A Semantic Theory of the English Auxiliary System
by
Erich Friedrich Woisetschlaeger

Submitted to the Department of Linguistics and Philosophy on October 29, 1976, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

In this thesis, a detailed analysis of the semantics of the English auxiliary system is worked out. Specific hypotheses are formulated for the relation of syntactic structures to semantic systems dealing with Tense, Mood, and Aspect. In addition to the formal principles set forth, some substantive semantic universals are proposed in the form of a small number of non-logical concepts used in specifying the meaning of the elements of the English auxiliary system.

Chapter 1 deals with the system of formal principles involved; among them are the Univocality Assumption and the Complement Convention, both of which constitute a departure from most traditional work in this area.

Chapter 2 presents an analysis of the progressive form of English. Two syntactic progressives are distinguished, one mapping into the domain of Aspect, and one mapping into a domain called Epistemic Status. The two categories of Epistemic Status, the phenomenal and the structural, are carefully characterized.

Chapter 3 contains an analysis of the English perfect in terms of Jakobson's category of Taxis. A new solution is advanced to the familiar problem of the interaction of the present perfect with specific past time adverbials.

Tense is analyzed within the conventional model of linear time in Chapter 4. Much detailed discussion is focussed on the claim that 'the meaning of English present tense is 'simultaneity with the time of the speech event'.

Chapter 5 examines the conceptual structure of modality in English, and a very simple, self-contained core system covering both root and epistemic modals is presented. Some interesting interactions of modality with other verbal categories are analyzed.

Thesis Supervisor: Noam Chomsky
Title: Institute Professor
ACKNOWLEDGMENTS

I am fortunate to be able to write this page at some distance from the life of a linguistics graduate student at MIT. Fortunate, because only from a little distance can one really appreciate the uniqueness of the community of scholars and the spirit of inquiry that Morris Halle and his colleagues have built and maintained there over the years. For this, for providing an atmosphere conducive to creative, and at the same time rigorous, thinking, I am thankful above all.

There are more specific debts of gratitude: To Noam Chomsky, Paul Kiparsky, and Haj Ross, the members of my thesis committee, for their support, and for their patience. The finished product is far from being a masterpiece; but much of what is solid about it they have helped make so.

To Morris Halle for a demonstration of what to aspire to in terms of inspired and inspiring teaching. For his caring, and his encouragement.

To Ken Hale for a glimpse of what being a mature linguist might be like.

They have all set standards of professional excellence that will provide a continuous challenge in my own work.

Next, I'd like to say "Thanks!" to my fellow graduate students. As fine a group of people as you'd ever want to meet.

A special handshake, with lots of feeling, goes to John Goldsmith though. Except that he wouldn't agree with everything that's in this thesis, his name might appear on the title page as a co-author. The many hours of discussion and exploration I spent with him are among the purest specimens of joyful inquiry I have experienced as a linguist.
Finally, this is surely a fitting place to thank Fran for her efficiency (mainly because one can hardly thank someone for their personality), and to thank Varia for typing most of the thesis, and for much more.
TABLE OF CONTENTS

Chapter 1: Introduction to the Verbal Categories .... 6

Chapter 2: The Progressive .............................. 28

Chapter 3: The Perfect ................................. 124

Chapter 4: Tense ........................................ 158

Chapter 5: The Modals .................................. 181

Footnotes .................................................. 207

Bibliography ............................................. 210

Biographical Note ................................. 214
CHAPTER 1.

INTRODUCTION TO THE VEREAL CATEGORIES

1.1. The general problem.

The first goal of this thesis is to provide a semantic analysis of the English auxiliary system. The question to which an answer will be sought is simple: What does the syntactic formula Tense - (Modal) - (have + en) - (be + ing) mean? What do its elements mean?

The most intriguing result of reflection upon this question is the discovery that it is not as simple and innocuous a question as it may seem to be at first blush. The problem that confronts us is not merely that it is difficult to find the right answer; it is more elementary than that. We first need to understand the question. We will understand it just to the extent that we know what would count as an answer; and there is little uniformity in the form of answers framed in current linguistic semantic theories.

Thus we cannot approach our immediate goal without first concerning ourselves with the larger issue of what kind of task it is. Let us then fix the conceptual framework within which we will pursue our goal as that of generative grammar. This means that we take our task to be the explicit description of the ideal speaker-hearer's intrinsic competence. Part of this intrinsic competence is the ability to assign a meaning to a potentially infinite class of sentences. And here lies the second goal of this work: to contribute to the answer to the question,
"How can this ability be described explicitly?"

Plainly, we will need the tools that any description requires. First, we need a vocabulary, and, secondly, we need a way to use that vocabulary. This corresponds roughly to the questions: How do you talk about meaning? and What can you say about meaning? We could illustrate this point by looking at its analogue in the theory of the base. The part of syntactic theory that is concerned with the vocabulary of the base is the theory of categories, and the part of syntactic theory that is concerned with the use of this vocabulary in the base is the theory of phrase structure.

When it comes to semantics, how do we choose a vocabulary, and how do we choose formal operations that employ this vocabulary? We will sidestep these questions for a moment while we stress that we must bear in mind that neither the choice of the vocabulary nor the choice of the operations on that vocabulary are arbitrary. Both are empirical issues, and there must, therefore, be true or false answers, not just more or less elegant ones. A second, and closely related, point to bear in mind is that these are universal questions. That is, they are questions about human language in general and not questions about any particular language. Thus the vocabulary available to any linguistic system is universally predetermined, although individual languages have some freedom in making selections from what is available.

Having duly acknowledged the empirical and universal nature of the task, we come back to the question of how to approach it. We will regard as a general guide for the purposes of this study the answer Chomsky (1965, p.18)
gave: "...the actual data of linguistic performance will provide much evidence for determining the correctness of hypotheses about underlying linguistic structure, along with introspective reports (by the native speaker, or the linguist who has learned the language )." This answer is more remarkable for what it leaves unsaid than for what it says, for it says nothing about there being any reliable technique for collecting data that bear in well-defined ways on questions of linguistic structure. We are thus faced from the start with the task of deciding how we are going to extract meaningful data in our particular domain. And we must delimit just what our domain is.

1.2. Locating the semantic domain.

The traditional labels for the domain we will be dealing with are Tense, Mood and Aspect. These three semantic categories are often regarded as a natural class, and it is to this intuition that we will turn first.

The verbal categories of Tense, Mood and Aspect form a natural class in the sense that they all presuppose a semantically complete specification of the 'bare propositional content' of the sentence. More precisely, the picture is this: Assume that there is a semantic subcomponent governing Tense, Mood, and Aspect (henceforth 'TMA component'). Whatever its internal structure, it will receive some kind of semantic object as its input, and yield a suitably modified kind of semantic object as its output. This input is a bare propositional content. The output is the end product of the semantic component as a whole; an assertion, a question, an exclamation, a command.
The notion of 'bare propositional content' is thus important. It refers to a structure which is built by the subcomponents of the semantic component other than TMA. Intuitively, this structure amounts to a full description of an event or a state of affairs, but without this description being related to the world. An analogy may be helpful here: We may compare the complete specification of the bare propositional content of a sentence to the complete specification of a coin, taken simply as an object (its weight, size, shape, the metal it is made from, the head and tail imprints, the nominal value, etc.). The failure to relate this bare propositional content to the world is analogous to not saying anything about whether or not the coin we have described is, was, or will be legal tender anywhere, nor saying anything about its purchasing power.

What enters into this specification of the bare propositional content? We give a list below of three broad semantic systems which we will take as exhaustive of propositional content. This sketch of the semantic component is not definitive. It does, however, reflect certain recent trends within generative grammar in the direction of investing the semantic component with some internal structure. For our purposes here it helps to make more concrete than does the coin analogy, what all goes into propositional content. The three systems are:

1. **Thematic Relations.** This semantic system deals with the verb and its arguments and is a theory of what relational patterns a verb can define on ordered n-tuples of noun phrases. The thematic roles of AGENT, THEME, SOURCE, GOAL, LOCATION, and
INSTRUMENT are part of the vocabulary of the system, and they enter into certain configurations as arguments of a limited set of abstract predicates like GO, STAY, BE, CAUSE, LET, etc. These abstract configurations define a set of basic conceptualizations of events and states, constitute the basis for certain limited kinds of inferences, and are held to remain constant across diverse ontological domains which include, e.g., POSITIONAL, POSSESSIONAL, and IDENTIFICATIONAL.

We will illustrate very briefly. Consider the thematic frame (1).

(1) GO (THEME, SOURCE, GOAL)

This frame represents the concept of an entity - the THEME - moving along a path, whose starting point is called the SOURCE and whose end point is called the GOAL. An inference schema based on (1) would, for example, say that if the THEME is at the GOAL at time $t_1$, then the THEME will have been at the SOURCE at some time $t_2$ which precedes $t_1$. The invariance of the frame across domains may be seen from examples (2)-(4), with the use of the verb go being positional in (2), possessional in (3) and identificational in (4).

(2) Max went from Paris to Brussels on a one-speed bicycle.

(3) The Picasso went to a Pittsburgh steel magnate.

(4) Soon after his conversion, Joseph went from being the life of every party to being an irritating moralist.

(2) ascribes to Max a change in geographical location. (3) ascribes
to the painting a change in ownership. (4) ascribes to Joseph
a change in identifying properties. Yet all are instances of the
frame \( x \textit{ went} (\textit{from} y) \textit{ to} z \). The theory of Thematic Relations
deals with the non-accidental nature of such convergences in
surface forms. Much of the work on this approach to the seman-
tics of the verb has been carried out by Gruber and Jackendoff.

2. **Lexical Semantics.** This part of semantics yields a
specification of word meanings. The study of semantic features
(cf. Katz (1966), Bierwisch (1967)) belongs here, as does the
study of semantic fields.

3. **Logical Form.** This part of semantics deals with a
variety of semantic structures, e.g. with the relation of articles
and degree adverbs to relative clauses and degree clauses,
respectively; relations of coreference, non-coreference, disjoint
reference, reciprocity, etc.; the interpretation of quantifiers
and of forms bound by them; the definition of the subject/predicate
structure on sentences; certain aspects of information structure
such as semantic focus and presupposition. (On Logical Form as a
linguistic level, see Chomsky (1975), Sag (1976).)

So much for an internal characterization of propositional content
at this point. Next we ask: What is its logical status? Is it a term,
or a predicate, or a sentence? It is none of these. The closest parallel
to it in the domain of nominal expressions is common nouns. The noun \textit{dog}
is neither a term, nor a predicate, and certainly it is not a sentence.
It has no independent logical status. Similarly, it can be said that within standard predicate calculus, propositional content has no logical status.

But if the output of the TMA component includes assertions - as we have said it does - which do have logical status as expressions that can be either true or false, then (given the fact that propositional content is a source of truth conditions) there must be mechanisms in the TMA component that construct logical structures. We will not give a full formalization of the entire TMA component, however. There are still too many issues that we know too little about.

We chose the phrase, 'Locating the semantic domain,' as the heading of this section. We have located the domain only very indirectly - by indicating its status relative to other domains and by giving a very compressed list of things that are not in this one. This will do for the moment. We will discuss particular categories that belong into the TMA component only to the extent that they figure in English; in this way any substantive characterization we propose for a given category will be supported (and could be challenged) by the range of data that is relevant to it.

1.3. **The Univocality Assumption.**

In our analysis of the English auxiliary we will adopt a methodological principle which we call the Univocality Assumption, and which we state in (5).

(5) **The Univocality Assumption**

Given a syntactic constituent, there is exactly one
semantic expression that specifies its meaning.

We need not bother to make (5) precise: the intent of (5) is clear enough for us to see that, as a factual statement, (5) is simply false. A look at any polysemous lexical item such as plane ("vehicle that flies"/"woodworking tool"/"level surface") shows that as long as we count each expression in quotes as a distinct semantic expression (5) fails for such cases. If, on the other hand, we treat (5) merely as a statement expressing the ideal state of affairs which we should try to approach in analysis, it appears to be nothing more than a guide for any good descriptive practice.

Now if the Univocality Assumption has only this very general force it should not be emphasized here. There is, however, reason to think that this emphasis is not misplaced.

First of all, we must bear in mind that the verbal categories are typically represented syntactically by grammaticalized forms. The property of being grammaticalized is not an 'either-or', but a 'more-or-less' one, and as a general rule we can say that a form is more highly grammaticalized the smaller the closed set of forms is to which it belongs. By this measure, the forms of the English auxiliary are highly grammaticalized.

Secondly, the following is a plausible principle:

(6) The more highly grammaticalized a morpheme is, the less it will tolerate polysemy.

Recall our premise that our ultimate goal in constructing a linguistic
theory is to explain how the grammar of particular languages can be learned. That is, we are trying to account for the fact that on the basis of limited exposure to primary data a language learner is able to construct a grammar that characterizes an infinite set of sentences. If we narrow this down to the specific question of how a semantic interpretation is assigned to the morphemes of the language it stands to reason that the most highly structured, and therefore optimal, linguistic system would be one in which to each distinct form a distinct meaning is assigned, or at least a distinct basic meaning, with some set of regular processes leading to derived meanings. Considerations of economy are presumably pitted against considerations of clarity to some extent, and they exert pressure in the direction of assigning a family of meanings to a form.

True polysemy is a different matter, however, where two or more entirely unrelated meanings are assigned to the same overt form. Nevertheless, as long as this happens in the major lexical form classes, e.g. nouns, the polysemy may be innocuous from the point of view of learning the different meanings of a polysemous form. Whatever the exact mechanisms are of attaining adequate dictionary entries, it is certainly a truism that extralinguistic and linguistic clues play a significant role. Without subscribing to any simplistic model of the acquisition of such dictionary entries, we may feel confident in saying that, for example, the entry for ball in the sense of spherical object, and the one for ball in the sense of a festive dance, will be acquired in sufficiently disparate extralinguistic, and linguistic, contexts that no problems are likely to arise.
The picture changes when we consider grammaticalized forms, members of closed sets of forms. The difference between these morphemes and lexical morphemes is that these grammaticalized morphemes are generally not related to any constant features of the extralinguistic or linguistic context. We have made the point in our discussion of the verbal categories (see section 1.2) that they do not deal with any features of the propositional content of a sentence, but deal rather with semantic features that are few in number, hence very broad, and that can co-occur with any propositional content whatever.

On the face of it then, these are extremely difficult forms from the point of view of learning their meaning. On the other hand, they are central forms from the point of view of syntactic well-formedness. They are either obligatory, or at least available, for every single utterance of a sentence of the language. Words whose meanings one is not sure of one may simply never use until one is clearer about them. With the grammaticalized morphemes this option is not available. Without them one cannot form well-formed sentences at all. Given this, there should be a very strong tendency in all languages to facilitate the semantic acquisition of these forms. As the child acquires the syntactic structure of the system of verbal categories, there ought to be a sharp delimitation of the range of accessible semantic systems that such a syntactic system may reflect. The more polysemy can be eliminated as a complicating factor, the more feasible this sharp delimitation will be. For if polysemy is a factor to contend with, which is to say that the learner must make allowance for the possibility that he may in fact be dealing
with more distinct morphemes than he can readily recognize, the learner cannot be sure just how rich a system of forms he is dealing with. And presumably, different degrees of elaboration in the system of syntactic forms may make rather different semantic systems accessible.

All of this is intended as a plausibility argument. From it we derive a theoretical bias: the forms which express the verbal categories will be unambiguous. It should be emphasized that, while the argument we have given about the theoretical costliness of ambiguity in the verbal categories is not conclusive, the consequence we have derived from it in terms of a research strategy is an extremely powerful one. We will be led directly to reject certain semantic analyses which, on different theoretical assumptions, would be entirely acceptable.

A case in point is an analysis of the English progressive given by Leech (1969), which is based on a conception of linguistic description as being primarily the perspicuous and systematic tabulation of linguistic facts. Within this kind of framework, there is nothing anomalous about saying that certain uses of the progressive signify limited duration, while others serve to throw emphasis on the ceaseless persistence of some process. These two uses are mutually contradictory, and could thus not both be derived from some unambiguous meaning of the progressive.

Similarly, McCawley's (1971) description of the present perfect in English as involving either universal quantification over an interval of time stretching from the past into the present, or existential quantification over such an interval, or an indication of the direct
effects of a past event still continuing, or, finally, the conveyance of hot news, is suspect. This is so not because the different uses are contradictory, which they are not, but because they are a somewhat heterogeneous collection. If both "indication of the existence of a past event", and "indication of the continuance of the direct effects of a past event" are to be separate senses of the same form, then these different senses really deal with non-comparable concepts. The Univocality Assumption, strengthened by principle (6), militates against such an analysis.

1.4. The Complement Convention.

A major assumption we will be making in this work is that the organization of the verbal categories is binary. In other words, the distinctions made within any verbal category fall into pairs of oppositions. In each such pair, there will be one member representing the marked value of the distinction, and one member representing the unmarked value of the distinction. The simplest case is that of a single non-null form contrasting with a null form; the non-null form naturally represents the marked value, the null form the unmarked value. In what follows, we will have this case in mind.

The relationship of marked to unmarked form will be as follows: the marked value will receive a positive semantic characterization. That is, the marked value will trigger a translation rule mapping into a semantic representation. Let us consider what such a rule tells us. Suppose it is a rule for future tense, and it tells us that the situation represented
by the propositional content of the sentence is asserted to prevail at a time later than that of the speech event. In addition to the specific temporal assignment made, however, the rule also tells us something else. It tells us what arena we are in: The theoretical construct invoked is a representation of time as a line, with the indexical constant of 'time of the speech event' defined on it. Moreover, it tells us that the kind of information the tense rule supplies is where on the time line a situation is located with respect to the speech event.

