. With the publication of Katz and Fodor's "The Structure of a Semantic Theory" (1963), the picture changed. Katz and Fodor argue that a grammar should be thought of as a system of rules relating the externalized form of the sentences of a language to their meanings. Hence a complete linguistic description must contain an account of meaning.

Katz and Fodor suggest that meanings are to be expressed in a universal semantic representation, just as sounds are expressed in a universal phonetic representation. Universality is necessary so that representations are language-independent; we must be able to compare meanings of sentences across languages. Put more strongly, to suppose a universal semantic representation is to make an important claim about the innateness of semantic structure. The semantic representation, it is reasonable to hope, is very tightly integrated into the cognitive system of the human mind.

Of course, compared to phonetic representations, semantic representations are only very indirectly accessible. It is fairly easy to talk about sameness and difference of meaning, but meaning itself, as generations of philosophers have known, is elusive. Thus the study of semantics has always been somewhat derivative, indirect, and fuzzy. It was Katz and Fodor's hope that by making semantics an explicit part of generative grammar, more incisive studies of meaning would be possible. And to some extent their hope has been realized, in that generative grammar has permitted the construction of more highly structured hypotheses about meaning.

It has generally been assumed that semantic representations are not formally similar to syntactic structures (the theory of generative semantics, however, denies this). A complete linguistic description, therefore, must include a new set of rules, a semantic component, to relate meanings to syntactic and/or phonological structure. This is the content of Katz and Fodor's slogan, "linguistic description minus grammar equals semantics" (where "grammar" is used to mean "syntax and phonology"). Katz and Fodor's phraseology, however, is unfortunate: their slogan seems rapidly to have acquired the negative connotation "Semantics is whatever you have to shove under the rug." This interpretation was perhaps a predictable outcome of the relative availability of syntactic and semantic formalisms; it is always less troublesome to defend a syntactic solution to a problem where the formalism is taken for granted than to solve it by developing a new semantic formalism that may not appear independently motivated. Thus research has been biased heavily in favor of syntactic solutions to problems.

It is the intent of this investigation to begin to right this imbalance. By now, many more grammatical phenomena have been studied in detail than in 1963, and much more is known about the criteria that must be met by a system of semantic representation and the rules relating it to syntactic structure. We will approach a number of these phenomena with the possibility of a semantic solution in mind. Insofar as possible, concrete semantic solutions will be proposed and defended with a rigor at least equal to that generally accepted for current transformational formulations. The variety of phenomena covered will permit an integration of the solutions into a more comprehensive and precise theory of the semantic component than has heretofore been proposed.

1.2 Semantic Representation and the Semantic Component

Katz and Fodor characterize semantic representation as a structured bundle of "semantic markers." From their relatively primitive model, a model of semantic representation has evolved in common use that treats semantic representation as something structured like a phrase-marker that is rather similar to syntactic representations, with perhaps some additional information added on somehow. What is taken for granted is that there is basically a single hierarchical structure into which the semantic material of the lexical items in the sentence is arranged.

We will take a fundamentally different approach here. In an attempt to account for a large range of semantic phenomena, we will find that these phenomena divide themselves into a number of independent groups for which rather different analyses are required (see section 1.5). To make clear the independence of these different aspects of semantic representation, we will separate semantic representation into four parts, including two hierarchical structures. Very crudely, the first hierarchical structure, the *functional structure*, represents relations in the sentence induced by the verbs, including such notions as agency, motion, and direction. The *modal structure*, the second hierarchical structure, specifies the conditions under which a sentence purports to correspond to situations in the real world. The *table of coreference* indicates whether pairs of noun phrases in the sentence are intended to be coreferential or not. The *focus and presupposition* designate what information in the sentence is intended to be new and what is intended to be old. The failure of earlier studies to properly distinguish these semantic substructures, particularly the two hierarchical structures, has been the source of much difficulty and confusion.

The commonly accepted view of the semantic component, proposed by Katz and Postal (1964), is that the only syntactic information used in determining semantic representation is the underlying (deep) structure. The motivation of this view, some objections to it, and some of its consequences will be discussed more thoroughly in section 1.3. What is noteworthy here is that this view is based on the assumption that functional structure is the sole source of semantic information. From the beginning of generative grammar, the idea that functional structure is preserved by transformations has been fundamental. In fact, one of the original arguments for the explanatory adequacy of a level of underlying syntactic structure, related to the surface by transformations, was that this under-
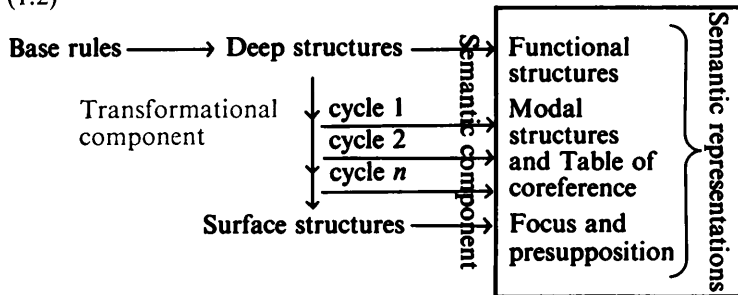
lying level expresses necessary generalizations about understood subjects and objects of verbs—for example the active-passive relationship or the ambiguity of sentences like *I found the boy studying in the library*.

The present study will propose a different view of semantic interpretation, namely, that various parts of semantic representation are related by the semantic component to various levels of the syntactic derivation. The difference between the two theories of the semantic component can be illustrated by the schematic diagrams (1.1), a grammar incorporating Katz and Postal's proposal, and (1.2), the alternative to be argued for here.

(1.1)

Semantic representations

↑

Semantic component

Base rules ———→ Deep structures

Transformational component

Surface structures
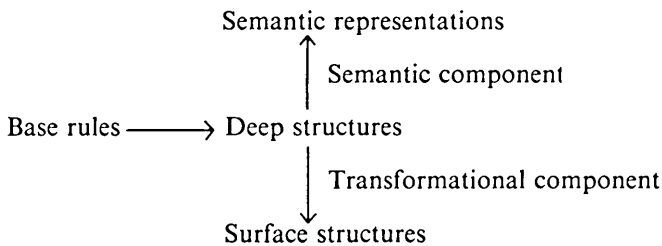
(1.2)

In the most recent variant of (1.1), generative semantics, it is claimed that the semantic component consists of at most an isomorphism, so that the base rules essentially generate semantic representations directly. In this case, given the nature of structures generated by the base within that theory, it can be said that there is no independent level of deep structure that is of any interest. But if (1.2) represents the organization of the grammar, one cannot claim that the "deep structures" we have been searching for are actually semantic representations, and that the level of deep structure can therefore be dispensed with. Such a claim is incomprehensible, since the

output of the base rules can represent only one of the aspects of semantic representation. On the contrary, in model (1.2) we must retain the conception of deep structure as representing a level of *syntactic* generality, the conception that originally motivated its existence.

Before going on to some of the basic issues involved in the decision to pursue (1.2) as a linguistic theory, it might be well to remark on one non-issue. To many people, (1.2) intuitively seems to be a repugnant way to organize the grammar. To some extent this feeling is natural, since (1.2) is more complex than (1.1). But it must be emphasized that the choice of a linguistic theory must be made empirically, not on the basis of intuition or personal preference. The reader is enjoined to weigh the considerable empirical evidence presented in this study before making his decision.

One particular point where intuition may be misleading is the question of performance models. The straight-line model of grammar (1.1) is admittedly very appealing in that it makes a performance model look easier to construct, particularly if we (pointlessly) reverse the direction of the upper arrow in (1.1). The performance model that comes to mind, however—one in which the language user actually performs derivations in his head—is open to serious doubt in any event (see for example, Fodor and Garrett 1966 and 1967, and Bever 1970). On the other hand, (1.2) requires a performance model in which some sort of parallel processing is taking place in the construction or interpretation of a sentence. It is hard to see immediately how such a performance model could work. But one must not let a lack of imagination dictate what are truly empirical decisions. From what is known about performance even in the temptingly easy-looking domain of phonetics, it is clear that the correct performance model will involve as yet undreamed-of subtleties. So much more should we expect nonobvious solutions in areas as abstract as syntax and semantics. The conceptual difference between (1.1) and (1.2) is undoubtedly trivial compared to the complexity of an adequate theory of performance.

1.3 The Katz-Postal Hypothesis

Because theory (1.1) has won such general acceptance, it is important to examine the reasoning that led to its adoption. Katz and Fodor (1963) suggest that semantic interpretation is performed by a set of *projection rules*. These rules add to the semantic representation of a sentence those parts of its content and organization not due to the lexical items, that is, the part of the interpretation traceable to the syntactic structure.

The syntactic structure of a sentence is generated by the application of a sequence of rules—first phrase-structure rules, then transformations. The differences in structure between two sentences are produced by differences in the sequences of rules which generate the sentences. Since two sentences containing the same lexical items can differ in meaning only if different sequences of rules have applied, it seemed natural to Katz and Fodor to attribute the meaning contributed by the structure to the operations of the rules themselves; that is, for each phrase structure rule and transformation in the grammar, there would be an associated projection rule telling how the phrase structure rule or transformation contributes to the meaning of sentences in which it operates. Thus projection rules can be divided into two classes, those that are associated with phrase structure rules (called type 1 projection rules or P1) and those associated with transformations (called type 2 projection rules or P2).

The type 1 projection rules create readings for a tree by starting at the lowest constituents in the tree and successively amalgamating readings of sister constituents to produce a reading for the mother constituent. Eventually, when the readings of all constituents have been amalgamated, there is a reading associated with the highest node, S. For each phrase structure rule telling how to expand a node into its constituents, there is a projection rule telling how to amalgamate the readings of the constituents to form a reading for the higher node. There are two examples of P1 rules in Katz and Fodor: a rule of attribution accounting for modifier-head relationships, and a rule combining the interpretation of a verb with the interpretation of its subject or object.

Type 2 projection rules are much less substantively discussed. Presumably a P2 rule shows how a given transformation changes meaning. When Katz and Fodor wrote their paper, P2 rules were absolutely essential, because the combining of kernel sentences into complex structures was regarded as a transformational operation: full interpretation of a complex sentence thus required projection rules to interpret the effect of the embedding transformations. Since then, however, recursion has come to be regarded as a property of the base component, so this particular type of P2 rule is no longer necessary.

Katz and Postal's *An Integrated Theory of Linguistic Descriptions* (1964) set out to show that no P2 rules are necessary, that is, that no changes of meaning are induced by transformations. First, let us consider obligatory transformations. Under the Katz-Fodor conception of projection rules, differences in meaning must be associated with choice-points in the der-

ivation. Since there is no choice whether or not to perform an obligatory transformation, there can never be two sentences differing in meaning solely because of the application of the transformation. Therefore there cannot be P2 rules associated with obligatory transformations. This leaves two cases: optional singulary transformations and generalized, or embedding transformations.