Assume now that Universal Grammar contains a rich enough semantic characterization of every verbal category so that what we have just described is generally true. That is, the form of a semantic translation rule belonging to the TMA component allows us to pick out exactly which semantic domain is addressed by the rule. There is a range of logically conceivable alternatives for making this suggestion more precise. We give a succinct statement of the strongest position compatible with our analysis of English in (7) below.

But first we need to make explicit the following subsidiary hypothesis: semantic translation rules have an additive effect. This statement, though concise, is rather obscure. What is meant is that the input expression on which a translation rule operates, appears intact in the output expression, with some symbols of the semantic language added on in some way. Schematically, we get the following picture. Let I be the input expression, O the output expression, c the verbal category to be translated, and c(I)→ O the canonical form of the translation rules; then, or so we hypothesize, O will be of the form XIV, where X and Y are
variables over strings of symbols of the semantic language. Given an input expression \( I \), then, \( AIB \) is a possible output expression, but \( AB \), where neither \( A \) nor \( B \) nor \( AB \) includes \( I \), is not. Given a rule \( M \) translating the marked value of a verbal category, and an output expression \( O \) analyzable as \( XIV \), we will call \( X...Y \) the augment of \( M \), and rewrite \( XIV \) as \( A(I) \), where \( A \) stands for the augment of \( M \). We are now ready to state the Domain Determinacy Hypothesis:

(7) The Domain Determinacy Hypothesis

Given a translation rule \( M \) for the marked value of a verbal category, an input expression \( I \), and an output expression \( A(I) \), \( A \) will contain semantic primitives, or semantic expressions, that determine which unique closed semantic domain \( M \) maps into.

We have already given an example of such a closed semantic domain. By 'closed' we mean here that it must be fixed what counts as exhausting the domain. In the case of Tense discussed above it is easy to specify what it would mean to exhaust the domain: if we add to the meaning 'later than the speech event', which we proposed for Future Tense, the meaning 'overlapping with, or earlier than, the speech event', we have covered every point on the time line. In more general terms we can say that the domain of Tense is exhausted if and only if every point on the time line is within the scope of at least one tense form.

For the kind of simple, two-valued system we are now chiefly addressing ourselves to, the Complement Convention shortly to be stated ensures that we not only know what would count as exhausting a given domain, but also that the domain will in fact be exhausted. This is not to be regarded as an accidental property of two-valued systems. Rather, it is a reflection of the general
condition on syntactic systems of the TMA component that they must exhaust the semantic domain associated with them. We might call this general principle the Domain Exhaustion Principle. (For an illustration of its use, see 2.3.2).

With this much said, we can now go on to state the Complement Convention:

(8) The Complement Convention

Given A as the meaning of the marked value of a binary opposition, the meaning of the unmarked value is the complement of A in the semantic domain to which A belongs.

The Complement Convention is our hypothesis about what the meaning of an unmarked form is. It does not say anything about how that meaning gets assigned.

The question is: Should we postulate a translation rule for the unmarked value, complementing the rule for the marked value. Notice that if there were such a rule, its form would be completely predictable, under the assumptions we have made. But what is a rule of grammar? It is the statement of a language-particular fact that is permitted, but not required by universal grammar. In other words, there are no rules of grammar of a particular language that are entirely predictable on the basis of linguistic theory. In particular, then, the fact that the unmarked value of a syntactically binary distinction is the complement of the meaning of the marked value, by virtue of being a universal, will not be expressed by a rule of the grammar.

It remains for us to make a proposal as to the specific mechanism that governs the assignment of this universally predetermined meaning of unmarked forms to sentences. We propose the following well-formedness
condition for semantic representations:

(9) The Semantic Closure Principle

If a grammar contains a semantic rule that maps into a semantic domain $D$, then all well-formed sentences are specified with respect to that domain $D$.

To be specified with respect to a semantic domain $D$ is simply the opposite of being vague with respect to $D$. If a form is specified with respect to $D$, there must be at least one meaning belonging to $D$ that the form in question does not have. The Semantic Closure Principle ensures that the Complement Convention, which merely determines a meaning and leaves the assignment of that meaning open, actually does enter into the interpretation of sentences. Specific consequences of the Semantic Closure Principle will be discussed later in the discussion, as such questions arise.

One consequence, however, is worth pointing out immediately. The position developed here does not allow systems with unmarked forms whose meaning includes that of the corresponding marked form. On general methodological grounds this is a desirable state of affairs since it means that we have built enough structure to rule out a class of logically entirely unproblematic systems. But there may exist factual problems. If Comrie (1976), e.g., is right in saying that "...in many cases the meaning of the unmarked category can encompass that of its marked counterpart," then (9) is plainly false. And Comrie is not alone. Jakobson (1971) takes an even stronger position on 'correlative distinctions'
If the marked form indicates the presence of P, then the unmarked form does not indicate the presence of P. If, in a given context, the unmarked form is taken to indicate the absence of P then this is merely one of the possibilities of interpretation provided for by the non-indication of P.

A case in point, discussed by Comrie, is the Italian phrase sto scrivendo ("I am writing") as opposed to scrivo ("I write"). As Comrie correctly observes, the first of these forms is very similar in meaning to that of the English Progressive form I am writing, but in Italian, as distinct from English, the Progressive form can always, without excluding progressive meaning, be replaced by the non-Progressive form (in this case scrivo). Comrie's conclusion, which by now has a familiar ring to it, reads as follows: "In such cases, we may say quite strictly that the marked category signals the presence of some feature, while the unmarked category simply says nothing about its presence or absence."

We have an interest in maintaining the Semantic Closure Principle, since we intend to use it in important ways, and we grant that there are apparent counterexamples. This is an uncomfortable position to be in, and so we proceed to give at least an outline of a resolution. We start with the following innocuous premise: not everything that looks like a correlative distinction is one, in fact. Recall that we very cautiously decided, at the beginning of this section, to limit our attention to cases where a single non-null form contrasts with a null form. In the
English auxiliary, have + _en and be + _ing are of this type. Under all analyses that treat these two forms as auxiliaries, they each form a class of one. This is the paradigm case in which the absence of a particular morphological form can be independently meaning-bearing.

Contrast this case with the opposite extreme, the case of a garden variety transitive verb, say _hug_. Since the position that _hug_ is inserted into, that of the main verb of the sentence, is filled from an open word class, the choice of not inserting _hug_ lacks any capacity of conveying an independent meaning. We hypothesize: the more highly grammaticalized a form is, the more likely it is to enter into a system of binary oppositions.

Suppose now that it were to turn out that the _sto_ of the Italian Progressive form _sto scrivendo_ is best analysed as a main verb. Then the fact that _sto scrivendo_ is related to the non-Progressive _scrivo_ in the non-privative fashion described by Jakobson and Comrie (i.e. the meaning of one form is part of the meaning of the other) would not be so surprising any more. If _sto_ is a main verb, it will be given a lexical meaning, it will not trigger any rules of the TMA component, and it will also not bear on the markedness issue.

1.5. An Ordering Principle?

The last question to be addressed in this introduction is this: is there any internal organization to the TMA component? In particular, are its rules ordered with respect to each other, and if so, how is the relative order of the rules determined?

We will present here the strongest position we feel justified in
formulating. Our justification ranges in strength from solid argument all the way to lack of obvious counterevidence. There are two parts to our position on this issue. The first concerns the matter of how the relative order of the TMA rules is determined, thus presupposing an affirmative decision on whether there is any ordering. This part is stated in (10):

(10) **The TMA Ordering Principle**

If a and b are each syntactic formatives representing a verbal category, and a is 'closer' to the verb stem than b, then the translation rule for a precedes that for b.

The notion of 'closeness to the stem' needs to be clarified. We will consider a to be closer to the stem than b if the path from a to the verb stem is shorter (in terms of intervening nodes) than that from b. If, by this criterion, a and b come out equally close, then closeness in terms of linear order counts. This will decide in all the cases we will encounter, even though it leaves one case unsettled. What (10) tells us, for example, is that both the Progressive and the Perfect are to be translated before the Modals.

The careful reader will have observed that the TMA Ordering Principle is a particular instance of the more general principle, subscribed to by Montague grammarians, that there is a one-one correspondence between syntactic and semantic rules. The truth of the general principle is by no means self-evident; the principle has in fact been argued to be false (cf. Chomsky (1975)). We take Chomsky to have shown
that the Montague principle is not right in all cases; but (10) amounts to claiming that it is not wrong in all cases, either. In any event, (10) is not deduced from the Montague principle.

The second part of our position on rule ordering limits the possible orderings among rules. We present it in (11):

(11) **The Verbal Hierarchy Hypothesis**

Let a and b each be verbal categories. Then if A and B are the semantic domains associated with them, universal grammar fixes the scope relationship between A and B.

In practical terms this means that if there is a rule R mapping into A, and a rule $R^1$ mapping into B, universal grammar determines which of them has to apply first. Assume R has to apply before $R^1$. This has further consequences, for only a syntactic structure in which a is closer to the verb stem than b is permitted. To see this, suppose it to be otherwise, with b being closer to the verb stem than a. By the TMA Ordering Principle, the translation rule for b, $R^1$, precedes the one for a, R. But, by assumption, R precedes $R^1$. Thus we get a contradiction. But we assume Universal Grammar to be internally consistent. We conclude that the case considered is not possible in a natural language.

The bearing that (10) and (11) together have on the learnability of systems of verbal categories is immediately apparent; one source of complexity, the lack of any consistent relationship between syntactic forms and semantic domains, is eliminated by them. Or, to put the matter in more positive terms: Given a successful syntactic analysis, the set
of accessible semantic analyses is restricted, and thus the correct analysis is more readily attainable. For example: (10) alone tells us to interpret the Progressive before the Perfect because the Progressive is closer to the verb. But (11) tells us much more, for it tells us this about have\textit{en} and be\textit{ing} as well: given any two semantic domains $D_1$ and $D_2$ that we may wish to consider as candidates for supplying the 'perfect meaning' and the 'progressive meaning', respectively, only one hypothesis about syntactic form/semantic domain-pairs is admissible. That is, given the Verbal Hierarchy Hypothesis, the two hypotheses roughly indicated by (12 i) and (12 ii)

(12) i. $(\text{have}\text{en}, D_1)$  
$(\text{be}\text{ing}, D_2)$  
ii. $(\text{have}\text{en}, D_2)$  
$(\text{be}\text{ing}, D_1)$

are not equally valued; one of them is simply not available.

In our analysis of English we will follow (10). In so doing, we will map syntactic formatives into semantic domains. The order in which we address these semantic domains is hypothesized by (11) to be fixed. (11) is extremely vulnerable, but there is some evidence that it is correct (perhaps only partially correct, if the relative order of only some pairs of semantic domains is fixed). Our analysis of English, for example, will place Aspect before Tense. Now consider the following statement of Jakobson's about the Russian verbal categories:
The shifters (tense) employ initial desinential suffixes, while the non-shifters (aspect) go farther back; they ignore the desinence, and operate with the stem ... its suffixes and prefixation.

This, too, points to Aspect before Tense. Or take the following statement from Blansitt (1975), summarizing some results of a survey study:

If progressive is signalled by a suffix, any tense, mode, or person suffixes will follow the progressive suffix.

Thus, here too, the aspectual suffix is closer to the stem than the tense suffix.

This line of investigation is of considerable intrinsic interest. Observations such as those just reviewed are waiting to be addressed by a theory. This work is intended as a serious contribution towards such a theory.
2.1. Introduction.

In Chapter 1, our discussion was geared to a characterization of a number of general design features of systems of verbal categories. When we developed the argument in favor of the Univocality Assumption, or our ideas surrounding the Complement Convention, our thoughts about the ordering of translation rules, we were not primarily guided by compelling evidence. Rather, what we tried to do was to generate hypotheses about desirable properties of verbal systems on the basis of our conception of the task at hand, especially in the light of the dearth of work in this particular area. We viewed our task as that of working towards an explanation of the learnability of such systems. In this endeavor we made the basic assumption that all generative grammarians make, viz. that no linguistic subsystem may vary arbitrarily from language to language, and that it is the main goal of linguistic theory to impose rather narrow limits on variability. The more we restrict the range of attainable hypotheses about the mental representation of the system underlying linguistic data, the more closely we approximate an explanation of learnability.

Nevertheless, enough variability must be permitted by linguistic theory to account for the fact that languages are not all cut to exactly the same mold, and that individual languages change through time, which
comprehensive treatment of English phrase structure worked out along the lines of Chomsky's (1969) Base Schema Hypothesis, usually referred to as the X-Bar Convention. In particular, we accept Jackendoff's proposal that within each of the four lexical categories (nouns, verbs, adjectives, prepositions) three subcategories be distinguished: (i) a major category (marked by Jackendoff with the feature [+Comp]), (ii) a minor non-specifier category (Jackendoff's [-Comp, -Det]), and (iii) a minor specifier category (Jackendoff's [-Comp, +Det]). We will interpret this framework as follows (no longer necessarily in agreement with Jackendoff): To say that the progressive be is a main verb is to treat it as belonging to the major verb category; to say that it is an auxiliary verb is to treat it as belonging to one of the minor verb categories.

We can determine the status of progressive be as an auxiliary unequivocally if we assume that (i) there can only be one specifier attached to any major category, (ii) that specifiers attach at the triple-bar level (in Jackendoff's theory all categories participate in the rule schema $X^n \rightarrow ... X^{n-1} ...$, with $n \leq 3$), and (iii) that non-specifiers attach at the double-bar level. Now if two forms are attached at different levels within a given phrase, then, provided they are attached on the same side, and barring any reordering rules, their linear order is thereby fixed. In English, auxiliary verbs are to the left of the main verb; consequently, a form attached at a higher level than another would be further to the left. Consider now Chomsky's (1957) auxiliary formula, $\text{Aux} \rightarrow \text{Tns-(M)-(have + en)-(be+ing)}$. The progressive be is rightmost in
this formula. Thus we know that it is not attached higher than the other auxiliaries. But by (i) and (ii) above, it would have to be attached higher if it were a specifier. Conclusion: the progressive auxiliary is a non-specifier.

To sum up this rather abstract discussion, we give in (1) a partial tree showing what we have arrived at with regard to the syntactic status of the progressive be as an auxiliary:

(1)

We now turn to the alternative analysis of progressive be, as a main verb. We will follow Jackendoff, and propose the structure in (2):
The decision to call the complement to be a prepositional phrase is based on the fact that the participial complements to be do in fact alternate with PP-complements. We will illustrate with (3) and (4):

(3) a. Tim was **working**.
    
    b. Tim was **at work**.

(4) John was **assembling a model plane** when I last saw him,
    
    and he's probably still **at it**.

The step of going from $P^1$ to $V^1$ in (2) above also requires some comment. Jackendoff's theory provides a recategorization schema of the form
\( x^n \rightarrow \text{AFF} \ y^n \), which is intended to be very restricted in its application. This restrictiveness is guaranteed, to some extent, by the stipulation that recategorization is subject to the presence of an affix, a kind of purely grammatical formative that stands outside of the categorial system, and furthermore, that recategorization must involve two categories at the same phrasal level. This restrictiveness is of considerable interest to us. Recall that we speculated in section 4 of Chapter 1 that the cluster of ideas represented by the Complement Convention and the Semantic Closure Principle were intimately connected with the notion of grammaticalization. The larger the class of formatives that could be inserted into a given position in a syntactic frame, the less grammaticalized that class of formatives, and the less applicable the theoretical constructs of 1.4. And perhaps also: The larger the syntactic frame that must be considered to determine whether a class of formative can be inserted, the less grammaticalized that class of formatives.

By these rough criteria, the progressive be would be highly grammaticalized under either syntactic analysis; the case illustrated in (1) we might want to call a case of full grammaticalization, the class of items insertable into a position marked \([-\text{Comp}, -\text{Det}]\) being very small indeed (as we will see); while the case illustrated in (2) might be called a case of partial grammaticalization, the class of insertable items (some verbs of temporal aspect, we will claim) being somewhat larger and the portion of the syntactic tree that has to be considered in order to pick the appropriate class being larger also.

We will not argue for the correctness of (1) or (2) on syntactic
grounds. We will, however, argue that both analyses derive a measure of support from semantic considerations. As we proceed, the air of paradox seemingly inherent in this state of affairs will hopefully disappear.

2.3. Aspectual (Main Verb) BE.

2.3.1. The notion of a subevent.

Sentences like (5) and (6),

(5) Walter was filing the day's mail.
(6) Walter filed the day's mail.

are the kind of minimal pair that linguists ponder in order to discover what the meaning contrast between the English progressive and the English non-progressive is. Confronted with (5) and (6), they would generally agree that (6) implies that Walter got all of the day's mail filed away, while (5) leaves that matter open. And as far as any useful information about the contrast examined here is concerned, what more is there to say? Nothing, we submit.

But we have insisted throughout that knowing what the facts are is only a precondition, albeit an important one, of our enterprise. Our goal is to help explain how native speakers acquire command of such facts.

It is with this goal in mind that we now introduce two theoretical constructs, one dependent on the other. The more basic of the two we will call an 'event' here, but will rely on the fact that event is, after all, an English word, to convey the meaning of the term. There is only one property of events as we will view them, that may not be self-evident. We will take events to be characterized by the fact that they
progress from an initial state to a terminal state. Thus, there are no infinite events. But this seems to be in accord with the normal use of the English word event. Take something like the earth rotating around the sun: for all practical intents and purposes this is conceived of as going on and on, forever; and we don't call the rotation of the earth around the sun an event. We would call one rotation an event. Yet to talk about one rotation, we've had to fix a beginning point and an end point. Similarly, the normal use of the word event presupposes some change through time. There is no event whose initial state and terminal state, as well as all intermediate states, are all identical. That is just what distinguishes events from states.

Consider now the event of Walter filing the mail. Suppose there are ten letters to be filed. The initial state will be when there is a pile of 10 letters sitting on Walter's desk, and the terminal state when the pile of letters is gone and each letter has been tucked away in a file.

Now if Walter files the letters one by one, there will be 9 intermediate states; one where there are 9 letters left to file, one where there are 8, etc. If we called what happens between initial state and terminal state an event, we will call what happens between two states at least one of which is an intermediate state a subevent.

We guarantee that every event has subevents in the sense just described by introducing the following postulate:

(7) There are intermediate states in every event.
This postulate is quite independent of the fact that in many cases, say in the case of Walter filing just one letter, we may not be interested in intermediate states, or find nothing in the situation that is a "natural" intermediate state (unlike the case of 10 letters filed one by one which we discussed above). We may feel that it would be inordinately hard to describe an intermediate state in many cases. But (7) only makes an existence claim, so that the absence of any stronger property of intermediate states (interestingness, naturalness, etc.) is of no consequence.

Recall now that we discussed the notion of 'bare propositional content' in 1.2. Let P stand for some such propositional content. Now let us use the expression 'an event of the type P' in a way analogous to the way we might use the expression 'an individual of the type DOG'. If P is "BILL KISS MARY", we arrive at something like the expression 'an event of the type BILL KISS MARY', or, in English, 'an event of Bill kissing Mary'. Having explained what a subevent is, we will use the notion to form the more complex expression 'a subevent of an event of Bill kissing Mary'. Finally, we turn this expression into a one-place predicate of the form 'x is a subevent of an event of Bill kissing Mary', just as we might have made up the predicate 'x is a dog'. Using our letter P again, our predicate is of the general form 'x is a subevent of an event of the type P'. Let this be represented by 'S_P(x)'.

We are now ready to turn our attention back to sentences (5) and (6). Assuming for now that the tense in both is represented by a two-place predicate 'DUR(x,y)' meaning 'x occurs during y', whose
first argument is an event and whose second argument is a time interval, we let \( P \) represent the propositional content of (5) and (6), and let \( t_1 \) stand for 'some interval of time wholly preceding the time of the speech event'. Given this, we propose (8) as our analysis of (5), and (9) as our analysis of (6):

\[
(8) \quad (\exists x) (S_p(x) \land DUR(x, t_1)) \\
(9) \quad (x)(S_p(x) \supset DUR(x, t_1))
\]

First we note the obvious: (8) and (9) differ solely in their quantifiers, which means that our analysis of (5) and (6) reduces the difference between them to logical primitives. If we can show that we have thereby actually explicated the difference, we gain strong support for the kinds of non-logical concepts we have chosen to express what is common to (5) and (6), in particular, the notion of a subevent.

Therefore, our second concern is to make sure that (8) and (9) say the right things about (5) and (6). The non-implication of completeness observed with (5) follows from (8) straightforwardly: (8) asserts about some (not necessarily the last) subevent that it occurred during the relevant time interval \( t_1 \). But completion is only entailed if the last subevent is asserted. Hence, (8) does not entail completion. (9), on the other hand, asserts about all subevents that they occurred during \( t_1 \), including the last subevent. Therefore, (9) does entail completion.

2 3.2. On Aspect in general.

Linguists generally agree that the difference between (5) and
(6) is an aspectual one. This general agreement suggests that Aspect as a semantic category must be reasonably well defined. Let us look at a short sample of general characterizations of Aspect in the literature. Jakobson writes that Aspect characterizes "the narrated event itself without involving its participants and without reference to the speech event" (Jakobson, 1971, p.134). Comrie says that "Aspect is ... concerned...with the internal temporal constituency of...one situation" (Comrie, 1976, p.5); this characterization is strong enough to make it "doubtful whether the definition of aspect given above can be integrated to include the perfect [i.e. English have + en. E.W.] as an aspect" (ibid, p.6). Bull (1960) holds that Aspect simply deals with the fact that events have a beginning, a middle, and an end, a position which appears to be close to what Jakobson would call an aspect that "quantifies" the narrated event.

We will take our cue from section 2.3.1. and propose that the following is a rather precise statement of what Aspect deals with:

(10) Aspect quantifies over the subevents of an event.

This statement seems to be compatible with Jakobson's, Comrie's, and Bull's positions; the merit we claim for it is its greater specificity.

Examples (5) and (6) are instances of the most general aspectual distinction found in natural language: 'imperfective' versus 'perfective'. We will use these two labels just for the cases of existential and universal quantification, respectively, over the subevents of an event.

Next we will deal with another pair of rather common aspectual
categories: 'inceptive' and 'completive'. Cleaving to our quantificational format, we represent the inceptive version of (5) as in (11):

\[(11) \ (\exists x) (S_p(x) \ & \ (y)((S_p(y) \ & \ x \neq y) \implies x \land DUR(x, t_1)))\]

Two new predicates must be clarified: "x \neq y" stands for "x is not identical to y," and "x \textless y" stands for "x precedes y," meaning that the final state of subevent x is not later than the initial state of subevent y.

Inspection of formula (11) shows that it fixes x as the first in the entire sequence of subevents. Interpreting P as before, we offer (11) as the analysis of the English sentence (12):

(12) Walter began filing the day's mail.

The completive version of (5) is represented in (13):

\[(13) \ (\exists x) (S_p(x) \ & \ (y)((S_p(y) \ & \ x \textless y) \implies y \land DUR(x, t_1)))\]

This formula asserts explicitly that the last in the sequence of subevents occurred in \(t_1\), and is our analysis of the English sentence (14):

(14) Walter finished filing the day's mail.

It is of some interest to examine at this point whether (8), (9), (11), and (13) yield the full set of entailments that seem warranted on the basis of sentences (5), (6), (12), and (14). It turns out that the formulas referred to above are adequate in this regard, provided we fix as a postulate that the occurrence of a subevent later in the event sequence entails the occurrence of all subevents earlier in the sequence,
but that the occurrence of a given subevent does not entail the occurrence of any subevents later than it. This postulate, which we might dub the 'Uncertain Future Postulate', brings out an important logical point about events: their occurrence is not entailed or presupposed by the fact that a definite description exists of them. To say that some action, for example, constitutes a subevent of an event e is not to presuppose or to guarantee that e itself exists, i.e. has occurred or will occur (as a whole). This situation is not unique to events. An exactly analogous situation obtains with certain 'objects of creation'. To say that a few dozen typewritten pages constitute part of my dissertation is not to say that the dissertation itself exists (as a whole), nor is there any guarantee that it ever will exist. All that seems to be required for the expression 'part of my dissertation' to be used felicitously is that it is reasonably well defined what would count as 'my dissertation', and, perhaps, that there is at least some reason to suppose that there will be such a thing if all goes well.

But to return to our main theme: Equipped with the Uncertain Future Postulate, and interpreting the time interval $t_1$ as before (bearing in mind that while this latter point is extraneous to Aspect, it is not extraneous to the matter of determining the truth-value of a proposition), we can derive just the right set of entailments (although to show this in detail would be somewhat complicated, since we have to make use of the built-in strict temporal ordering of subevents with respect to one another). This set of entailments is as follows: (9) entails all of (8), (11), and (13); from (13), all of (8), (9),
and (11) follow; (8) and (11) imply each other; from neither (8) nor (11) do either (9) or (13) follow.

The addition of specific time adverbs reduces the set of entailments, but in ways that appear quite compatible with the aspectual analysis pursued here. For example, while Walter finished filing the day's mail entails Walter began filing the day's mail, there is no entailment from Walter finished filing the day's mail at 4:30 p.m. to Walter began filing the day's mail at 4:30 p.m. No rigorous account will be given here of facts such as these; it is rather transparent that there are no problems of analysis lurking here. What little discussion there is of the interaction of time adverbs with aspectual distinctions in section 2.3.3.1 supports this impression. It is shown there that the present analysis of Aspect appears to be logically adequate in a number of interesting examples. In addition, the general outlines of a logic of time adverbs are presented there, completing the theoretical framework within which a detailed analysis can be worked out.

The formulas in (8), (9), (11), and (13) exhaust the aspectual core distinctions. All other aspectual specifications deal with the temporal contiguity of subevents. They can all be analyzed by reference to the notion of a \textit{pause}. This is intended very much in the sense of the English word. In our case of Walter filing the mail, if he has filed five letters and then sits down to enjoy a cigarette before he files the sixth, that's a pause. In the language we have developed to introduce subevents we would say that a pause is a time interval throughout which there is only one and the same intermediate state. During a \textit{pause}, the event itself does not advance toward the terminal state.
To make our exposition of the remaining aspectual categories perspicuous, then, we introduce the one-place predicate 'x is a pause' ('P(x)') with the meaning just described, and extend the predicate 'precedes' to involve the initial point of a pause in addition to the initial state of a subevent, and the final point of a pause in addition to the final state of a subevent.

The category 'cessative' will then be analyzed along the lines of (15):

(15) \( \exists x \) (\( \exists y \) (Sp(x) & P(y) & x\( \leq y \) & Dur(x, t_1)))
(15) says that some subevent occurred, followed by a pause. This is our analysis of the English sentence (16):

(16) Walter stopped filing the day's mail.

From this it is but a small step to the category 'continuative'. Where (15) asserts a pause, (17), our analysis of the 'continuative', negates it:

(17) \((\exists x)(\exists y)(Sp(x) \& P(y) \& x<y \& DUR(x,t_1))\)

This makes explicit the aspectual meaning conveyed by (18):

(18) Walter kept filing the day's mail.

A small step in a different direction takes us from 'cessative' to 'resumptive'. The former has a pause following a subevent; the latter reverses the temporal order, as in (19):

(19) \((\exists x)(\exists y)(Sp(x) \& P(y) \& y<x \& DUR(x,t_1))\)

This gives us the analysis of the English sentence (20):

(20) Walter resumed filing the day's mail.

The formulas in (15), (17), and (19) bear a formal resemblance to the ones in (11) and (13), in that they involve the predicate 'precede'. It appears that (8) and (9), which make no reference to the sequential arrangement of subevents, have their formal analogues in the "pausal" aspects as well. Jakobson mentions a category of 'determinate aspect',


"signaling the integrity, unbrokenness of $E^n$ [=narrated event, E.W.]". We would represent this as in (21):

\[(21)\quad (x)(y)((Sp(x) \& P(y)) \supset (DUR(x,t_1) \& \neg DUR(y,t_1)))\]

Finally, we could conceive of one other fairly natural category, which we might dub 'interruptive', and which would have (22) as its analysis:

\[(22)\quad (\exists x)(\exists y)(Sp(x) \& P(y) \& DUR(x,t_1) \& DUR(y,t_1))\]

(21) asserts the absence of any pauses, and (22) says that there are some. Neither (21) nor (22) contains the predicate 'precede', just as neither (8) nor (9) does.

This completes our sketch of aspectual categories. We will now construct a hypothesis about the internal structure of the entire domain of aspect. This is a hypothesis about a substantive semantic universal. Both the Domain Determinacy Hypothesis and the Complement Convention of Chapter 1 presuppose the existence of such substantive characterizations of verbal domains. We present the hypothesis about the structure of Aspect in the form of a tree (fig. #1: (23) The structure of ASPECT). The branching nodes are numbered, and we will specify below what distinction is made by each bifurcation. The terminal nodes all carry an informal label for the category in question, and the number under which we gave a characterization of the category in the text. Some path segments are marked m (for 'semantically marked'). This means that the branch so marked can be left out of an aspectual system, without violating the overarching principle that the domain of Aspect be
The structure of ASPECT

1
  2   m   3
   m   m
5 'determinate' 'perfective'
4 'imperfective' 'interruptive'
  (8)   (22)
  'inceptive' 'completive'
  (11)   (13)
  'cessative' 'continuative'
  (15)   (17)
6
  7
8 'resumptive'
  (19)
exhausted by every aspectual system. This marking has not been carried through to the extent that might be possible; as things stand now the only claim made is that an aspectual system containing just 'imperfective' and 'perfective' is complete.

Now for the enumeration of the distinctions made at each node of (23), going top to bottom and left to right:

(24) 1: existential vs. universal quantification over subevents
2: independence from vs. dependence on sequential information (required by 'precede')
3: reference to pauses vs. no reference to pauses
4: no reference to pauses vs. reference to pauses
5: no reference to pauses vs. reference to pauses
6: first vs. last subevent
7: subevent before pause vs. subevent after pause
8: pause asserted vs. pause negated

What is not brought out very clearly by (23) and (24) is that this entire structure is built out of only three non-logical concepts: that of a subevent, that of a pause, and that of temporal precedence. Furthermore, what the central concept in the domain of Aspect is--a subevent--is not made explicit. This latter fact bears directly on the Domain Determinacy Hypothesis, which must be taken to be instantiated for Aspect as follows: if the augment of a semantic translation rule R contains the predicate 'x is a subevent', then Aspect is the semantic domain that R deals with.
We can now also make the Complement Convention explicit. Given the binary organization we have assumed for the semantic domains in the TMA component, there will be exactly one sister node for any node in a semantic structure. The complement of a node is that unique sister node; for example, in (23), 3 is the complement of 2, etc.

This concludes our discussion of Aspect in general. We now turn to a discussion of some questions about Aspect in English.

2.3.2. The aspectual system of English.

Consider the set of verbs begin, commence, start, keep, continue, go on, resume, stop, finish. They all have in common the fact that they take participial complements. They further have in common the fact that alongside such participial complements, they all either take ordinary NP or PP complements. Participial complements are a case of highly restricted verbal complementation (that is to say, participial complements are a very limited type of verbal complement in English, as compared to infinitival complements with to, and finite complements). In section 2.2., we noted that in Jackendoff's theory of the base such restricted verbal complements that alternate with non-verbal complements (NP's or PP's) and always lack the verbal auxiliaries are best analyzed as instances of the recategorization schema $X^n \rightarrow_{\text{AFF}} Y^n$. The participial complements to the abovementioned set of aspectual verbs meet all the criteria for being analyzed in this way, and we propose (25) as their syntactic analysis:
Our interest in assuming a structure like (25), in spite of the obvious problems with it, is that it makes explicit that the verbs that enter into it are highly grammaticalized. The occurrence of recategorization in their complements must, after all, be governed by a special syntactic feature of those verbs, unlike normal complementation, where full clauses are embedded, with the verb in addition determining the complementizer type. Any other syntactic analysis that makes this difference
similarly explicit would serve equally well for our purposes. For as we have pointed out before, only highly grammaticized verbal categories are 'factored out' to be dealt with in the TMA component, rather than simply entering into the determination of propositional content.

Let us assume that the aspectual verbs are sufficiently grammaticalized to be factored out in the sense mentioned. Then the set of verbs usually referred to as verbs of temporal aspect (e.g. Newmeyer (1969), Emonds (1970)) make up the following subtree of the tree in (23):

(23')

```
  5
 / \
 6   7
/   /
begin finish resume
/   /
start commence stop keep

/   /
commence quit continue

/   /
go on
```

All the 'sequential' aspects, as we might call the categories under 5, are lexically represented; so, on the face of it, 5 is syntactically marked. Its complement is 4 of tree (23), which includes as an obligatory category the imperfective. But taking 5 as a marked category and 4 as its unmarked complement, whose meaning is not given by rule but by the Complement Convention, is not a theoretically available analysis.
It is here that the Domain Exhaustion Principle, briefly mentioned in section 1.4, acquires direct relevance. That principle requires that a domain which a grammaticalized set of forms \( F_g \) maps into must be exhaustively covered. That is to say, every semantic category \( C_s \) which belongs to such a domain must either be the meaning of a form which is in \( F_g \), or \( C_s \) must be a subcategory of some semantic supracategory \( C_s' \) such that \( C_s' \) is the meaning of a form in \( F_g \). Now if the domain of Aspect cannot be exhausted without inclusion of the category 'perfective', and if the Complement Convention can only pick out sister nodes in semantic systems as semantic complements, it follows that any aspectual system that (23') is a part of, must mark imperfective aspect, so that perfective aspect can be integrated into the system by application of the Complement Convention. If in addition to those included in (23') there is also an overt form for imperfective aspect, 'perfective aspect' will be the meaning of the null-member of the set of aspectual verbs. Otherwise, perfective aspect would remain unexpressed, contrary to the Domain Exhaustion Principle, resulting in a defective system.

The conclusion for English is that the progressive \textit{be}, which we have shown in 2.3.1 to be imperfective semantically, is the missing member of the set of verbs of temporal aspect, with which it shares the syntactic analysis as a main verb with participial complements. This pattern of participial complementation is the syntactically marked value of the English aspectual system; the absence of participial complements, i.e. the case where the verb that corresponds to the main predicate of the propositional content is also the matrix verb syntactically, is the syntactically unmarked value of the aspectual system, and by the principles elaborated
above, must be interpreted as perfective.

2.3.3. The English perfective-imperfective opposition in more detail.

2.3.3.1. In this subsection we will simply underscore the validity of the claim that the English progressive/non-progressive alternation is interpreted by an imperfective/perfective distinction. First, we consider a passage from Leech (1971), which makes the point about implied completion:

(26) "'The man was drowning'./'The man drowned'.
To the first sentence one could add ... 'but I jumped into the water and saved him': but not to the second, which implies that the man actually died." (p.16).
and, similarly,

"'I was reading a book that evening' and 'I read a book that evening'. The Simple Past here suggests that the speaker reached the end of his book before the end of the evening; completion in this sense is not implied by 'was reading'. (p. 17)

Leech's comment on the first pair of examples can be taken to show that the non-progressive forms imply completion in the logical sense of implication, rather than in the looser sense of a conversational implicature, since the implication is not cancellable without the speaker contradicting himself.