Embedding transformations have already been disposed of by assigning their recursive properties to the base. Optional singularities fall into three types: deletion, insertion, and order-changing rules. Deletion rules can be kept from changing meaning if we adopt *recoverability of deletion* as a condition on transformations (Katz and Postal, p. 81). The known insertion rules all insert meaningless particles such as *do* and inflectional markers, so they do not change meaning. Order-changing rules preserve understood grammatical relations. Since the Katz-Fodor theory tacitly assumes that grammatical relations (as generated by phrase-structure rules and interpreted by P1) are the only kind of structural information relevant in semantic interpretation, order-changing transformations preserve meaning as well.

Thus under Katz and Postal's assumptions about semantics, no P2 rules appear necessary. Furthermore, Katz and Postal show that stating a P2 rule is tantamount to restating the operation of the transformation with which it is associated, so that the whole concept of P2 is suspect. Therefore, the whole burden of semantic interpretation falls on P1, which are associated with the phrase-structure rules of the base component.

Katz and Postal's argument hence leads to the following conclusion:

Katz-Postal Hypothesis, weak form (KP1):
Semantic projection rules operate exclusively on underlying phrase-markers; hence transformations do not change meaning.

From the conclusion that all information required for the operation of projection rules is present in underlying structure, it is a simple rhetorical step to

Katz-Postal Hypothesis, strong form (KP2):
All semantic information is represented in underlying structure.

Given the Katz-Fodor conception of projection rules associated with the choice-points in derivations, and given the limited power proposed for

projection rules, Katz and Postal make a very good case for even the strong form of their hypothesis. The Katz-Postal Hypothesis is, of course, theory (1.1) of section 1.2.

KP2 makes a very strong claim. Just how strong it is can be seen from the following paraphrase: every nonlexical semantic difference must be represented as a deep structure difference. But even this strong a claim seems to have been accepted nearly universally.

KP2 follows from KP1 if the semantic component is composed entirely of simple combinatorial rules like those given by Katz and Fodor. But suppose that some projection rule actually *adds* meaning, that is, it adds semantic markers to an interpretation other than those contained in the lexical items in the sentence. Suppose further that this rule is optional, and that the information it adds is not essential for a well-formed interpretation. Then it is possible for there to be two interpretations for a single sentence, differing only in whether this optional projection rule has applied. The logical possibility of such a situation shows that KP2 and KP1 are not equivalent: projection rules may operate exclusively on underlying phrase-markers without all semantic information being represented there. Hence KP2, the universally accepted form of the hypothesis, does not follow from KP1, the form Katz and Postal state and claim to prove.

KP1 itself is open to question as well. Katz and Postal's treatment of apparent counterexamples is largely correct, but there is at least one serious error, and subsequent research has uncovered a great many more areas of the grammar where KP1 cannot be justified on independent syntactic grounds. Much of this book will be devoted to discussing these areas.

The error in Katz and Postal's analysis is in their discussion of negation. In the grammar of *Syntactic Structures*, negation was simply added by an optional transformation, which obviously changed meaning. Katz and Postal point out (p. 74) that the analyses of Lees (1960) and Klima (1964) independently motivate a negative morpheme in underlying structure, so that no transformation which changes negative sentences to positive sentences is necessary. However Katz and Postal overlook the fact that even with the negative morpheme, Klima's system of rules is not meaning-preserving: he derives from the same underlying structure the nonsynonymous sentences *Not much shrapnel hit the soldier* and *much shrapnel didn't hit the soldier*. We will discuss examples of this sort and attempts to preserve KP1 in spite of them in Chapter 8.

How can we simultaneously reject KP1 and accept Katz and Postal's

arguments against type 2 projection rules? It turns out that in fact there is no consistent way to characterize the way the passive, for example, changes meaning, so there cannot be a projection rule for the passive. Rather, the generalizations that we will observe concern derived structure configurations, regardless of what transformations are involved in the derivation. Hence the semantic rules we propose here will be of a type not envisioned by Katz and Fodor—rules that interpret derived structure rather than derivations.

It is important to notice the significance of deep structure in a theory incorporating KP2. KP2 claims that all meaning differences are represented in deep structure. Since projection rules therefore cannot add or change any elements of meaning, deep structures must in fact represent logical structure rather than syntactic structure if any conflict between the two arises, regardless of complexities thereby incurred in the transformational component. The need for "abstract" deep structures exemplified in G. Lakoff (1968a, 1970a, 1971), Ross (1969a, 1970), Bach (1968), R. Lakoff (1968), and McCawley (1970) is a natural consequence of KP2.

As deep structures become more and more "abstract," they gradually become denuded of syntactic significance. Distributional facts that can be captured naturally by a relatively "shallowly" conceived base must instead be explained by somewhat arbitrary restrictions on transformations. Thus aspects of linguistic description that can be accomplished within the relatively limited power of a context-free phrase-structure grammar must instead be delegated to a much more powerful transformational component, complete with a full and extremely powerful theory of exceptions. The concomitant limitation of the base is hardly commensurate. An oft-repeated argument that the semantic component is simplified by the use of abstract deep structures is not justified: the transformational component, it can be argued, would be correspondingly simplified by writing certain processes as rules in the semantic component. And in fact, we will show that for many grammatical processes a semantic account adds less total machinery to the grammar than does a syntactic account. Without formulating serious proposals in both frameworks, no such decision can be made.

Following the tendency toward abstract deep structure to its logical conclusion, the deep structures grow closer and closer to one's intuitive notion of "semantic representation." This has to be the case, since the power of the semantic component is so severely constrained by KP2. One

is thus led rather persuasively to the idea of dispensing altogether with the autonomous level of deep structure, generating instead semantic representations, and proceeding directly to surface structures by means of transformations. Such a proposal is made in McCawley (1968a, 1968b, 1970) and discussed further in G. Lakoff (1970b) and Postal (1970).