Our next point is that the non-progressive does, in fact, only imply completion; it does not assert it. This is what sets perfective aspect apart from completive aspect. To make this point, we must say something about temporal adverbs. We will assume that every assertion in natural language must be interpreted with reference to some time interval over which an asserted state of affairs is claimed to be true, or through which an asserted event is claimed to occur. Let this time interval be $T_1$. Temporal adverbials designate a specific time interval (not just length of time, but an interval locatable on the time line), call it $T_2$. The unmarked relation between $T_1$ and $T_2$ (indicated by the propositions at, on, in, during, or by the absence of a proposition) is as follows: $T_1 \subseteq T_2$.

Now consider (27) and (28).

(27) John finished reading the voluminous novel at midnight.

(28) John read the voluminous novel at midnight.
On our analysis, (27) refers only to the last subevent of the larger event of 'reading a voluminous novel', while (28) refers to all subevents. This makes (28) pragmatically incongruous, because 'midnight', no matter how flexibly interpreted, is too short an interval to accommodate the reading of the novel from start to finish. It can, however, accommodate the finish alone, and there is consequently nothing wrong with (27).

Similarly dependent on the interpretation of temporal adverbials is the next example, again taken from Leech (1971):

(29) The following sentences illustrate the lack of certainty about completeness in another context:

I was reading from 10 p.m. to 11 p.m./I read from 10 to 11 p.m. The Simple Past tells us that the speaker started to read at 10 o'clock and finished at 11 o'clock. The Progressive, however, does not specify either the time of beginning or the time of completing the activity; all we know is that reading was in progress for that hour. Hence it would be a fitting answer from a suspect being interrogated by a detective. The detective would ask 'What were you doing between 10 p.m. and 11 r.m.? being uninterested in whether the activity persisted after that period or not; and the suspect would reply in kind.

This is just what our analysis entails. The progressive form leaves open the possibility that some reading preceded or followed the time interval specified; the non-progressive form does not.

2.3.3.2. This subsection is devoted to the discussion of the following interesting contrast (the description is once more taken from Leech (1971)):
"... the Progressive often forms a 'temporal frame' around an action denoted by a non-progressive form; in this case, whereas the relationship of meaning between two neighboring Simple Past forms is usually one of 'time sequence', the relationship between a Progressive and a Simple Past form is one of 'time-inclusion'. The contrast can be studied in these two sentences:

When we arrived she made some fresh coffee.
When we arrived she was making some fresh coffee.

The first example tells us that the coffee making followed the arrival; the second, that the arrival took place during the coffee making.

Leech's description of the contrast between his two sample sentences is quite accurate. However, the informal account in terms of neighboring tense forms is somewhat beside the point. The examples do not primarily illustrate a point about neighboring tense forms, but rather something about the semantic properties of the temporal connective when. The temporal connectives of English have recently been examined in detail by Heinämäki (1974). Heinämäki proposes that the central uses of when can be described in terms of the following truth conditions:

\[
(31) \quad A \text{ when } B \text{ is true if and only if }
\]
\[\begin{align*}
& (i) A \text{ is true at an interval } J, \\
& (ii) B \text{ is true at an interval } K, \text{ and } \\
& (iii) \text{ a. if } A \text{ or } B \text{ (or both) is durative, then there is an interval } I \text{ such that } I \subseteq J \text{ and } I \subseteq K. \text{ (If } A \text{ is an accomplishment, then } J = I; \text{ if } B \text{ is an accomplishment, then } K = I). \\
& \quad b. \text{ If neither } A \text{ nor } B \text{ is durative, then } F(K) < G(J), \\
& \text{ where } F(K) \text{ is the final point of interval } K, \text{ and } G(J) \text{ is the initial point of interval } J.
\]

The paradigm case of (31iiia), that of two 'durative' clauses, gives rise to a situation of temporal overlap, of which Leech's 'time inclusion' is special case. A representative example is (32):

(32) It was raining in New Orleans when we were there.

The schematic representation of this is (33):

(33)

\[ \text{A when B} \]

\[ \text{A It was raining in New Orleans.} \]

\[ \text{B We were in New Orleans.} \]

As Heinämäki puts it: "...some subinterval of the period defined by the main clause is also a subinterval of the period defined by the when-clause."

The case covered by (31 iiib) is that of a 'time sequence'. We give (34) as a typical example:

(34) When John pushed the button, the bomb exploded.

According to Heinämäki, "the natural interpretation is that the events described by the clauses happen in succession: the one mentioned in the when-clause takes place first, then the one expressed in the main clause." Schematically, we have (35):

(35)

\[ * \text{A The bomb exploded.} \]

\[ * \text{B John pushed the button.} \]
This general account of when-clauses (there are altogether eight distinct schematic configurations covered by (31)) describes the facts of English when-clauses exhaustively. But what exactly does it say about the meaning of when? It says that it is a semantic property of when to define either a relation of temporal overlap on two clauses, or a relation of temporal non-overlap. What Leech (see (30) above) took to be a fact about adjacent tense forms has here become a fact about the meaning of a temporal connective. The question of whether there is an interesting relationship between the two cases does not arise in any serious form.

Given that the temporal connectives are a small closed set of forms, just this question does arise in our framework, however; the temporal connectives are highly grammaticalized, and thus fall under the Univocality Assumption. We give the following as our univocal meaning specification for when-clauses:

\[(36) \quad A \text{ when } B = \text{df.} \]
\[(\exists x)(\exists y)(T(x) \& T(y) \& \text{DUR}(A,x) \& \text{DUR}(3,y) \& y<x)\]

\(T(x)\) stands for 'x is a time interval', \(\text{DUR}(x,y)\) and \(y<x\) are interpreted as before, except that \(\text{DUR}(x,y)\) is somewhat broader, admitting in the first argument position not just events, but states of affairs and facts as well. We will examine the adequacy of (36) only to the extent that it sheds light on the imperfective/perfective distinction, which, as we have made clear at the outset, is limited to event descriptions, and is undefined for state descriptions. The extension of (36) to states, while
unproblematic, can only be discussed towards the end of Chapter 4, after some necessary groundwork has been laid.

So let us look at the sentences in (37) and (38) in relation to the analysis of *when* given in (36):

(37) When Jim broke the new typewriter, Mary just fixed the old one.

(38) When Jim broke the new typewriter, Mary was just fixing the old one.

In both (37) and (38) the matrix clause, as well as the embedded clause, refer to events; thus the aspectual imperfective/perfective distinction is applicable. The *when*-clause is non-progressive in both cases, and consequently receives a perfective interpretation. The matrix clause is perfective in (37), and imperfective in (38). Let $F$ represent the propositional content of the matrix, and let $B$ stand for the propositional content of the *when*-clause. Putting together what elementary analysis we have developed for Aspect and for *when*-clauses, we get (39) as our analysis of (37), and (40) as our analysis of (38):

\[
(39) \ (\exists x)(\exists y) (T(x) \land T(y) \land (z)(S_F(z) \supset DUR(z,x)) \land
\begin{align*}
(w)(S_B(w) \supset DUR(w,y) & \land y<x) \\
(40) \ (\exists x)(\exists y) (T(x) \land T(y) \land (\exists z)(S_F(z) \land DUR(z,x)) \land
\begin{align*}
(w)(S_B(w) \supset DUR(w,y)) & \land y<x)
\end{align*}
\]

This looks more forbidding than it is: (39) says that all subevents of
typewriter fixing follow all subevents of typewriter breaking. This is clearly an instance of what Leech calls 'time sequence'. (40), on the other hand, says that some subevent of typewriter fixing follows all subevents of typewriter breaking. This does not guarantee temporal overlap, but merely admits the possibility of temporal overlap. But this is no defect of our analysis. (40) differs from (39) in using an existential quantifier instead of a universal one. And by the Gricean conversational postulate that the speaker is presumed to say as much as he knows, an existentially quantified expression is inappropriate where the universally quantified expression is known to be true. Thus, it is inappropriate for a speaker to say "Some ravens are black" if he knows that all are. In sum, while (40) does not guarantee temporal overlap on logical grounds, it can hardly be used to refer to anything else, for the pragmatic reason just mentioned.

The case of when-clauses is a rather instructive one. We have reduced a superficially rather puzzling semantic bifurcation between 'time sequence' and 'time inclusion' to the interaction of a highly motivated and conceptually simple aspectual analysis, a single and, again, rather simple semantic analysis of a temporal connective, and a very general conversational principle. A successful analysis of this kind provides a considerable measure of support for the elements it rests on. Our main interest in this subsection has been to provide such support for our analysis of the imperfective/perfective distinction.

2.3.3.3. Evidence of a rather different kind comes from the interaction of Tense with the progressive/non-progressive alternation. It is a
well-known fact that if we compare certain progressive/non-progressive pairs in their past tense forms with their present tense counterparts, we notice a rather obvious lack of symmetry. While the progressive forms, say He was leaving and He is leaving, differ in their temporal reference, but are otherwise semantically parallel, the corresponding non-progressive forms, He left and He leaves, are not similarly parallel. This observation is commonplace, and pairs like (41) and (42)

(41) John played a superior game at centerforward.

(42) John plays a superior game at centerforward.

illustrate the difference clearly: (41) can be taken to refer to a single event, implying its completion in the past, but (42) cannot be taken to refer to a single event, implying its completion in the present. Sentences like (42) are said to have only a generic sense or use. Sentences like (41) often have that generic sense, too, but in addition they have the specific sense just described.

We cannot approach this problem without giving at least an outline of Tense, and of the English tense system. The theoretical construct that distinguishes Tense from other verbal categories is that of $t_0$, 'time of the speech event'. Most of the discussion of Tense concentrates on the kinds of relations that natural languages can define with respect to $t_0$. We will focus on a particular property of $t_0$ itself: How big is it? Our answer is that $t_0$ is always a minimal time interval, i.e., it has the properties of a psychologically reasonable 'point in time'
(the 'unreasonable' interpretation we want to avoid is that of a
temporal entity with location but zero extension). Philosophically,
this notion of an 'atom of time' is not a very comfortable one, but for our
purposes here we will brush philosophical scruples aside, and proceed
under the presumption that our description of \( t_0 \) as such an atom of time
is intelligible.

Now on to tense in English. We assume, without further argument
at this point (but see Chapter 4), a three-tense system for English,
consisting of Future, Past, and Present. Our analysis of future tense
is roughly as in (42a):

\[
(42a) \quad \text{FUTURE}(p) = \text{df.} \quad (\exists x) (T(x) \& \text{DUR}(p,x) \& t_0 < x)
\]

\( T(x) \) stands for '\( x \) is a time interval', as before; the interpretations
of \( \text{DUR}(x,y) \) and \( x \sim y \) are also unchanged. We use \( p \) to stand for some
untensed proposition. Past tense is exactly like future tense, except
that the final conjunct in the definiens is \( x < t_0 \). Present tense is
also identical except for the final conjunct, which is \( x = t_0 \). The new
predicate \( x \sim y \) means '\( x \) is simultaneous with \( y \)'.

The stage is now set for an analysis of the case of asymmetry
which we described earlier in this section. Let us deal with the
present-tense case first. We will use \( P \) to represent the propositional
content of (43) and (44) below:

(43) John is playing a superior game at centerforward.

(44) John plays a superior game at centerforward.
and propose to analyse (43) as in (45), and (44) as in (46):

\[(45) \ (\exists x)(T(x) \& (\exists y)(Sp(y) \& DUR(y,x)) \& x=t_0)\]

\[(46) \ (\exists x)(T(x) \& (\forall y)(Sp(y)\supset DUR(y,x)) \& x=t_0)\]

As it turns out, of the two formulas (45) and (46), only the first is satisfiable. To verify this claim, we have to show that (46) is not satisfiable, and that (45) is.

To see why (46) is not satisfiable, we must first fix in mind the hypothesis that \(t_0\) is an atom of time. Recalling our discussion of subevents in 2.3.1, we further assert that no more than one subevent of a given event can occur during such an atom of time. This assertion is trivially true: we have tailored our characterization of minimal time intervals so as to guarantee its truth. By postulate (7) in 2.3.1, which said that there were intermediate states in every event, we are guaranteed that the domain of the universal quantifier in (46) has at least two members. But now (46) says that there is some time interval simultaneous with \(t_0\), and therefore atomic, such that at least two subevents occur in it, which, given what we have just said, is impossible.

(45), on the other hand, is satisfiable, since the occurrence of at least one subevent at a minimal time interval is possible. The fact that it can at most be one, is logically irrelevant.

Contrast the case of present-tense (43) and (44) with the case of past-tense (47) and (48) (to which future tense is analogous in all relevant respects):
(47) John was playing a superior game at centerforward.

(48) John played a superior game at centerforward.

(49) is the analysis of (47), and (50) that of (48):

(49) \( (\exists x)(T(x) \& (\exists y)(S_p(y) \& DUR(y,x)) \& x<t_0) \)

(50) \( (\exists x)(T(x) \& (y)(S_p(y)\supset DUR(y,x)) \& x<t_0) \)

(49) says that at least one, and (50) says that all, subevents of the event in question occurred before \( t_0 \). Since there is an unlimited supply of time along that portion of the timeline, both (49) and (50) are, of course, satisfiable.

One point of special interest arises with (49), our analysis of past-progressive sentences. (49), as distinct from (50), does not imply that the event as a whole was completed before \( t_0 \). This is an entirely desirable consequence of our analysis, since sentences like (51) are free from contradiction:

(51) He was writing a letter to his congressman when I talked to him an hour ago, and he's still at it, the old fool.

This is to be contrasted with (52):

(52) #He wrote a letter to his congressman, and he's still at it, the old fool.

From (52), a contradiction can be inferred, since the first conjunct implies that the letter-writing was finished before the time of the
speech event, while the second conjunct asserts that it is still in progress at the time of the speech event.

The account we have provided here for a rather curious fact of English depends heavily on the claims we have made about the English tense system, and about the properties of $t_0$. On the face of it, these claims, especially those about $t_0$, are quite outrageous. The results of section 2.4 will, however, go a long way towards showing how these claims can be sustained rather naturally.

Our discussion of Aspect in English ends here. We have proposed an analysis of aspectual $be + ing$ as a main verb with a syntactically highly marked participial complement, treating it as a member of a small class of aspectual verbs that share this syntactic peculiarity. We suggested at the end of section 2.2 that there were semantic considerations that lent some support to a dual syntactic analysis of progressive $be$. To the second syntactic alternative, progressive $be$ as a true auxiliary, and its semantics, we now turn.

2.4. Epistemic (Auxiliary Verb) BE.

2.4.1. The notion of a structural fact.

Consider the sentences (53) and (54):

(53) Sam drives a truck for the ABC Company.

(54) Sam is driving a truck for the ABC Company.

Both (53) and (54) are appropriate to describe a situation in which Sam
works as a truck driver, and receives a weekly paycheck from the ABC Company for his services. Both are appropriate in a situation where, as far as the ABC Company is concerned, Sam is just another name on the payroll, gets paid his union wages, etc. In short, if we just look at the bare facts of Sam's work contract, his obligations to perform various duties, his rights to refuse excessive overtime, his pay scale, etc. we find no criteria such that two different types of situation can be characterized, one to which (53) is more appropriate, and one to which (54) is more appropriate.

And yet there is a difference between (53) and (54). (53) says more about Sam than does (54). The non-progressive form categorizes Sam in terms of his station in life: he is a truck driver. A host of more or less explicit assumptions about Sam are bound up in this categorization: assumptions about Sam's educational background, about his earning power, about where he is likely to spend his vacations, and about the kind of restaurants he is liable to patronize. If we were asked to describe the kind of man Sam is, (53) would be a substantial piece of information, because of its potential for characterizing the man.

It is in this regard, in the suitability of the sentence for representing a state of affairs as an integral component in a more or less well-organized conception of some entity, that (53) differs quite sharply from (54). (54) is apt as the description of a man who goes to night school in order to become an accountant, but also holds a daytime job because he has to eat. This is perhaps an overdramatization of the
contrast, but in the right direction, nonetheless. Whatever the nexus of associations and assumptions that is evoked by characterizing a man as a truck driver, that entire nexus is left unevoked by (54). This is the weakest statement we can make about (54): that it is neutral on this point of evoking any larger conceptual structures that go beyond the bare facts expressly stated in the sentence. Given the availability of the contrasting form (53), and the Gricean principle mentioned in 2.3.3.2, the implicature of (54) will generally be that what it describes is not a characterizing property of Sam's.

Let us call the kind of fact conveyed by the non-progressive (53) a structural fact. We choose the adjective structural to highlight the way such non-progressive sentences suggest that the facts described in them are part of a larger organization of related facts, part of a structured conception of something.

We can now say that the difference between (53) and (54) may be described, quite informally, as follows: (53) describes a structural fact, and (54) does not describe a structural fact. Without any further explication of the precise meaning of structural fact, there are two points we can make even now: first, the distinction we have made is a binary one, and secondly, over the set of sentences that describe facts it is an exhaustive one.

The simplest syntactic arrangement for the representation of semantic domains with such a structure is to have a marked form, which is related by semantic rule to the one member of the binary opposition, and an unmarked form, which is related to the other member by the
Complement Convention. We propose that just such a syntactic arrange-
ment represents the abovementioned semantic contrast in English.
We have touched on the syntactic structure involved before, and present
it here again:

(55)

The stem be is this syntactic structure alternates with no other stems;
it forms a substitution set of one, contrasting only with its absence.
This is the highest degree of grammaticalization for any grammatical
formative, and there is thus no question but that the principles pro-
posed for the TMA component in Chapter 1. should fully apply to this
case.

The first of those principles that we will address ourselves to
is the Domain Determinacy Hypothesis. Exactly what kind of domain are
we dealing with? And what is its distinctive vocabulary, comparable
to the way that the predicate 'x is a subevent', e.g., is distinctive of Aspect?

2.4.2. A sketch of the domain of Epistemic Status.

The contrast between (53) and (54) we have just discussed is a specific instance of the major distinction made in the semantic domain of Epistemic Status, which we will now describe in more general terms.