Implementation of such a theory of "generative semantics," however, seems to lead to an extremely unconstrained conception of transformations. Furthermore, G. Lakoff (1969, 1970b) has argued that even less constrained devices, derivational constraints, must be widely used in a grammar based on generative semantic principles. The conceptual generality of a grammar containing only base rules, transformations, and derivational constraints has been claimed (for example, by Postal, to appear) as a great advantage over the theory proposed here, which contains several types of semantic rules in addition to phrase-structure rules and transformations. Lakoff (1970b) points out that semantic rules operating on derived structure are expressible as special cases of derivational constraints, concluding that they do not form part of a substantively different linguistic theory.

However, these arguments miss the point. To see why, we must consider more carefully what factors enter into the choice between two competing theories.

1.4 On Choosing Between Two Theories

Let us now make clear what empirical issues are involved in deciding between models (1.1) and (1.2). I take a linguistic theory to be an abstract representation of human language ability.[1] It defines a set of grammars of individual languages, each of which is claimed to be a possible human language. Therefore the adequacy of a linguistic theory is measured by how well the class of grammars it defines matches the class of human languages.

If it were shown that the class of grammars defined by theory (1.1) were identical to the class defined by theory (1.2), we would of course decide in favor of (1.1) on aesthetic grounds. Similarly, if the set of grammars de-

[1]The possible degree of abstraction of linguistic theory is easily underestimated. Kepler's laws of interplanetary motion are certainly abstract in that they are mathematical relations relating position and velocity to time; they do not describe the inertial and gravitational mechanisms governing the motion. Yet they constitute a strong explanatory theory relative to other theories of Kepler's time. It may well be that the present stage of linguistic theory is equally abstract. This does not make the theory less interesting.

fined by a theory incorporating only phrase-structure rules and no trans-
formations were identical to the set of grammars defined by the theory of
transformational grammar, we would decide in favor of the former. When
defending a structurally complex theory against a structurally simple one,
one must show that the two theories do not define the same set of gram-
mars, and that the complex theory proves a more accurate model of the
data. To defend the simpler theory, one need only show that the two
theories are empirically equivalent.

There are three kinds of differences that we will adduce in favor of theory
(1.2) (corresponding to Chomsky's (1965) levels of observational, de-
scriptive, and explanatory adequacy). First, if it can be shown that theory
(1.1) cannot define a grammar that generates the full range of English
sentences, theory (1.2) may include such a grammar. That alone would
justify the more complex theory. In practice, however, this argument is not
so easy to advance, because of the nearly unlimited ability of any theory of
this scope to accommodate awkward points. For example, Chomsky (1970b)
points out that one can always simulate derived structure rules of semantic
interpretation in a theory of the form (1.1), by generating constituents of
arbitrary structure in the base, filtering them in the base for the desired
semantic property, then using a filtering transformation at the desired
point of the derivation to match these arbitrary structures with the derived
structure. In the course of this study, we will see many examples of such
attempts to save theory (1.1).

A second difference we may find between the two theories is in their
ability to express significant generalizations about the language. When
applying the alternative theories to simple cases, it often turns out that they
require virtually the same amount of machinery. When we dig deeper,
however, it often turns out that they make slightly different but crucial
predictions. For example, it will be shown in Chapters 4 and 5 that in theory
(1.2), the three fundamental rules Pronominalization, Reflexivization, and
Complement Subject Interpretation can be ordered together, enabling us
to capture the substantial similarity in their environments. This generaliza-
tion cannot be captured in theory (1.1) without additional constraints of a
brute force nature. Many such differences arise in the course of this in-
vestigation. They may be relatively small points in the entire description of
the language, but given two theories as sophisticated as those we are com-
paring, it may often be the small points of generality that decide between
them.

The third and perhaps most important difference we may find between the theories concerns the classes of possible grammars defined by the theories. A linguistic theory claims that the grammars it defines all correspond to possible human languages. Hence a theory can be defective in that it defines grammars which, insofar as we can determine, do not correspond to any human language. To clarify terminology that is often misinterpreted, such a theory can be said to be "too powerful" in that it defines too large a set of grammars, or "too weak" in that it fails to constrain the class of possible grammars sufficiently. The addition of several new kinds of semantic interpretation rules in theory (1.2) appears superficially to make (1.2) a more powerful, or less constrained, theory. But it will turn out that we will be able to place substantial limitations on the power of each type of rule in the grammar, including transformations, so that the end result will be a smaller class of possible grammars, or a less powerful theory, than theory (1.1).

I will illustrate first a very specific example of a heavier constraint possible in theory (1.2), then go on to somewhat more general and speculative discussion. In the standard theory of pronominalization, perhaps the most natural under theory (1.1), it is claimed that pronouns are transformationally reduced forms of fully specified NPs. Transformations producing pronouns must, among other things, verify that a potential pronoun is coreferential with its antecedent. Therefore this theory must specify that coreferentiality of two NPs is a possible condition on transformations. What is not explained is why only pronominalization-like rules ever make use of this kind of condition. Why, for example, is there no rule that moves an NP if it is coreferential with an NP elsewhere in the sentence? In the interpretive theory of pronominalization developed in Chapter 4, where pronouns are generated by the base and semantic rules determine their antecedents, there is no need for transformations ever to refer to coreference conditions. Hence we can deprive transformations of the ability to refer to coreference and construct a weaker theory in which a movement rule dependent on coreference cannot be stated. Thus the interpretive theory, while it has an extra conceptual device, that is, semantic rules which establish coreference on the basis of derived structure, describes a more constrained set of grammars.

A more general constraint on transformations permitted by theory (1.2) concerns the integrity of lexical items. Chomsky (1970a) proposes the Lexicalist Hypothesis, roughly, that transformations do not perform de-

rivational morphology. This hypothesis in turn leads to constraints on the
form of the base and to generalizations about certain transformations (see
also Bowers 1969a and Jackendoff 1968c and 1971a for generalizations
expressible under the Lexicalist Hypothesis). Theory (1.2) is consistent with
the Lexicalist Hypothesis, and perhaps with a stronger position: trans-
formations cannot change node labels, and they cannot delete under
identity or positive absolute exception (see sections 5.2 and 6.9–10). This
means that the only changes that transformations can make to lexical items
is to add inflectional affixes such as number, gender, case, person, and
tense. Transformations will thus be restricted to movement rules and
insertion and deletion of constants and closed sets of items. We will refer
to this set of constraints as the Extended Lexical Hypothesis.

Another heavy constraint on the power of transformations is proposed
by Emonds (1970). Roughly, he proposes that with a certain specifiable
class of exceptions, the output of a transformation must be a structure that
can be independently produced by a base rule. Thus, for example, he
claims that it is no accident that the deep object of a passive sentence comes
to occupy subject position rather than perhaps a position between the
auxiliary and the main verb, where no noun phrase can be generated in
the base. Likewise, it is no accident that the deep subject of a passive ends
up in a prepositional phrase which is like all other prepositional phrases,
rather than in some altogether new kind of constituent. The exceptions to
Emonds's generalization are transformations that operate only in a special
class of clauses, primarily main clauses; these transformations each per-
form one of a very small class of possible operations. Clearly this hypoth-
esis puts very strong constraints on the notion "possible transformation."
It seems much more likely to be true of the transformations needed by the
theory proposed here than it is of those needed by a theory incorporating
the Katz-Postal Hypothesis.

Of course, these heavy restrictions on transformations must be accom-
panied with concomitant restrictions on possible semantic rules, if the
number of possible grammars is to be reduced. But the rules to be proposed
here fall into a small number of very restricted types, and places where they
apply in the syntactic derivation are similarly restricted. One would hope
that the rule types and their orderings would be universal, for example that
the coreference rules, whatever their exact form in a given language, would
be rules of semantic interpretation operating at the end of each transforma-
tional cycle, as theory (1.2) suggests they are in English. If this claim is

correct, the number of possible grammars is greatly reduced, since all the ordering problems of the standard transformations dealing with coreference are no longer open to question.

Thus it appears quite possible that the addition to linguistic theory of semantic rules applying to derived structure, though it produces conceptually more complex grammars, results in fewer possible grammars, and grammars that capture more generalizations. These are the empirical criteria that measure linguistic theories.

The remainder of this chapter will lay the groundwork for the development of a semantic component.

1.5 Elements of Semantic Interpretation
Let me sketch some of the things which must appear in semantic representations. Katz and Fodor are interested in such properties of readings as synonymy, analyticity, anomaly, and truth conditions, as well as the actual content of the reading. Here we will be primarily interested in the content of the reading, and how it is derived.

First, to strike a discouraging note, it is not even clear that one can construct a formal object which corresponds to the intuitive notion "semantic interpretation of a sentence," because of the infinite divisibility of many semantic properties and the (perhaps undecidable) problem of choosing what information is part of the reading and what information merely follows from the reading (see for example Wittgenstein 1958 and Quine 1960). Much of the difficulty in defining semantic readings arises in trying to represent the meanings of lexical items. Here, however, we will be more concerned with how the meanings of lexical items are combined to form meanings of sentences on the basis of syntactic structures, in other words, the contribution of structure to the meanings of sentences. And within this domain, I think it is possible to separate out certain discrete aspects of meaning and deal with them coherently. This is not to imply that specifying the meanings of lexical items is any less important. It is just a separate and perhaps more difficult problem.

The aspect of semantic representation that is perhaps most closely linked to syntactic structure is the *functional structure* of a semantic reading. We can think of verbs as semantic functions of one or more variables, the readings of syntactically associated noun phrases providing semantic values for the variables. Under this assumption, each verb in the deep

structure of a sentence presumably corresponds to a function in the seman-
tic representation. The embedding relations of functions in the semantic
representation will presumably mirror the embedding relations of verbs
(and other functional words) in the deep structure. This part of semantic
representation was recognized by Katz and Fodor; in subsequent work it
was assumed to be the only contribution of syntactic structure to semantic
representation.

A refinement of this aspect of meaning might provide for a partial
analysis of verbs into semantic subfunctions such as *causative*, *directional*,
and so forth, giving the semantic representation of a verb some internal
functional structure. Such an analysis can provide a way of grouping verbs
into natural semantic (and syntactic) classes and thus explain certain
similarities in behavior. The proposals of Katz (1966) and Gruber (1965,
1967a, 1967b) are attempts to analyze verbs in this fashion. The "higher
pro-verbs" of G. Lakoff (1971) and McCawley (1968b) and the case gram-
mar of Fillmore (1968) are attempts to represent this internal structure
externally, as a part of syntax. But basically there is no disagreement on the
claim that these semantic properties can be represented structurally, and
that it is the deep structure which determines them. In Chapter 2 we will
discuss further this aspect of meaning.

Other elements of semantic representation do not lend themselves to
being represented in trees or functional form. One example is coreference
relations among noun phrases. Although the determination of coreference
relations does depend on syntactic structure, the semantic notion "NP^1 is
(non)coreferential with NP^2" has nothing to do with the functional struc-
ture of sentences. For example, to say that *John* is the subject of *knew*, that
you wouldn't believe him is the object of *knew* and that *him* is the object of
believe in the sentence *John knew you wouldn't believe him* is to say nothing
about whether *John* and *him* are to be understood as the same individual.
Rather, an independent device is necessary to express coreference relations.
Referential indices, introduced in *Aspects of the Theory of Syntax*, are one
such device. Here we will use a different formalism, a *table of coreference*
independent of the functional structure. Each entry in the table will contain
a pair of NPs and a relation "coreferential" or "noncoreferential" obtain-
ing between them. For certain formal reasons which will appear in Chapter
4, this notation has additional advantages. What is important to observe
for the present is that however coreference is marked, it is clearly not the

same kind of semantic information as functional structure. Chapters 4 and 5 will develop the rules deriving the table of coreference.