We will maintain that there are two ways of describing the world. One may describe what things happen in the world, and one may describe how the world is made that such things may happen in it. This dichotomy is exhaustive: there isn't anything to describe outside of either the phenomenal content, or else the structure, of the world. Phenomena are the objects of observation, structure is the object of knowledge. The former are what our perception of the world deals with, the latter is what our conception of the world deals with. The relationship is like that of data to a theory. And so, at the most elementary level, what we are saying here is that one can either talk about data or one can talk about a theory of those data. The important thing to see at this point is that they are really very different things.

These remarks give some hint as to the relation between these two categories of phenomena, on the one hand, and structure, on the other hand. Made explicit, the relationship is this: Phenomena provide the evidence for assuming a particular structure, or if a structure has already been assumed, phenomena provide an empirical test of the
correctness of those assumptions. Not all phenomena, or all structures, participate in this relationship, of course. There may be phenomena that are considered to be mere happenstance --- data, as it were, that no theory bothers with. Structure, on the other hand, may have no phenomenal correlates of the kind described, as would be the case with structures stipulated by convention, such as a code of etiquette for example, which is related to phenomena in a rather different way.

The notion of structure is not a temporal notion. To be sure, any description of the constitution of an object presupposes that the constitutive properties of the object do not randomly fluctuate in time: structure is characterized by non-randomness. But structure is not immune to change - it is not permanent. So the relationship of structure to passage of time is a somewhat trivial one: structure persists, unless the entity in question changes.

The concept of change is, in fact, quite intimately related to structure. We normally use the word change to refer to structural change. Mere variation in phenomena does not constitute change. Consider a singer's voice. All we have to judge this object, 'the singer's voice', by is the sounds we hear. Nevertheless, if we say of a singer that his voice has changed, we do not, for example, refer to a situation where the man is suffering from a cold so that the sounds we hear are different from the ones we are used to. We would not even say "His voice has changed, but as soon as he is over his cold it'll change back to normal again", if there were nothing wrong but the usual hoarseness attributable to the cold. On the contrary, if somebody were
unaware of the cold and blurted out, "Oh, but his voice has changed!", we could say reassuringly, "Don't worry, his voice hasn't changed at all. He's just suffering from a little cold". This, then, is a perspicuous case of variation in phenomena not counting as change. The assumed structure is to some extent immune to the challenge of phenomena that are, taken at face value, counterevidence.

Now suppose the cold goes away, but the hoarseness doesn't, and gargling doesn't help either. Then we might well conclude sadly, "His voice really has changed." In phenomenal terms, it is the same hoarseness we hear, yet we arrive at a different judgment. But not based on temporal factors --- what is relevant is such things as knowledge about colds and their effect on the sound of voices.

This discussion of the relation between structure and change is of some importance, because judgments about change are as close as we will come here to a diagnostic for determining whether we are dealing with a structural or a phenomenal description. The test is simply this: If \( F(a) \) is true at some time \( t_1 \), and \(-F(a)\) is true at \( t_2 \), \( t_1 < t_2 \), then \( F(a) \) is a structural description if, on the basis of \(-F(a)\), the statement a has changed is true at \( t_2 \). As a true diagnostic, this test is of limited value, unfortunately. The problem with it is that in order to use the word change in the spirit illustrated above, one must already 'have' the concept of 'structure' that the test is supposed to lead to. Take the following examples as representative of this difficulty: Suppose in our town things are different from last year in that the following statement is now true which wasn't true a year ago: "Our
school admits minority students." Has our school changed? We do not hesitate to answer in the affirmative, because a matter like a school's admissions policy can be a very important ingredient in our conception of what our school is, or ought to be. Now let us contrast this with the case of the novel situation be this: "Our school has minority students among its students." Would we now say that our school has changed? Probably not, certainly not in the same way we could where the entire admissions policy, and not just some accidental facts about actual admissions, was at stake. A man might infer from bare facts about admissions that there has been some deeper underlying change, but one would surely trivialize the statement "Our school has changed" if one had in mind nothing but those bare facts. Keeping this qualification in mind, we can put our diagnostic test to good practical use, if only because change is much more common a word than structure, and thus easier to have sound intuitions about.

These introductory remarks set the stage for our analysis. We will now turn to a detailed discussion of a variety of examples, and by employing the phenomenal/structural distinction over a substantial range of data, we will bring out its nature more clearly.

2.4.3. The English structural/phenomenal opposition in more detail.

2.4.3.1. Consider the pair of sentences (56) and (57):

(56) The engine isn't smoking anymore.

(57) The engine doesn't smoke anymore.
Our semantic analysis of this pair of sentences is that (56) describes the engine in question phenomenally, and (57) structurally. Let us consider a scenario to make the distinction clear as it applies in this case.

Suppose your car has been smoking a lot recently, and, knowing a lot about automobiles, you decide to repair it yourself. You pinpoint the source of the trouble in a defective hose, and you replace it. You can now definitely assert, "The engine doesn't smoke anymore," with --- or without --- letting it run smokelessly. To say, "The engine isn't smoking anymore," you would certainly have to start the engine first, and your comment would be just an observation, rather than a claim about it being repaired, as (57) is.

On the other hand, imagine that you are driving down the highway with the engine smoking. You and your passenger comment on this --- "Look, the engine is smoking" --- and your passenger dozes off to sleep. He wakes up, and noticing no more smoke, he may well say (56), "The engine isn't smoking anymore." It would be distinctly odd to utter in this context the non-progressive (57).

What makes these two cases different is the presence or absence of 'repair', which is an example of the abovementioned more general notion of change.

Having first considered the example of a smoke-free engine, a material object, we choose a non-material object for our second example. Non-material objects show both phenomenal and structural properties, too, quite parallel to the case of the engine. Here we do not discover the
properties that constitute the object; rather, we invest it with them. Let us consider the pair of sentences suggested to us by Susan Martin:

(58) This law raises the price of oil by 10¢ a gallon.

(59) This law is raising the price of oil by 10¢ a gallon.

Here, as above, the non-progressive, (58), describes the structural properties of the law; (59), the progressive, describes phenomenally. The strong properties of the law are those which, if altered, would lead us to say the law itself had changed, as we explained above. The phenomenal properties will lack this characteristic.

Sentence (58) is an observation about what was written into the law by the legislator that drafted it; we could as well say of a bill --- not yet passed --- that it "raises the price of oil by 10¢ a gallon," just under the condition that this is indeed a provision of the proposed law. (59), on the other hand, refers to the observable consequences of the law, leaving quite open whether the price hike was the original intended result of the law. On this analysis, then, the progressive should make little sense when speaking of a bill, which rarely has observable effects:

(60) This bill is raising the price of oil by 10¢ a gallon.

And in fact (60) is anomalous; it is very hard to conceive of a situation where it is appropriate. Perhaps the only such would be if rumors of the bill circulating in the legislators' cloakroom were having certain
effects on the oil market; but again, it is the observable effects which are reported by the progressive form in (60).

The previous examples about a car engine and a federal law have dealt with what we might call *internal structure*. Whether an engine smokes or not has to do with its internal constitution, and whether a law raises prices or not has to do with its contents. But our conception of structure, or what characterizes an object, is not confined to internal structure; it extends to structured external relations. Consider the following two sentences:

(61) The statue of Tom Paine stands at the corner of Kirkland and College.

(62) The statue of Tom Paine is standing at the corner of Kirkland and College.

Now (61) does not say anything about the internal makeup of the statue in question. Rather, it relates the statue to a location, and says that there is something structural about this relation.

Let us now examine how the phenomenal/structural distinction applies here. In both sentences, we are dealing with a human artifact, a statue. Now, a lot of statues may get sculpted just to satisfy a creative impulse, and they may end up sitting any old place where they are in nobody's way. They would have *no* structured relations to a location; they have a location as a matter of simple observation, but whatever that location happens to be, it could just as well have been
any number of other places. Their location is the kind of fact that no theory, no organizing principles, no structure, is concerned with. (61) and (62) deal with a different kind of case. It's the case where the location of the statue is part of its concept. This would be true where, e.g., the City Beautification Committee commissions a statue so that it may grace some busy street corner and foster civic pride. Now the location of this statue is not a matter of indifference. The City Council will hold a public hearing, and by due administrative procedure the statue will be assigned to a chosen location. And that's the kind of situation which (61) reports appropriately: the statue is in its rightful place. Sentence (62), on the other hand, should be appropriate if a bare eyewitness account is to be given (which is in turn a pragmatic matter: when is a bare eyewitness account as far as one is willing to go in describing the location of a statue or even an appropriate, or sensible act?).

We will take up in some detail now the point about the non-temporal nature of the phenomenal/structural distinction. This point deserves attention, because it is the point where we part company with all other descriptions and analyses of the English progressive we are aware of (to mention but a few, Jespersen (1961), Leech (1971), Joos (1964), Fiengo (1974)). The standard analysis of the progressive/non-progressive distinction is in terms of the temporal notion of duration, and the conventional meaning assignments are "limited duration" for the progressive, and "duration not limited" for the non-progressive. "Duration not limited" does duty for timeless truths, for
generic truths, and for what is often called habitual aspect, to name the central cases. "Limited duration" is most commonly interpreted in the sense of 'the state of affairs described is only temporary'. Joos (1964) calls this category "temporary aspect", but he has a rather different interpretation of this notion. He says that temporary aspect "does not necessarily signify anything about the nature of the event...", but that instead "it signifies something about the validity of the predication, and specifically it says that the probability of its validity diminishes smoothly from a maximum of perfect validity, both ways into the past and the future towards perfect irrelevance or falsity." So "limited duration" for Joos is not to be taken as a property of the event or state described, but as a property of the validity of the predication. Nevertheless, even for Joos, the temporal nature of the category is basic: "The meaning of our temporary aspect is limitation of duration" [Joos' emphasis].

Recall the interpretation that was proposed for the contrast between (61) and (62) above in terms of the phenomenal/structural distinction. We should be able to construct examples where our approach makes different predictions as to appropriateness from those made by the alternative time-oriented approach. Consider (63) and (64).

(63) The statue of Tom Paine now stands at the corner of Kirkland and College, but everybody expects the new Administration to move it.

(64) The statue of Tom Paine is standing at the corner of Kirkland and College, and nobody thinks the deadlocked
City Council will ever find an appropriate place for it. Our analysis squares well with the well-formedness of both (63) and (64). The main difference between the two cases is that the first conjunct of (63) describes the current location of the statue as the properly designated one, while the first conjunct of (64) describes the current location pure and simple; the speaker, by choosing to use the weaker form in (64), strongly suggests that the location has not been determined by the proper procedures. The second conjuncts are entirely compatible with the first conjuncts in both cases; the duration of either state of affairs is independent of what we might call its administrative status.

The standard analyses in terms of temporal duration, on the other hand, predict that both (63) and (64) should be anomalous. In (63) the second conjunct implies that change is imminent, while the first conjunct is claimed to signal 'duration not limited': This ought to be incongruous. (64) is just the opposite case: the second conjunct expresses a prognosis of permanence, while the first conjunct is claimed to express 'limited duration'. Again, the sentence as a whole ought to be strikingly incongruous.

Such examples can be multiplied. Take the verb own. This verb is concerned with a 'social artifact'. That is, if John owns a car, this is a fact determined by a fact about the structure of the (social) world, not a fact about its phenomenal proceedings. Nevertheless, ownership can be very transitory. But no matter how hard we try to make up a scenario which is favorable to the temporal account of the progressive, say, by assuming that our friend John trades in his cars for new ones as fast as other people change their underwear, we still would not say, "John is owning a
a Mercedes at the moment."

A final example of this type. Organizational hierarchies have the same status vis-a-vis the phenomenal/structural distinction as do relations of ownership: they are structural by their very nature. We construct them; there are no data of observation from which we derive organizational hierarchies as a hypothesis about how the world works. They are stipulated, rather than discovered. But they can be unstable, and rapid change is possible. Consequently, a given state of affairs may be limited in duration, or, equally well, the probability of a predication about these things being valid may decay rapidly in both temporal directions away from the status quo. Again, no matter how much organizational reshuffling we assume has taken place and has yet to take place in our test scenario, a sentence like "The Interior Department is at present including the Consumer Protection Office" cannot report what is properly reported by "The Interior Department includes the Consumer Protection Office at present." This follows from our analysis, and is inconsistent with the 'temporal duration' analysis, however construed.

Having argued that the meaning of the progressive does not involve any temporal notions, we may confront the traditional analyses in terms of temporariness from a different perspective. We may ask what there is about the relevant data that makes the limited duration analysis such an attractive one, to the point of seemingly unanimous agreement on some version of it. At the root of an answer to this question lies the observation that it is simply quite natural to think of
structure as relatively enduring, and of phenomena as relatively
transitory. But this is a matter of statistical tendencies, not of entailment. A structure that was fixed one way at one time, can change or be changed, and so become fixed a different way. On the other hand, phenomena that just happen to be a certain way at one time, may happen to persist that way through time indefinitely. We have just given examples illustrating the fact that structure can be subject to change. An example of a phenomenon persisting would be "A photograph of Grandma and Grandpa is sitting on the mantelpiece, as it has been for 30 years," describing a family's living room.

Such examples notwithstanding, the usual case is that of structure being fixed, and phenomena being fluid. Only structures can be immutable, which is the way we conceive of the abstract structures of mathematics, logic, science. This is almost a tautology, for if a phenomenon proved immutable, we would conclude that there are structural grounds for this. This is the domain of inductive reasoning, in which repeated similar observations lead to the formulation of generalizations. The individual observations are at the phenomenal level; the generalization makes a claim about the nature of the thing observed, hence constituting a structural description. But, as we will see in connection with a later example, there is nothing that compels the inductive leap.

Let us consider how our position compares to the more traditional one in a specific case. How do we evaluate the following passage from Leech (1969), for example (the symbol '[+SITU]' represents a semantic feature ascribed to the progressive. E.W.)?
(65) [+SITU] signifies not just duration, but limited duration. In this it contrasts with the unrestricted present: 'I live at Highgate' and 'I am living at Highgate' differ precisely in the suggestion that in the second case the residence is temporary.

Leech's description of the contrast between the two sentences he gives may hit the nail on the head. Extensive questioning of native speakers may lead to the conclusion that they would all describe the contrast exactly as Leech does. Does this not suffice to show that Leech is plainly right? Our answer is neither a simple yes nor a simple no. We grant that Leech has given a very careful and accurate description of the minimal pair he mentions. We deny, however, that he has given a successful analysis of the contrast he describes so well. His analysis is essentially to say that the most palpable difference between the two sentences examined—the suggestion of temporariness vs. the absence of such a suggestion—is a direct reflection of a semantically primitive notion of 'limited duration' which is associated with the progressive, and is not associated with the non-progressive. But we have just argued with a handful of examples that this analysis fails in both directions. Neither is the progressive incompatible with extended duration, nor is the non-progressive incompatible with limited duration (see esp. (63), (64)).

Our analysis makes a much less concrete claim about pairs such as 'I live at Highgate'/'I am living at Highgate.' In regard to the non-progressive sentence it says that my living at Highgate is a structural fact, a fact that characterizes me. A critic might well say that the problem with this answer is that he cannot determine what it means, exactly, in contrast, e.g., to Leech's answer, which he finds attractively
precise. Our response to such criticism is twofold. We grant that, in a general way, the criticism is well taken, since our characterization of the structural/phenomenal distinction is not yet rich enough to permit a precise determination of how it is to be applied in any arbitrarily chosen case. On the other hand, however, we do not grant that the criticism is on the mark with its major thrust, which is the demand that a semantic analysis of a sentence be as specifically informative as are the competent native speaker's descriptions. In our view, a semantic analysis may underdetermine many specifics that native speakers routinely supply; in fact, some of the ways native speakers do supply such specifics have been systematized by Grice and others in the form of conversational postulates. So we would not consider our analysis deficient if it said no more about 'I live at Highgate' than something like 'his residence at Highgate characterizes the speaker'. Whether a residence has to be permanent, or how permanent it has to be, in order to characterize somebody may be left quite reasonably to the native speaker to evaluate, a task in which he may have to draw extensively on his knowledge of his culture.

We leave our digression on time now, and continue with our explication of the notions phenomenal and structural. The point we had arrived at was that the structure of an object may include relational properties as well as internal ones. We interpreted this notion of a structural relational property with reference to examples (61) and (62), which had to do with the location of a statue.

Just in case that discussion is taken to suggest that for a
location to be treated as a structural fact, there must be some social conventions that fix it, we will explicitly deny any such suggestion. For a statue to acquire a 'fixed' location that becomes part of the concept of that statue, some such social decision may be necessary. On the other hand, there are no special conditions to be satisfied for us to consider the location of the Rock of Gibraltar to be part of our concept of it. Nor should we find it surprising that there is this difference between readily movable human artifacts and large-scale features of our physical environment. So we expect to find (66), rather than (67).

(66) The Rock of Gibraltar stands at the entrance to the Mediterranean Sea.

(67) The Rock of Gibraltar is standing at the entrance to the Mediterranean Sea.

The location reported by (65) is not assigned, it is discovered. But it is certainly not capricious.

The next set of examples will extend the range of objects which are described in either structural or phenomenal terms to human beings. Here 'structural' is perhaps best translated into 'characterizing', and 'phenomenal' into something like 'observed behavior'. This translation has no theoretical status: it is meant as a heuristic guide. It is interesting to note here that with some judicious placement of emphases, we can trace our analysis, as it applies to people, back to Jespersen. He observes that in the progressive "very often a transitory
condition or behavior is meant in contrast to the person's habitual or real character" [emphases added. E.W.]. We will illustrate two other points as well with the examples to follow: first, that connotations of a form, or 'expressive coloring', can often be directly accounted for on the basis of the meaning of the form in conjunction with some simple assumptions about conversational principles, and secondly, that the analysis naturally extends to copular sentences with AP-complements.

The minimal pair we will be discussing is (68) and (69).

(68) He is polite.

(69) He is being polite.

First, imagine a situation where two friends are talking about a third person who both of them know well. This third person - call him Alvin - is considered by both of them to be an impolite and inconsiderate fellow. Suppose they see Alvin, however, giving up his seat to an elderly lady on the subway they are riding on. Then, we observe, the only appropriate way for them to comment on their observation is to say (69), "He is being polite"; to say (68), "He is polite," would be inappropriate, despite the evidence in front of their eyes.

This leads to a straightforward account of the fact that there is often an implication of insincerity (or play-acting) when adjectives (or predicate nominals) are used in the progressive. The point is a Gricean one: you say as much as you know, and if you know the person well enough to be able to judge him polite, then you wouldn't normally
weaken your statement by using the progressive. Therefore, if you do
use the progressive, there is an implication that you may know enough
about the fellow to be unimpressed by the evidence before you.

Take a similar situation on the subway, except the two friends
don't know the person they observe giving up his seat. Then it seems
either "He is being polite" or "He is polite" can be used felicitously,
though not interchangeably. When one says, in such a situation, "He
is being polite", then the suggestion of insincerity disappears, again
for the obvious Gricean reason. What use of the progressive here
suggests is simply that you are not prepared to extrapolate beyond the
evidence available to you. Exactly the opposite is true in the case
where you say "He is polite." This is not simply understood as a comment
on observed behavior, but makes a claim beyond, to the character of the
person.

The above example is an instructive one from the point of view
of showing how to maintain the distinction between the core semantic
contrast -- phenomenal vs. structural -- and any pragmatic inferences
that may be drawn from it. In the case of Alvin and his courtesy, an
inference of insincerity could be drawn, using Gricean pragmatic
principles and keeping the meaning of the grammatical form constant.

By maintaining this distinction, and holding the structural/
phenomenal contrast constant, we arrive at an explanation for the use of
the progressive in order to achieve a rather different conversational
effect. For example, we might say of someone,

(70) Old Lilly is always feeding the pigeons in the park.
The progressive is used in (70), even though the non-progressive is possible, as in (71).

(71) Old Lilly always feeds the pigeons in the park.

There is a sense of mild reproof in (70), a suggestion that the speaker is irritated, which is absent in (71). Why should this be? On any account of evidence and knowledge, it should be true that if we know enough to say (70), we thereby know enough to be warranted to say (71). Yet the speaker chooses the weaker, non-characterizing form (70), as if the speaker was avoiding actually characterizing Old Lilly in this way, and rests content with reporting what may be distasteful behavior to him. (Recall our earlier observation that the inductive leap cannot be compelled.)

Jespersen (1961) gives some rather subtle and complex examples of this general type (e.g. "I am always saying what I shouldn't say; in fact, I usually say what I really think.") and comments that "the combination of an expanded tense with always and its synonyms very often gives an emotional coloring to the sentence", and in some cases "there is a kind of mild reproach." But he also recognizes that "it is not always easy to apply the rule of 'frame-time' [Jespersen's main hypothesis about the meaning of the progressive. E.W.] to them." In fact, it is never easy.

This kind of example is also a source of problems for Leech's analysis. Leech considers sentences like "Death is getting nearer every day" and proposes that with them "the effect of [+SITU] [Leech's semantic feature for the progressive. E.W.] ... is to throw emphasis on the
ceaseless persistence of the process." He then concludes that in such
examples "the idea of 'persistence' ... is strong enough to cancel out
the second 'connotation' above, that of 'limited time extension'.
For this reason it may be best to consider it a separate meaning of the
continuous form, not merely a separate connotation of the same meaning."

It is somewhat puzzling that Leech notes, but for all practical
intents and purposes disregards, the fact that this 'inexorability'
sense of the progressive seems to arise only in the presence of adverbs
like always, continually, forever, which one should think express
'persistence' quite straightforwardly. Instead of trying to relate
'inexorability' to these adverbs, he locates the source of this sense
in the semantic feature [+SITU]. But the more serious problem with
Leech's analysis is that it makes it hard to understand how the meaning
of the progressive form could be learnable. The task that the language
learner is faced with on Leech's account is the assignment of two
contradictory meanings to a single grammatical form within a single
semantic domain. In terms of the argument we gave in Chapter 1 in
favor of the Univocality Assumption, this is the worst possible way to
design the syntax and semantics of a verbal category.

We again have insisted on the difference between an apt descrip-
tion of a class of sentences, and a successful analysis. Not the least
of the aims of this thesis is to show how the central question asked by
the generative grammarian, "How can human beings learn a language?",
can have genuine consequences for the evaluation of linguistic analyses
in the notoriously 'fuzzy' and intractable domain of the verbal
categories, while based on elementary assumptions about learnability.

2.4.3.2. In the preceding subsection we have made a first attempt to explicate the structural/non-structural ('phenomenal') distinction introduced in 2.4.2. Before we go on to a fuller range of examples, it will be useful to give at least a sketch of the logical character of the elements of Epistemic Status. The main purpose of this logical excursus is not to develop a rigorous logic of the domain of Epistemic Status, but to introduce some elements of the conceptual structure involved in a reasonably well-organized fashion, so that the remainder of our exposition, which is rather complex, can be made as clear as possible.

Recall that we introduced, in section 1.2, the rough notion of a 'bare propositional content' which we said was presupposed by all the semantic rules that belong to the TMA component. In section 2.3.1 we suggested that in something like the fashion that a common noun *dog* is related to a class of individuals, a propositional content is related to a class of events, processes, or states. Let us use as a general term for the latter class of abstract entities the term 'fact'. As we did before, we will let $P$ stand for a propositional content. Now we coin expressions of the form "the fact-type $P$", which would be spelled out as slightly ungrammatical phrases of the form "the fact-type of Walter filing the day's mail", "the fact-type of the water boiling", "the fact-type of the king being dead", etc., and which we will abbreviate as "$p^*$". Next we introduce, in the same informal spirit, a two-place
predicate, "x is structural with respect to y", which we might alternatively have expressed as "x is part of the structure of y", "x belongs to the concept of y". The first argument place is restricted to expressions of the type p*, while the second argument place may be filled by any term, without any restrictions as to type. The expression, "The fact-type of John being generous is structural with respect to John" illustrates the kind of propositions we have arrived at at this point. The next step is to embed this type of proposition in turn in a fact-nominal, deriving expressions of the form "the fact that the fact-type of John being generous is structural with respect to y." Finally, we leave the second argument position of the embedded proposition open, and we get "the fact that the fact-type of John being generous is structural with respect to y." This last kind of expression we symbolize as "[s(p*,y)]", where the brackets indicate the fact-nominal, s the two-place predicate "x is structural with respect to y", and p* the expression "the fact-type P"; y is a variable.

Expressions of the form "[s(p*,y)]" are terms that we will admit into the first argument place of our predicate DUR(x,y). When we introduced this predicate we were concerned with an analysis of events, and gave its approximate meaning as "x occurs during y". Since facts do not 'occur', but rather 'hold' or 'obtain', we amend the above wording to "x occurs, or obtains, during y".

The second major predicate in the domain of Epistemic Status is 'x is phenomenal with respect to y'. To avoid unnecessary confusion, we will use the letter n (for 'non-structural') to symbolize this predicate,
thus obtaining expressions of the form "[n(p*,y)]".

As stated at the outset, this excursus into symbolization is entirely subordinated to the aim of achieving greater ease of exposition by making more perspicuous the main features of our conceptual scheme. We now go over these main features one by one, pointing out in the process how our symbolization reflects them.

1. The main point is that we take Epistemic Status to say something about the entire proposition expressed by the sentence. The non-progressive in its epistemic sense tells us roughly this: "The sentence you are looking at is a description of structure, not of phenomena." This point comes out quite clearly in [s(p*,y)], where p* (which has in it the propositional content of the sentence) is an argument to which the predicate s ascribes structuralness. Epistemic Status builds considerable logical structure.

2. The second point flows from the first one. "The sentence you are looking at is a description of structure" is not quite what we have in mind. A sentence cannot just be a description of structure, it must be the description of the structure of something. We have treated all the examples in 2.4.3.1 as though there were a simple rule which determines that it is always the referent of its surface subject t\at a non-progressive sentence describes t\he structure of. This impression is misleading, however. In section 2.4.3.3 we will present a large number of examples whose successful analysis along the lines of our proposal depends on there being no such grammatical rule. We will conclude that sentence grammar does not determine what entities a structure-describing
sentence ascribes that structure to. In the expression "[s(p*,y)] this is indicated by the variable in the second argument place.

3. The third point is that what is asserted as obtaining, or being true in such a sentence is not the propositional content itself, but rather the predication of structuralness. This, too, is captured by "[s(p*,y)]"; the notation clearly suggests that the brackets have 'scope' over the asterisk. Therefore, when we say "DUR([s(p*,y)], t₁)", we assert that what obtains during the time interval t₁ is the fact that p* is part of the structure of y; we assert nothing about the propositional content directly. This third point differs from the first, it should be noted. Point one about Epistemic Status predicking something of an entire proposition could be true if the resulting logical structure were something like "p is true, and p is structural". Here we insist that we unequivocally get something more like "that p is structural is a fact". This difference has important consequences, and we will explore its consequences in detail in Chapter 4.

2.4.3.3. In section 2.4.3.1 we gave a range of illustrative examples, varying them along such dimensions as the category of object involved (physical objects, non-material objects, human beings), and a variety of connotative effects (temporariness, sincerity, emotional coloring). One dimension, however, we have kept constant; in all examples it was the subject of the sentence that provided the entity that was described in either structural or phenomenal terms. But we have just pointed out in the above logical excursus that this is a fact about the arrangement of our exposition only, not a consequence of our analysis. Our actual
position is that an appropriate choice for the variable $y$ in $[s(p^*, y)]$ is fixed by the speaker within the context of a discourse; it is thus not a concern for a sentence grammar.

A position of this general type has been advanced recently by Lasnik (1975) in a characterization of the semantic status of non-indexical personal pronouns. According to Lasnik the pronoun he just means a male human being, and does not refer or co-refer to any particular person as a matter of sentence grammar. Reference is assigned by the speaker (subject to a non-coreference condition) in accordance with pragmatic conditions of cooperation and the like. Lasnik's general point is that this is no less true of noun phrases like John, although it may seem to be because the number of possible referents has been reduced drastically.

This is, then, a natural point for us to broaden our range of examples and to include cases where something other than the subject of the sentence is described structurally or non-structurally.

There is a rather ill-defined class of somewhat exotic uses of the non-progressive that have received some attention in the recent literature. The unifying property of these uses is, in traditional terms, that they allow simple present tense forms to be used non-generically. Braroe (1974) notes that "the non-progressive PRESENT of verbs of the kind that can show a formal contrast progressive vs. non-progressive given PAST, can apparently be used non-generically in contexts where the speaker reports on an event to which only he, and not the hearer, has sensory access". This use, call it the telecommunicative use, is often illustrated by reference to sports commentaries, as in (72).
Brown chases a well-hit fly ball to deep center, and finally runs it down at the warning track.

Joo (1964) gives a very different general characterization of the specific uses of the non-progressive. He proposes that specific non-progressives are in order when "sharp focussing on a single time is part of the message; allowing the validity to spread out within a temporary aspect frame, however narrow, would then spoil the message: asseveration and demonstration uses of the generic aspect are obvious cases." Leech (1969) takes almost exactly the same tack as does Joos. He allows that the non-progressive Present can be interpreted, besides the options of an Unrestrictive Present and a Habitual (Iterative) Present, as an Instantaneous Present. Of the latter he says,

The present tense interpreted in this way refers to an event psychologically perceived as taking place in its entirety at the moment of speech, or more briefly, 'single-event-now'. The instantaneous present is restricted to certain contexts (occurring, for example, in sports commentaries, cooking demonstrations, and ceremonial utterances) because very rarely are happenings brief enough to be thought of as being started and finished in the very instant of utterance.

We propose a very different account of these uses of the non-progressive. Take the case of an experiment in chemistry class, which the teacher accompanies with a commentary. The teacher's commentary is not a description of what he observes himself doing; rather, it is a description of the structure of the experiment that he is performing. Consider this example:

And now I take the flask of sodium nitrate, and add its
contents to the mixture in this beaker. Now I place the beaker on the bunsen burner, and heat its contents to a vigorous boil...

Suppose we imagine him continuing with (74)

(74) Now I take my handkerchief out of my pocket and wipe the perspiration off my forehead.

The effect this would have would be to make the actions described in (74) part of the experiment. An attentive student would conclude that it is perhaps important that no little drops of sweat drip into the boiling mixture, and, on re-executing the experiment, he would bring everything to a boil, reach for his handkerchief, and wipe his forehead dry.

One might think that this is merely a matter of 'relevance' in some rough and ready sense of the word. The experimenter wouldn't mention things that aren't important to the proceedings, and indeed, if he turned (74) from the non-progressive into the progressive and uttered it at the same point of his demonstration, the entire sequence would sound quite unnatural.

But let us consider the scenario underlying (73) and (74) from a different perspective, say, the perspective of the telecommunicative situation. We assume that the speaker gives a description of the experiment to a listener who can not see what is going on. We further assume that the speaker knows the experiment, i.e. he can tell what's really part of it and what's not. And finally, we assume that the
telecommunicative situation is such that it exerts pressure against long pauses. We might say that a cooperative speaker would keep up a continuous stream of narration so that the listener need never even begin to suspect that something is being withheld from him.

Now we can make some unequivocal claims. If the speaker reports

(75)

(75) And now he takes the flask of sodium nitrate, and adds its contents to the mixture in the beaker. Now he places the beaker on the bunsen burner, and heats its contents to a vigorous boil. Now he takes his handkerchief out of his pocket and wipes the perspiration off his forehead.

everything in (75) describes how the experiment goes. If, on the other hand, the speaker reports (76)

(76) And now he takes the flask of sodium nitrate, and adds its contents to the mixture in the beaker. Now he places the beaker on the bunsen burner, and heats its contents to a vigorous boil. Now he is taking his handkerchief out of his pockets. Now he is wiping the perspiration off his forehead.

there is a break in the action, as it were, at the point where the chemical mixture has been brought to a boil. The part about the handkerchief is still relevant to the report here, because we have attributed to the hearer a need to know about everything that's going on. But the speaker, who, as we have stipulated, knows the experiment, must shift
from the structural to the phenomenal mode if he doesn't want his report to be misleading.

Before drawing any conclusions, let us consider a third case. We again assume the telecommunicative situation, including the hearer's need to know. But now we remove the assumption that the speaker knows the experiment; even more, we assume the speaker doesn't even know that what he is observing is an experiment. Then we have removed the object whose structure might be described, and only the purely neutral phenomenal description, as in (77), should be appropriate.

(77) And now he's taking a flask off a shelf, and he's pouring its contents into a beaker. Now he's putting it on a bunsen burner. Now the stuff is boiling. Now he's taking his handkerchief out of his pocket. And now he's wiping the perspiration off his forehead.

In (77), as in (75), everything is on a par. But while in (75) the parity was in terms of the structure of an experiment, the parity in (77) is at the level of observed actions.

Now for some initial conclusions. The set of examples from (73) to (77) shows that the progressive/non-progressive alternation is independent of the telecommunicative/non-telecommunicative contrast. They also show that neither the progressive nor the non-progressive are sensitive to the question of what happenings may be brief enough to be thought of as being started and finished in the very instant of utterance. But most importantly, they show something positive as well. They show
that our account of the progressive/non-progressive alternation in terms of a semantic phenomenal/structural distinction extends quite naturally to cases where the object whose structure is being described is not represented by any NP present in the sentence, but is drawn from the discourse.

This last point is very important. For how does one know when one can consider some object not mentioned in the sentence as the thing that the sentence is really about, i.e. its semantic topic? As it turns out, and as everybody who has looked at this matter agrees, the contexts that can provide such a semantic topic are very restricted, and the object that serves as the semantic topic highly structured. Experiments are a very special type of happening, not just any kind of actions strung together. The same goes for sports games, cooking demonstrations, magic tricks, ceremonies, rituals. They are all well-defined contexts. It is this kind of context that, e.g. , fixes the difference between a baseball player just running from first to second base, and him stealing second base. This last comment forms a natural bridge to the second point to be made --- these special contexts typically have a good deal of internal structure. Games are paradigm cases, with their inventory of permissible moves and with the definition of a declared objective of the game, in many cases. This point about the structuredness of the contexts involved is central to our account, for where there is no structure, one cannot give a structural description. On any other account, the restrictiveness with regard to context is a separate fact to be noted, but not particularly connected to the
meaning of the form. At least, it hardly seems likely, for example, that the contexts in question are just those in which the action is so fast and furious that the spoken word can barely keep pace.

Another important example of an abstract structure in this special context category is the example of a plot. 'Plot' is our name for the internal structure of plays, movies, operas, and the like. When McCawley (1971) notes, with reference to example (78)

(78) I've just seen a movie in which someone steals the crown jewels.

"that this is one of the few contexts in which a non-habitual present tense is possible in English," the force of his comment is that (78) represents an oddity, that it is semantically aberrant. But (78) is not an example of poetic license; on the contrary, the sentence would be very odd indeed if is r*ealing were to be substituted for steals. In our terms, there is perhaps something unusual about sentences like (78), viz. that their topics are not overtly represented in them, as a rule; but semantically, these cases are quite regular.

Notice, in this connection, that films and photographs differ in just the respect we are interested in here: films have a plot, photographs do not. Thus, there is structure of a kind to describe in the one case, which is simply not there in the other. We predict use of the non-progressive in the former case, and use of the progressive in the latter - a prediction which is borne out, as shown by (79) and (80).