Another element of semantic interpretation which has nothing to do with the functional structure is *focus and presupposition*. Various concepts have been discussed under these names and also under such names as *topic-comment* or *thematic structure*.[2] Here we will use *focus of a sentence* to mean "the information in the sentence that is assumed by the speaker not to be shared by him and the hearer"; *presupposition of a sentence* will mean "the information in the sentence that is assumed by the speaker to be shared by him and the hearer." Changing the focus and presupposition of a sentence, for example by introducing emphatic stress, does not change the understood functional relationships between verbs and their arguments: *JOHN saw Bill*, *John SAW Bill*, and *John saw BILL* differ in focus and presupposition, but John is performing the same action with respect to Bill in each case. Chapter 6 will discuss the rules which determine this element of the interpretation.

The scope of negation and quantifiers is another independent aspect of semantic interpretation. Negation and quantifiers may appear to be part of the functional structure, since we can set up expressions in the predicate calculus in which negation and quantifiers appear to have function-like behavior. Such an approach has been taken in Carden (1968) and G. Lakoff (1971). However, the syntactic and semantic behavior of negation and quantifiers is sufficiently different from that of verbs that this claim cannot be made lightly. Chapter 7 will argue that this aspect of interpretation is more adequately represented in a second hierarchical semantic structure, the *modal structure*, which has considerably different properties than the functional structure.

Three semantic properties having to do specifically with reference must appear in semantic representations. The first is specificity of indefinite NPs (see Baker 1966 and Dean 1968): *Fred wants to meet a voluptuous blonde* is ambiguous as to whether or not the speaker can point out the girl Fred wants to meet. The second is genericity: *A unicorn is a dangerous beast* expresses properties of the species *unicorn*, not of some individual. The third is referential opacity: *John thinks that the book that was burned was not burned* is ambiguous in that it can ascribe to John either an inconsistent or an incorrect belief, depending on whether he is to be responsible for the

[2]See Halliday (1967). Halliday's use of the term *theme* is emphatically not to be confused with the use to be introduced in Chapter 2 as an expression of functional structure.

correctness of the description *the book that was burned.*[3] Attempts to treat these properties in terms of functional structure (for example Baker 1966 and Bach 1968) have inevitably required deep structures far removed from the surface and powerful transformational apparatus. Chapter 7 will suggest that the formalisms of the modal structure are well adapted to express the first two of these properties of semantic representation.

Finally, the illocutionary force of a sentence is an aspect of semantic interpretation that has recently been open to heated discussion. It has been argued, most comprehensively by J. R. Ross (1970a) and R. Lakoff (1968) that whether a sentence is a declarative, an imperative, or an interrogative should be represented explicitly in the functional structure. Anderson (1968b), Fraser (1970), and Culicover (1970) argue against this position in several ways; Culicover goes on to propose an approach that treats them as independent from the functional structure, thereby gaining several important generalizations in the transformations. In Chapter 7, we will see a number of indications that illocutionary force is a further element of interpretation that can be represented in the modal structure.

The semantic interpretation of a sentence, then, is to be viewed as a collection of information of various sorts about different aspects of the meaning. To say that because of its complexity, this view is inferior to a position claiming that all meaning can be represented as functional structure is only to assert prejudice. Again it must be emphasized that the decision is empirical; it must be based on the relative adequacy of the linguistic theories which entail these views of semantic representation.

1.6 Well-Formedness Conditions on Semantic Interpretations
The box on the right-hand side of diagram (1.2) represents the collection of elements of meaning assigned to a sentence by the semantic interpretation rules of the grammar. What is not necessarily determined by the grammar is whether this collection of disparate elements actually forms a sensible meaning. To determine this there must be a set of well-formedness conditions on semantic interpretation. Some of these conditions are parts of the grammar, but others shade off into pragmatics or knowledge of the real world. At this point I will present only two examples of well-formedness conditions, both fairly obvious, and mention a few more to be developed later on.

[3]Referential opacity has been discussed by many philosophers, including Frege, Russell, Carnap, and Quine. Janet Fodor (1970) discusses it in the framework of generative grammar.

A first example of semantic well-formedness conditions might be *selectional restrictions*. Under this hypothesis, (1.3), for example, would be generated by the grammar and receive an interpretation, but the interpretation it receives would be nonsensical.

(1.3) Colorless green ideas sleep furiously.

This position is taken by Chomsky in *Syntactic Structures*, but not in *Aspects*. Jackendoff (1966a) and McCawley (1968a) argue for a return to Chomsky's earlier position. (Ironically, G. Lakoff (1968a) points out that his arguments for the "abstract" source for instrumental adverbs hold only if one accepts the position of *Aspects*, which Lakoff has since given up in favor of the position taken here.)

There are a number of reasons for taking this position. To claim that sentences with selectional violations are not generated by the syntax makes it impossible to produce perfectly acceptable sentences like (1.4).

(1.4) It's crazy to talk of rocks eating.

If, instead, sentences with selectional violations were generated by the syntax but received no reading, (1.4) could not be interpreted. If they were interpreted, but the interpretation they received were simply ANOMALY, containing no information of the lexical items and their semantic relations to each other, then we would predict (1.4) and (1.5) to be synonymous: both would mean "It's crazy to talk of ANOMALY."