(79) a. In Bill's snapshot Mary is riding a horse.
b. In Bill's snapshot Mary rides a horse.

(80)  

a. In Bill's latest movie Mary milks a goat.

b. In Bill's latest movie Mary is milking a goat.

Next, we use the notion of a plot to reinterpret the facts covered in the following remarkable passage from Leech (1969), who, as we mentioned above, subscribes to the idea that the uses of the non-progressive under discussion here are felicitous, provided the events described can be perceived as a complete unit in the instant of speech:

The instantaneous present ... is stylistically restricted, because for obvious reasons it has dramatic overtones which do not suit it to ordinary colloquial speech or discursive prose. Compare I am opening the door with I open the door. Where the former is an unremarkable description of a present happening, the latter is more dramatic, and seems to demand the accompaniment of an exclamation mark or some spectacular gesture. (p.139)

On Leech's account, the reasons for the dramatic overtones are not really that obvious. Our example (73), dealing with a chemistry teacher and an experiment, which falls under Leech's Instantaneous Present, could involve a very bored teacher and a very bored class, and be devoid of any suggestion of drama. What is the matter with Leech's I open the door is not that it is dramatic by its very meaning, but rather that it is so undramatic that it is very unusual without a context to place it in. The sentence means, among other things, that it is a description of the structure of something. The speaker can readily be eliminated as that something. So the pragmatic search is for some a to substitute for the x in the structure of x. Plots are a possible candidate, but --- or so we interpret Leech's remark about 'some spectacular gesture' --- if the
fellow means to be play-acting, he ought to act as though he were.

This is analogous to the point we made earlier in connection with the interpretation of the progressive when it co-occurs with adverbs of the _always_-class. The logic of the theoretical approach pervading so much work on natural language semantics seems to be this: Suppose form A has the general meaning M. But in the presence of formative F, or when used in context C, A does not seem to mean M. Rather, it appears to mean N, where N need not be related to M in any particular way. Solution: A has a subsidiary meaning N, limited to context C, or to sentences containing F. The merit of Leech's work, and the reason it serves so frequently as a foil to our own analysis, is that the scope of Leech's work is quite comprehensive, and that problems of the kind sketched above are dealt with explicitly.

We have treated the steps of an experiment or a demonstration as being constitutive of their structure, and we have proposed that the plot elements of a dramatic work be viewed analogously. Let us return now to the case of sports commentaries and determine what elements of structure they deal with. Consider the following contrasts between things we imagine being said by a sportscaster:

(81) Vail slides head first into third base, and the umpire calls him safe.

(82)#Vail is sliding head first into third, and the umpire is calling him safe.

(83)#The sun shines out here at Fenway Park, and the crowd hums with excitement.
The sun is shining out here at Fenway Park, and the crowd is humming with excitement.

What do (81)/(82) describe so that the non-progressive version (81) is clearly more natural? And what do (83)/(84) describe so that the opposite state of affairs prevails? The difference is that (81) deals with the moves of the game, with the structural building blocks that define that particular game while (84) does nothing of the sort: it is concerned with phenomenal circumstances surrounding the game. We see that this case has much in common with that of experiments or movies. The individual actions or events are not just significant as occurrences in their own right: they gain added significance by being embedded in a sequence of actions and events, which itself is an abstract whole. These abstract entities are psychologically real enough; there are words that name them.

Extending our analysis, we consider next the use of the present tense that is traditionally referred to as the 'historic present'. The historic present is widely viewed as being due to some kind of poetic licence. The writer, it is held, aims at making his account of some past event more vivid by presenting it as a currently ongoing event. Interestingly, not all linguists are in agreement. Joos (1964), e.g., observes astutely that this account of the historic present is unconvincing, not on psychological, but on empirical grounds. He notes that a narrative in the historic present is systematically different from a description of things going on at the time of speaking.

Joos juxtaposes two very similar prose passages (Joos 1964, p. 131)
which we include here as (85) and (86).

(85) While Superintendent Hannam is examining the cupboard, the doctor walks slowly across the room to an identical built-in cupboard on the left-hand side of the fireplace, opens the centre compartment and puts his hand inside, then he takes out two objects which he puts inside his left-hand jacket pocket.

(86) "Hannam is examining the cupboard. Now the Doctor is walking slowly across the room to an identical built-in cupboard on the left-hand side of the fireplace. He is opening the centre compartment and putting his hand inside; now he is taking out two objects and putting them into his left-hand jacket pocket.

According to Joos, a speaker of English would have little difficulty recognizing the fundamental difference between (85) and (86): only (85) can be a narrative, i.e. tell a story, and only (86) can be a running commentary on what is happening before the speaker's eyes. Joos gives a very revealing summary of this and of related observations (for 'generic' simply read 'non-progressive', and for 'temporary' read 'progressive' - E.W.):

This can be stated as a simple transformational rule, thus: Every generic aspect of here-and-now reference remains generic with real past reference; but the temporary aspect is changed to generic aspect for each event that advances the plot of the narrative and remains temporary for each event that is rather background to the plot-advancing events without itself advancing the plot. [the emphases are Joos']

Now notice that the difference Joos underscores is exactly analogous to the contrast we noted with examples (81)-(84): the 'moves of the game' correspond to the 'plot-advancing events', the phenomenal circumstances' correspond to the 'background'. The conclusion within our
framework is inescapable: the 'historic present' is, contrary to Leech (1969), for example, not a somewhat parasitic device "whereby past events are understood to be described, as a kind of dramatic licence, as if they were happening at the present time". Rather, it is what grammarians call it when somebody describes the structure of a narrative. And the facts are simply these: to describe structural plot elements, the non-progressive is used. As a separate grammatical category, the historic present has the same status in the grammar of English as the abovementioned category labelled 'telecommunicative': none.

If the above analysis of the historic present is right, then Jespersen is wrong in saying that "the present tense is used in speaking of the past." Sentences of the kind in question do not speak of past time; they speak of abstract entities called plot lines or narratives, in the present time. But Jespersen is also arguably wrong when he says that "the present tense may be used in speaking of some future time." Rather, the sentences in question do not speak of future time; they speak of abstract entities called schedules or programs.

Consider sentence (87), for example:

(87) The last train to Winchester leaves at 7 p.m.

It is often supposed that (87) is a 'generic' sentence, and that the non-progressive is used because it describes repeated events. Actually, the repetition is irrelevant; it's knowing the arrangement that predetermines actual events that matters. So we might well say, on a trip to Europe,
(88) Look! Our itinerary says that we fly to London tomorrow, and then go to Paris by train on Saturday.

Repetition is just as irrelevant here as it was in the case of the description of the structure of games and plots. The structure of a baseball game has a unique enactment: that game. The structure of the plot of a stage play may have no enactment, or hundreds of them. How often an abstract structure is realized has nothing to do with whether that structure exists.

True, it is in the nature of train schedules that we would lose faith in them in the trains listed on them never ran, while we wouldn't say that a play couldn't really be a play if it doesn't ever get performed. This has nothing to do with the meaning of the non-progressive, however. Whether or not some real-world consequences follow from the existence of a schedule might be a topic in a Theory of Schedules, but not in a theory of verbal categories.

In yet another respect much of the discussion about the so-called 'programmed future' belongs more properly into a theory of schedules than into a theory of grammar. It is a rather well-known fact, pointed out by Lakoff (1971), for example, that paradigms like (89)-(91) are found quite regularly with the programmed future.

(89) The Yankees play against the Red Sox tomorrow.

(90) The Yankees play well against the Red Sox tomorrow.

(91) The Yankees beat the Red Sox tomorrow.
This kind of paradigm becomes a problem of grammar if one assumes that all of (89)-(91) are derived from structures including the future marker will by a rule of will-Deletion which apparently is subject to some complicated conditions on applicability. The question shifts wholesale when we ask instead why neither the event in (90) nor the one in (91) are schedulable under normal assumptions. The rule about schedules seems to be --- and this is not a statement about grammar --- that it is fine to schedule activities, but not so sensible to schedule the outcome, success or failure, of those activities. By scheduling something, one does not presuppose it will take place, but by scheduling an outcome, one does presuppose that the event itself will take place. Thus the sentence "The Rutland Tigers receive the A.C. Murphy Trophy tomorrow" is fine if the Rutland team has already won the championship, but not if we only know that the winner of tomorrow's deciding contest will get the trophy, and are furthermore convinced the Rutland team will defeat their opponents. An adequate theory of schedules will explain all of this and yet be sophisticated enough to make allowance for the fact that there are cases, such as train schedules, where it is not only in order to fix in advance that a train will travel from Boston to New York, say, but also fix when it will arrive in New York.

Jespersen has collected some nice examples that are practically tailormade for a theory of schedules. Jespersen says that "very often... the present tense alternates with shall or will", apparently making the erroneous assumption that this is a case of free alternation. His examples include (92) and (93).
(92) We dine together tomorrow and next day by invitation; but I shall alter my behavior to him, till he begs my pardon, or else we shall grow bare acquaintance.

(93) Tomorrow morning I leave England. You will never see me again. This is the last time I shall ever look on you.

Dinners and departures are paradigm cases of schedulable events. Barely getting to know someone, or never seeing someone again, just aren't the kinds of things people put on a schedule.

It should not be thought that these repeated references to a theory of schedules are not serious just because 'theory of schedules' sounds a little funny. No, just as it takes a theory to tell us whether a linguistic phenomenon is syntactic or semantic, whether it is a fact about the base rules or about the way transformations work, etc., so it takes a theory to tell us whether a fact is a fact about grammar at all. The observation that our semantics of the non-progressive does not directly "handle" the distinction between, say (89) and (90), implies nothing as to the adequacy of our theory. If it were to be claimed that the contrast between (89) and (90) is a semantic fact, such a claim could not be supported just by intuition, but only by an explicit semantic theory that shows what kind of semantic fact it is.

The natural companion to our discussion of the 'programmed future' is an examination of the 'futurate' use of the present progressive. Our characterization of the phenomenon category within the domain of Epistemic Status is not yet rich enough to say whether, or how, this
category figures in the semantic analysis of sentences like (94):

(94) We are having some friends over to our house this evening.

But from the Univocality Assumption we can deduce that there ought to be a semantic account of (94) in terms of the concepts of Epistemic Status.

We approach this problem by considering the minimal pair (95) and (96).

(95) Bill flies to Buffalo tomorrow.

(96) Bill is flying to Buffalo tomorrow.

If this is a semantically minimal pair, similar to (61) and (62), or (68) and (69), then there ought to be something about (96) that is capricious, fluid, unstable, accidental, which can become fixed in some kind of general framework, in which case (95) is appropriate. Given that (95) describes a schedule, the question becomes: what is it that schedules fix? Do they fix future events? Well, not really, since putting something on a schedule does not in and of itself have any effect on a future event. It is for this reason that simple truth-value intuitions flounder somewhat in the case where we assert (95) today, and it turns out tomorrow that the flight Bill was to be on has had to be cancelled on account of thick fog, and Bill never gets to Buffalo. Does this mean that (95) was false? One would be much happier to call (35) false if we had gotten mixed up about dates and said (95) although the trip was actually booked for the day after tomorrow. Moreover, our
analysis of such sentences as (95) already makes an unequivocal claim
here. (95) is true just in case it describes faithfully what the
schedule is, regardless of whether the schedule is ultimately adhered to.

So if schedules do not actually fix future events (what Leech
calls 'future-as-fact'), what do they fix? We will say that they fix
intentions. They shift intentions from a plane of instability, from
the private sphere of the intender, to a conventional framework. Set-
ting up a schedule is a way of relinquishing free choice. Having an
intention is not.

The conclusion that follows from this is that (96) is properly
to be analyzed as a description, valid for the time of the speech
event, of an intention. This is not quite accurate yet, for in fact we
need not postulate any such entities as intentions to serve as the things
that are described in terms of the phenomena emanating from them.
Rather, the intentions are the phenomena, and they emanate from human
beings. This means that phenomena can be understood as objects of
introspection, as well as in the more usual way, as objects of what we
might call extrospection. Intentions, along with emotions, hopes,
wishes, thoughts, make up the internal phenomenal contents of human beings.
This access to an inside world of phenomena is limited to human beings,
and is extended, by analogy, to other animates. Non-animates lack this
dimension, and with them, phenomena are limited the way we ordinarily
think of phenomena, to sensorily perceptible things.

This is not an a priori truth; it is a hypothesis. The claim
implicit in formulating this hypothesis is that the semantic category
phenomenal means this, universally. Consider in this connection the following summary of the semantic conditions for the felicitous application of the rule of Subject Postposing in French, given by James C. Atkinson in a careful study of this topic:

...the phenomenon will be one immediately perceptible to the senses: something that can be seen, heard, smelled, or felt --- all those realities occurring in the immediate physical environment which ... we are privileged to experience directly without the mediation of any speaker ... There are, however, exceptions, the most important being descriptions of emotional or intellectual processes, which (like the experience of taste) obviously cannot be perceived by another. (Atkinson (1972), pp.16,17)

This passage could almost serve as a succinct statement of the basic meaning of the category phenomenal, except for the ascription of exception status to the inclusion of emotional and intellectual processes: the exceptionality is only apparent.

We have maintained throughout our discussion that the entity of which the structural or the phenomenal description is given may or may not be mentioned in the sentence. This is true here as well. An example that is most plausibly analysed as describing the subject of the sentence is (97).

(97) Can I offer you another cup of coffee? No, thanks.

We're leaving now.

But the case where the relevant intentions are not those of the subject of the sentence can also be illustrated quite readily. So, the head of a household could tell the movers,
(98) The refrigerator is going over there for now.
pointing to a corner of the kitchen. Or a mother may tell her kids,

(99) You're going to bed now --- immediately.

This latter type of case seems to require that the intender have
enough authority for his or her intentions to carry some weight, so
that the actual realization of the intention is reasonably assured. This
observation in turn seems to be related to the fact that in cases where
there is no mediating chain between intention and execution, as in (97),
there is often a suggestion of imminence: the intention is concerned
with something close at hand, so that its fulfilment is reasonably
assured. But this is a tendency, not an iron-clad fact about this use
of the progressive. As Leech (1971) points out, a sentence like
"When I grow up, I'm joining the police force" is not impossible.

While this matter of imminence is of some interest, it does not
test our analysis. There is, however, the following simple test of
our analysis that suggests itself: If we are right in relating the
futurate uses of the progressive to present intentions, then there ought
to be future events outside of the scope of the progressive, as
follows: The less an event can be thought of as being determinable by
intentions, the more unacceptable we should find futurate progressive
sentences referring to such an event. Leech, whose analysis of this
particular use of the progressive comes very close to ours in its
predictions, notes that sentences like (100) and (101),

(100)#The sun is rising at 5 o'clock tomorrow.
It is raining tomorrow.

are indeed absurd, since they deal with events thought of as being determined by natural law.

This brings us to the end of section 2.4.3.3 The main concern of this section has been to analyse a considerable range of cases in all of which something other than the surface subject of the sentence constituted the entity structurally or phenomenally described. We have argued that a number of special semantic categories associated with the Present Non-Progressive, such as the Instantaneous Present (for sports commentaries), the Historic Present (for storytelling), or the Programmed Future (for schedules), are spurious. The semantic contribution made by the Non-Progressive is constant across all these cases; The differences are due not to facts of the grammar, but to facts of the world.

But this is not to be taken as saying that the categories of Epistemic Status are impervious to all grammatical facts. In the next section we will consider a number of examples that suggest that there is some interaction with lexical semantic properties of verbs.

2.4.3.4. Consider the difference between (102) and (103):

(102) The Red Sox were winning in the eighth, in which Rice hit a two-run homer.

(103) The Red Sox won in the eighth, in which Rice hit a two-run homer.
The intuitive difference between the two sentences is that (102) could be a report simply of how the scoreboard read at some point during the eighth inning, while (103) is an analysis of the entire game, pinpointing its decisive stage. In both sentences we are talking about a baseball game, and the sentences clearly make different claims, yet the distinction made earlier (in connection with (81)-(84)) between 'moves of the game' and 'surrounding circumstances' appears not to help at all here. There is a clear sense of more information backing up (103) than (102), and there are entailments derivable from (103) that do not follow from (102): only from (103) can we infer that the Red Sox were victorious. But just how this is related to structure, or to phenomena, is less than obvious.

The solution to this problem starts with the observation that the verb win stands in an entirely different relation to the description of a baseball game than does the verb slide, or even the verb phrase slide into second base. The former does not, while the latter does, describe part of the observable proceedings in the ballpark; but winning describes the outcome, or the results, of those proceedings. The question of structure vs. phenomena shifts wholesale. In the case of an action, the question simply is whether it is part of some larger sequence of actions or not: if yes, we tag it 'structural', if no, we tag it 'phenomenal'. In the case of 'winning', the question cannot be whether it is part of some larger organized whole, for it is that by definition. If winning is concerned with some target state resulting from a well-defined activity, what could it mean to ask whether a given
instance of that state is 'structural' with respect to that activity? The only way sense could be made of such a question would be if there were 'non-structural' instances of the target state. As vague as this is, let us ask: What is the target state a team has to attain in baseball in order to win a game? Answer: It has to have scored more runs than the opponent. But this is not enough: the additional condition is that it be in that happy state at the official end of the game.

We propose -- somewhat tentatively --- that the lexical entry for the verb win (and, likewise, for other achievement verbs like reach, discover, arrive, find, defeat, etc.) is a complex one, specifying both that there is an arrangement of circumstances which describes the target state, and that there is some criterion by which, out of a number of possible realizations of the target state one is picked as decisive. What 'decisive' means in each case, i.e. what counts as success, depends on a selectional property of each verb; thus with winning the target state is embedded in a contest, with finding it is part of a search, with reaching it is part of a journey, etc. But the important thing to take note of is that the two parts into which the meaning specification of achievement verbs is divided on this proposal, differ just along the lines of the structural/phenomenal distinction. In introducing this distinction we said that phenomena were the objects of observation, and structure was the object of knowledge. This fits the case under consideration: the arrangement of circumstances which makes up the target state is an object of observation, while the criterion
that fixes when such a target state \textit{counts} is an object of knowledge. This conception will now be confronted with a set of concrete examples. Consider sentences (104)-(107):

(104) Going into the final turn, Lucky Lady was winning by three lengths.