(1.5) It's crazy to talk of Bill elapsing.

We see therefore that sentences with selectional violations must receive interpretations; hence they cannot be filtered out before readings are completed.

Furthermore, there are cases where selectional restrictions cannot be applied until the readings of an indefinite number of constituents have been amalgamated to form a reading. Compare (1.6) and (1.7):

(1.6)   I ate something that was the result of what Bill acknowledged to be a new baking process.
(1.7) *I ate something that was the result of what Bill acknowledged to be a syntactic transformation.

The selection in these sentences is between *eat* and the final NP in the sentence, which is deeply embedded. To capture this at the stage of lexical insertion, as proposed in *Aspects*, would involve duplicating all the machinery of the semantic component. Sometimes the selection is based on semantic properties that can be identified in general only with an entire sentence, not with some particular formative. For example, it will be argued in Chapter 5 that *shout* in (1.8) requires the complement sentence to be something that *Bill* can cause to happen.

(1.8) I shouted to Bill for Harry to leave.

It has often been noticed that this selection imposes a constraint on possible verbs in the complement:

(1.9) *I shouted to Bill for Harry to $\begin{cases} \text{know the answer.} \\ \text{be tall.} \\ \text{have black hair.} \end{cases}$

What has been overlooked is that choice of complement subject is relevant. By changing the complement subject in (1.9) appropriately, the sentence is acceptable.

(1.10) I shouted to Bill for the next recruit to $\begin{cases} \text{know the answer.} \\ \text{be tall.} \\ \text{have black hair.} \end{cases}$

Thus the selection must be dependent on the reading of the entire sentence, not just the verb or subject. This is a further argument that it must be performed on completed semantic readings.

   Finally, violation of selection restrictions can occur either on the basis of knowledge of the language or on knowledge of the real world. (1.11), if uttered while pointing to a man, seems to be the same sort of violation as (1.12), even though (1.11) depends on facts external to the language and (1.12) does not.

(1.11)   That person over there is pregnant.
(1.12) *That man over there is pregnant.

And as with the analytic-synthetic distinction, it is impossible to tell where linguistic knowledge leaves off and extralinguistic knowledge takes over.

(1.13) is presumably ruled out on the basis of one's knowledge of the language (although even this is open to question). On the other hand, the very similar (1.14) can be ruled out only on the basis of a mathematical theorem.

(1.13) Irving drew a circular square.
(1.14) Irving constructed a five-sided regular polyhedron.

The only level of derivation at which linguistic and extralinguistic facts can be brought to bear on sentences in identical fashion, as appears necessary in (1.11)–(1.14), is the level of semantic representation. Thus the most general solution to the problem of selection seems to be a well-formedness condition on semantic representation.

Another well-formedness condition will be used extensively in Chapters 4 and 5: the Consistency Condition on coreferents. This condition states simply that if two noun phrases are marked coreferential by the grammar, they must in fact be able to represent the same individual. (1.15) is an obvious case of its application.

(1.15) *The old man saw herself.

The only possible antecedent for *herself* is *the old man*; we will assume that the semantic component actually produces this interpretation. But since the two noun phrases differ in gender, they cannot denote the same individual, so the Consistency Condition rules out this interpretation of the sentence. Since it is the only possible interpretation, the sentence is rejected as unacceptable.

By adopting well-formedness conditions on interpretations, it will often be possible to avoid complex constraints on interactions between various rules. In this way the rules of the grammar will be able to apply freely, producing readings without regard to their acceptability. For example, in Chapter 3 we will show that the rules generating and moving adverbs, which have generally been treated as extremely idiosyncratic, can be made perfectly general by subjecting the resulting semantic interpretations to independently motivated well-formedness conditions: if an adverb occurs in an incorrect position, it will be integrated into the interpretation of the sentence in a way incompatible with its possible range of meanings. Another well-formedness condition, the Thematic Hierarchy Condition developed

in Chapters 2 and 4, explains a number of exceptions to the passive and replaces complex constraints on the application of transformations (the Crossover Condition). The interaction of two independent well-formedness conditions developed in Chapter 5 constrains the selection of coreferents for complement subjects in a way that eliminates a complex system of constraints on transformations. The result in each of these cases is that the transformations can be permitted to apply without constraint because the resulting semantic interpretation can readily be ruled out. Thus the rules already necessary for semantic interpretation can replace constraints whose only motivation is the need to rule sentences out.

1.7 Assumptions about the Syntax and the Lexicon

We will assume a base component of the usual form, a context-free phrase-structure grammar whose initial symbol is S. However, there are two important differences between the base component we will use here and that of *Aspects of the Theory of Syntax*.

The first difference concerns the process of lexical insertion. In *Aspects*, category nodes such as N and V are expanded into *complex symbols* which express subcategorization and selectional restrictions. Then lexical items are inserted by context-sensitive rules which prevent deep structures from violating these restrictions. Inasmuch as we have argued that these restrictions are properly implemented in the semantic component, a simpler lexical insertion process is possible. We will assume that lexical insertion rules insert lexical items freely under category symbols, eliminating the notion *complex symbol* altogether. In other words, we will assume that *Colorless green ideas sleep furiously* does have a well-formed deep structure.

The second difference concerns the use of node symbols. One of the consequences of the Lexicalist Hypothesis of Chomsky (1970a) is that syntactic nodes are to be represented as matrices of distinctive features. This change in the conception of syntactic nodes is parallel to the replacement of the IPA alphabet by matrices of distinctive features in phonological theory; the arguments are of similar form, based on cross-classification of categories with respect to rules of the grammar. We will make particular use of the concept of syntactic nodes as distinctive feature matrices in dealing with the analysis of adverbs and reflexives in Chapters 3 and 4 respectively. Hence we will assume that the base component and all other rules which refer to phrase-markers have the distinctive feature mechanism

available as a means of capturing generalizations; such symbols as *NP* and *VP* will be considered as abbreviations of feature matrices, just as *ü* and *č* are treated as abbreviations of feature matrices in phonology.

Aside from the introduction of syntactic distinctive features, the transformational component will be assumed to be substantially the same as proposed in *Aspects*. Transformations will be assumed to be strictly ordered and applied according to the principle of the transformational cycle: the complete sequence of transformations is applied to the most deeply embedded *S*, then repeated, each time applying to the next most deeply embedded *S*. In Chapter 4 we will show that at least some NP nodes govern a cycle as well.

Finally, a little attention must be given to the internal structure of the lexicon. Under the Extended Lexical Hypothesis, transformations cannot perform derivational morphology. How then can we capture the semi-productivity of morphological processes? At the time of Lees's *Grammar of English Nominalizations* (1960) there was no possibility but a transformational solution, since the concept of a lexicon had not been proposed. But as with semantics, the potentialities of a lexicon were not explored even after the theoretical framework was available. Thus at present most well-known proposals about derivational morphology are couched in transformational formalisms, for example, G. Lakoff (1971) and Chapin (1967). One notable exception is Gruber (1967b).

Such solutions are not available to us. Rather, it is necessary to list, for example, both a verb and its nominalization in the lexicon. To capture the relation between them, there must be a way to express the fact that there is less independent information in a pair of lexical items consisting of a verb and its nominalization than in a pair consisting of a random verb and noun. One way to capture this redundancy is to consider the measure of information a simple counting of features, but to eliminate all or some of the features of the nominalization to capture the generality; this method, however, is not consistent with our assumption that both the verb and the nominalization are fully specified in the lexicon. Alternatively, one could propose that the regularities are expressed within the measure of information itself, as *redundancy rules* that say that certain shared features of the nominalization do not count as independent information. Such a solution will be assumed here.

We will suppose, then, that part of the lexicon (or information measure for the lexicon) is a set of morphological and semantic redundancy rules, parallel in function to the morpheme structure rules, that specify inde-

pendent phonological information content. These redundancy rules will enable us to express the concept "separate but related lexical items" without the use of transformations. Jackendoff (to appear) explores this assumption in some detail, with application to a wide range of word-formation processes; it is shown that these redundancy rules make somewhat different and more satisfactory predictions than transformations in a number of cases.

1.8 A Remark on Motivating Rules

In a theory of grammar that minimizes the power of the semantic component, the base, and the lexicon, such as the theory of generative semantics, there is only one way in which similarity in meaning or co-occurrence restrictions between two constructions can be captured: a transformation. In a theory permitting a number of different kinds of rules, such as the theory to be explored here, there are many ways of capturing generalizations. In addition to transformations, there are all the different kinds of semantic rules operating at different levels of the derivation. Generalizations can also be captured within the lexicon, by means of the redundancy rules mentioned in section 1.7. Furthermore, certain generalizations can be expressed by treating the nodes for lexical categories as feature complexes, then stating base rules, transformations, and semantic rules so as to refer to more than one major category at a time.

With all these different kinds of rules at our disposal, several very different analyses will often come to mind for the same phenomenon, each of which seems equally capable of expressing the proper generalization. How do we decide which account is to be preferred? There can be no sort of principle that says, "Always choose an X rule if you have a chance": it is not difficult to construct algorithms to reduce all rules to any chosen type, given exception machinery of sufficient power, such as G. Lakoff's (1971). Rather the decision will be made on the basis of how the rules interact with each other most naturally and how appropriate the power already proposed for a particular type of rule is for handling something new. In general, similar processes should be handled by similar kinds of rules, to limit the total power of the theory.

Also of prime importance in motivating a particular treatment of a phenomenon is how it is reflected in the lexicon. If a process takes place only for certain lexical items, or varies over several classes of lexical items, we should choose the way out of handling the process that least increases the independent information content of the lexicon. The use of exception

features is the worst possible solution, in that it represents an arbitrary bifurcation of the lexicon, and so every marked feature represents independent information. Interpreted another way, the use of exception features makes the claim that each exceptional lexical item must be learned individually. On the other hand, if the difference in grammatical behavior has something to do with the meaning of the items in question, then that is the best possible case, since the rule has only to refer to the properties already present—if the meaning of the item is learned, its behavior is known automatically.

Unfortunately, this latter case is also the least formalizable, since we often do not have a principled way of expressing the meaning. For the sake of stating a rule, however, it seems to me perfectly adequate to provisionally adopt an arbitrary feature, if we have clear intuitions about when this feature is present, and if it is fully understood that it has no life independent of the complete reading in which it is embedded.

Our investigation will be organized as follows: Chapter 2 introduces a formalism for expressing functional structure and defends it with respect to other current proposals. Chapter 3 is a detailed investigation of adverbs and adverbial phrases, a demonstration of the efficacy of a theory employing projection rules and syntactic distinctive features to capture generalizations. Chapters 4 and 5 are concerned with coreference. The former develops an interpretive theory of pronouns and reflexives; the latter extends this analysis to the deleted complement subject and explores its extensive consequences on the complement system as a whole. Chapter 6 discusses focus and presupposition, with a systematic semantic analysis of some intonation contours. Chapter 7 introduces the modal structure as a representation of specificity and extends it to several other phenomena. Chapter 8 is a detailed study of negation and its interaction with modal structure and focus and presupposition. Chapters 9 and 10 present conclusions and consequences of the proposed semantic theory for the transformational component.