(105) Going into the final turn, Lucky Lady won by three lengths.

(106) Thanks to a tremendous finish, Lucky Lady ultimately won by three lengths.

(107) Thanks to a tremendous finish, Lucky Lady was ultimately winning by three lengths.

(104) merely reports what meets the eye: one of the horses is in the phenomenal configuration which, when paired with a structural fact about the race --- the significance of the winning post --- counts as winning. (106) reports that both that phenomenal configuration and the structural requirement for winning, namely that the configuration in question hold at a designated point, were satisfied. (105) appears to report that too, but contradicts satisfaction of the structural requirement with its adverbial \textit{going into the final turn}; hence its incongruousness. (107) is incongruous for the opposite reason: the force of \textit{finish} and \textit{ultimately} is to suggest that the winning post was passed, the force of the progressive is to suggest that only the condition of \textit{being ahead} was met.
This account of (104)-(107) appears adequate, despite its informal nature. We can gain an even deeper understanding of the paradigm in (104)-(107) by contrasting it with the standard paradigm for another achievement verb -- find -- which never allows the progressive when referring to a single event. With find, acceptability judgments pattern as in (108) and (109):

(108) #When we extended our search to the nearby forest, we were finding little Tom.

(109) When we extended our search to the nearby forest, we found little Tom.

We suggest that the phenomenal part of finding consists in recognizing the thing or person looked for. To call an act of recognition phenomenal is not unreasonable. We had occasion once before to hypothesize that perceptual and psychological processes were phenomena, albeit internal ones. Our account of the futurate use of the progressive was based on this hypothesis, and we adduced Atkinson's semantic analysis of the semantic effects of Subject Postposing in French as an extremely interesting piece of independent corroboration.

We further suggest that the structural part of finding consists in the stipulation that only the first act of recognition relevant to the ongoing search for x is to be counted as finding x. This makes it clear why there are no progressive uses of find with respect to a single search. For if a given act of recognition is indeed the first, the structural description of the search must be chosen (that is, the
speaker can say no less than what he knows). If a given act of recognition is not the first, however, then the structural condition for finding is not met; but a merely phenomenal description of the search is also inappropriate, because there is, at that point, no longer any search to describe, by definition. Contrast this with the case of a race, where the phenomenal condition of being ahead obtains while the race is still on.

The last class of examples we will draw on in order to elucidate the phenomenal/structural distinction involves perception verbs. We will start with yet another minimal pair:

(110) Little Mary saw ships on the horizon long before her Grandpa did.

(111) Little Mary was seeing ships on the horizon long before her Grandpa was.

The judgments about (110) and (111) that are to be explained are that (110) simply suggests that Little Mary sees a lot better than her Grandpa, while (111) suggests that both are prone to hallucinations, but especially Little Mary. The puzzling fact is that the progressive suddenly takes on an irrealis dimension; the ships in (111) aren't real, nor are the noises in I'm hearing all sorts of strange noises. There is no contradiction in saying He was seeing a big white rabbit that didn't exist, or so it seemed from his behavior. But a contradiction can be derived from #He saw a big white rabbit that didn't exist, or so it seemed from his behavior.
We propose an analysis quite close in spirit to that advanced for (104)-(107). What we need to do is assume that the perception verbs (and here we are only talking about the so-called stative ones) can be analysed as referring to the following situation: (i) the perceiver has a sensory experience, and (ii) that sensory experience is physically related to some feature of the environment. The part in (i) is phenomenal: sensory perception is on a par with emotions and intellectual activity in this regard. The relationship referred to in (ii) is something hypothesized, not something experienced: thus it is structural. The part in (ii) is what Gruber (1966) analyzes in Look and see, when he talks about paraphrasing John sees a cat by John's gaze goes to the cat, while the part in (i) is concerned with what happens when John's gaze actually meets the cat.

Suppose now we say that (ii) refers to the structural condition that must be satisfied for (i) to count as an act of perception. Then the non-progressive, describing the structure of an act of perception, entails that the object of perception exists. The progressive, on the other hand, deals only with the phenomena of perception; thus, by a Gricean inference, we conclude that the structural conditions for those phenomena are not satisfied. But a sensory experience that is not in the normal way related to a feature of the environment is a hallucination; in particular, there is a strong implication that there is no 'object of perception'.

The set of examples discussed in this section differs from the examples previously used in one important respect. We had not assumed
before that the categories 'structural' and 'phenomenal' were referred to in the lexical entry for a verb. When we analyzed sentences (61)-(64), and (66)-(67), all involving the main verb stand, our discussion was carried on under the presumption that the verb was neutral vis-a-vis Epistemic Status, and that considerations of Epistemic Status were only introduced at the stage at which the auxiliary be + ing is interpreted. Schematically, we may represent the difference between a verb like stand and a verb like win as follows, letting M stand for the lexical meaning of the verb, and $S_i$ for a semantic feature that enters into M in some way:

\[(112) \quad \text{a. } \text{stand:} \quad M = (S_1, S_2, \ldots, S_n) \]
\[(112) \quad \text{b. } \text{win:} \quad M = \begin{cases} M_{\text{structural}}: (S'_1, S'_2, \ldots, S'_n) \\ M_{\text{phenomenal}}: (S''_1, S''_2, \ldots, S''_n) \end{cases} \]

The main point is that the entry for win is more highly structured than that for stand. Interesting questions arise when we consider the interaction of semantic rules mapping into Epistemic Status with sentences containing verbs inherently specified for Epistemic Status. This will be the topic of section 2.4.3.5.

2.4.3.5. Implicit in the schematic sketch of (112) is the existence of two more possibilities: verbs whose meaning is exhausted by structural meaning components, and verbs whose meaning is exhausted by phenomenal meaning components. Both possibilities are in fact realized in English. The second case, exclusively phenomenal meaning, is the
more interesting, and we will discuss it first.

2.4.3.5.1. In our discussions of win and find we offered specific proposals as to what was to be considered their phenomenal meaning content. In the case of winning, the phenomenal part of the meaning can be described as being ahead, while in the case of finding, the phenomenal meaning is something close to recognizing. But if being ahead and recognizing pick out the phenomenal part of winning and finding, respectively, it follows that their own meaning is exclusively phenomenal.

Consider now the following paradigm:

(113) Lucky Lady{is \textit{was}} ahead (by three lengths).

(114)*Lucky Lady{is \textit{was}} being ahead (by three lengths).

(115) Mary{\textit{doesn't} \textit{didn't}} recognize John without his beard.

(116)*Mary{\textit{isn't} \textit{wasn't}} recognizing John without his beard.

The facts are plain: both verbs are quite incompatible with the progressive, and fully acceptable in the non-progressive, regardless of tense. How is this to be explained?

We have not developed enough of the formal structure of the semantic component to be able to give a formally rigorous account of the above facts. Nevertheless, the solution is clear in its outlines. Recall the Semantic Closure Principle of Chapter 1, which said in essence that every sentence is assigned to a particular category in a
semantic domain D, provided there is a semantic rule that maps into D. This is a well-formedness condition on semantic representations. The only further principle we need to make explicit in order to present our solution is (117):


Well-formedness conditions are met minimally.

By our assumptions about the predicates be ahead and recognize, they are lexically marked 'phenomenal'. This lexical marking of the predicate determines a semantic property of the entire proposition, namely that in terms of Epistemic Status it is a phenomenal description. (This is the step where formal rigor is most clearly lacking, although it is an intuitively entirely plausible step). As a consequence, a sentence in which either of the above predicates occurs is assigned to the category 'phenomenal' in the domain of Epistemic Status. If the predicates occur with be + ing, the rule assigning phenomenal status to the proposition will apply. But this last step is a violation of (117). The proposition is already marked for the category 'phenomenal'. Re-marking it for the same category is not consistent with satisfying the Semantic Closure Principle minimally. Thus sentences containing phenomenal predicates plus the auxiliary be + ing are ill-formed.

2.4.3.5.2. We have now dealt with predicates that are lexically marked as 'phenomenal'. The class of predicates left to discuss are those that are lexically marked 'structural'. An example of this type,
which we have previously discussed in connection with the notion of temporariness, is the verb own. Relationships of ownership are structural by their very nature. They depend entirely on legal conventions, and they are not characterized by any associated phenomenal perceptibles.

The simplest answer to the question as to why a sentence like (118)

(118)*John is owning a brand-new Mercedes.

is ill-formed is that the TMA component assigns a meaning which is inconsistent with the lexical meaning of the predicate. Not only is this the simplest answer; it is also the correct one. In particular, it would be wrong to conclude that the ill-formedness of (118) is due to principle (117) in conjunction with the Semantic Closure Principle, analogously to the case of (114) and (116). It might be thought that there is an entirely parallel argument here: the lexical marking is transferred to the sentence, and thus the sentence is specified for structural Epistemic Status. At a later stage, there is again a specification for Epistemic Status, this time for phenomenal status. But, it might be said, the sentence had already met the well-formedness condition at the earlier stage, thus letting the rule apply would violate the Minimal Satisfaction Principle (henceforth MSP).

A simple consideration shows that this cannot be the correct interpretation of the MSP. If it were, no semantic rule could assign a more specific meaning in a domain for which a proposition is antecedently marked in a more general way. The way to interpret the situation correctly is as follows: the point at which the Semantic Closure
Principle goes into effect is the point where the semantic rule referred to in the formulation of the principle has just had a chance to apply. At that point, the sentence either satisfies the Principle, or it does not. If it doesn't, the minimal steps that ensure satisfaction are taken, and the derivation goes forward. If it does, then it does so either minimally, or not minimally, with respect to the semantic category the sentence is specified for at that point. If satisfaction is minimal (which means that there is no earlier stage in the derivation at which the well-formedness condition is satisfied with respect to the same semantic category as the one assigned at this later stage), all is well, and the derivation proceeds. Otherwise, the sentence is ill-formed.

An important example that supports this view of how the Semantic Closure Principle and the Minimal Satisfaction Principle interact has already been discussed in section 2.4.3.4. Achievement verbs are lexically specified for both epistemic categories. Yet they do occur with be + ing, because the semantic rule triggered by be + ing picks just one of the two categories, cancelling the other, and thus has genuine semantic consequences.

2.4.3.6. In the preceding section we have discussed two classes of predicates with the surface property of being resistant to the progressive. While this surface property has been known to grammarians for a long time, it has never been successfully analyzed. The conventional wisdom (cf. Lakoff (1966)) has been that there is a syntactico-semantic feature [stative] that distinguishes a class of verbs. This property of
stativity has been an extraordinarily elusive one to characterize. Lakoff's (1966) set of diagnostic environments basically revolves around the notion of agentivity, and there is thus some suggestion that stativity and agentivity are incompatible. A rather different generalization with regard to stative predicates can be extracted from Fiengo (1974) who states that "a predicate allows Progressive Aspect if and only if it allows locatives". From this it follows that stativity and the specification of a location are incompatible with each other. And there are numerous other attempts in the literature at giving criteria for membership in the class of stative verbs.

Our analysis shows immediately why there are formidable difficulties inherent in characterizing stative verbs. They can hardly be said to form an intuitively natural class. The property they have in common --- being intrinsically specified for Epistemic Status either one way or the other --- is highly abstract, and completely inaccessible to any attempt at extracting shared substantive semantic features. The shared property is a purely formal one.

Nevertheless, observations such as those made by Lakoff, for example, may be quite accurate. But rather than constituting an even partial analysis of the data, they can be argued to follow from the proper analysis. Lakoff's focus on agentivity as an explanation is misguided, but his intuition about the data is sound. Agentivity is equivalent to the capacity to use one's free will. It is inherent in the notion of free will that this free will can be exercised capriciously, or in accordance with some guiding principles. But free will, by
definition, neither entails complete randomness of behavior, nor does it entail blind adherence to some or other principle. Thus it can never have the effect of making its exercise in principle describable only in phenomenal, or only in structural terms. Both options must be available, or we wouldn't conceive of the phenomenon in terms of being the product of a free agent. Thus the fact that agentive predicates are never restricted to being used in the non-progressive follows from our analysis of Epistemic Status (quite apart from the aspectual progressive/non-progressive contrast).

2.4.3.7. The last item of business in this chapter will be an examination of the relation between the aspectual be + ing and the epistemic be + ing.

There is an important syntactic difference between them in point of 'degree of grammaticalization': the aspectual progressive is clearly less grammaticalized than the epistemic progressive. This is correlated with a difference in generality of applicability as follows: the aspectual progressive is limited to being applied to propositional contents that represent events, while no restrictions as to type of propositional content are imposed on the epistemic progressive.

The syntactic difference between the two forms has another clear consequence in our framework. By the TMA Ordering Principle, the rule interpreting the epistemic progressive takes scope over the rule interpreting the aspectual progressive, since the aspectual progressive is closer to the verb stem. This leads quite naturally to the question: What happens when the epistemic progressive and the aspectual progressive
co-occur? An example of this type would look like (119):

(119)*Walter is being filing the day's mail.

This is an ill-formed sentence, and in a framework that explicitly recognizes two progressives, this ill-formedness demands an explanation.

There is an almost trivial answer to this problem: an event is an object of observation, and not of knowledge. This is putting matters rather abstractly, but the point is hardly controversial. An event viewed in and of itself is something that takes place at the phenomenal level; hence, the description of an event as such is a description of a phenomenon, not of any structure. We propose that this truism should be reflected in the semantic representation of event-descriptions by a redundancy mechanism: Every sentence that receives an analysis in terms of subevents, i.e. an aspectual analysis, will be taken to concern itself just with the event at hand, and will consequently be marked as phenomenal (by exactly which formal operations is a rather premature question, given the state of the art). Granted this, the account of the ill-formedness of (119) is precisely parallel to the explanation we gave for the ill-formedness of "Lucky Lady is being ahead" once we had established that the predicate 'be ahead' was marked as phenomenal.

This account predicts that the epistemic progressive will not only be incompatible with the lexical item representing imperfective aspect, but with any lexical item of Aspect. This prediction is borne out.

(120) a.*Walter is stopping filing the day's mail.
b. *Walter is keeping filing the day's mail.
c. *Walter is beginning filing the day's mail.

This result is rather striking: it complements rather nicely some observations made by Milsark (1972) in regard to the Doubling Constraint proposed by Ross (1972). Milsark points out that the Doubling Constraint cannot be a prohibition against a certain inadmissible sequence of morphemes, since there is a systematic difference between participial complements and gerunds in that the former are subject to the constraint, and the latter are not. We have provided a well-motivated account for the central cases -- participial complements -- that the constraint does apply to; the surprising thing is that the explanation for a fact that seems so transparently phonetically inspired could derive from rather abstract semantic considerations.
3.1. The syntactic analysis of the Perfect.

Syntactically, we will treat the English perfect as an auxiliary of the same type as the epistemic progressive. Thus, the structure we will assume is as in (1):

(1)

The perfect and the epistemic progressive can co-occur as sister constituents. Their order relative to each other is presumed fixed by the Verbal Hierarchy Hypothesis; in other words, it is not a syntactic problem. The Verbal Hierarchy Hypothesis has some intuitive support here.
As we go away from the verb stem, the verbal categories acquire ever widening scope: Aspect is concerned with the internal constitution of an event, Epistemic Status is concerned with the relation of a propositional content to knowledge and observation, while the perfect, as we shall see, introduces a second proposition into the analysis. The horizon broadens from step to step.

The perfect shares a property with the progressive: *have* + *en* forms a substitution set of one. We have already commented on this kind of situation in connection with the epistemic progressive in 2.4.1, and those comments apply here as well.

3.2. The notion of a tactic relation between propositions.

Consider sentences (1) and (2).

(1) Bruno has dined at Mama Anna's many times.

(2) Bruno dined at Mama Anna's many times.

There is one respect in which (1) and (2) are the same: the many occasions of Bruno dining at Mama Anna's that must be established facts for either (1) or (2) to be true all lie in the past from the vantage point of the speech event. Moreover, they could be the very same occasions in both cases. This is a rather curious state of affairs: we have a pair of sentences that differ from each other with respect to two grammaticalized morphemes -- tense and the perfect -- and yet, they appear to have identical truth conditions.

This focus on what occasions of dining a speaker may have in
mind when uttering (1) or (2) is somewhat misplaced, however, if we want to discover in which respects (1) and (2) are different from each other, as they surely are, and in particular what role the perfect plays in this. But if we shift our focus and ask not about truth conditions as much as about appropriateness conditions for the utterance of (1) and (2), respectively, the differences become readily apparent. We notice that (1) could not be used if Bruno was dead; (2) could be used in such a case. Similarly, (1) could not be used if Mama Anna's restaurant had burned down or had been closed down, or had simply changed hands and was now called 'Chez Antoine'; (2) would be felicitous in all these circumstances. There is a suggestion about (1) to the effect that Bruno most likely will dine at Mama Anna's again some time, which is absent from (2). There is a difference in regard to felicitous continuations of a discourse which begins with either (1) or (2): to add to (1) "...so he can depend on getting a good table in the back of the room when he goes there" would be entirely natural, but it would be quite incongruous with (2).

These differences are sometimes all subsumed under a general statement saying that the perfect indicates "the continuing relevance of a previous situation" (Comrie 1976, p. 56). This statement is not quite complete: relevance is a dyadic relation. We ask: 'relevance to what? Well, to the matter at hand, to some current situation. How much of this does the formula in (3) express (taking 'p' here as a short-hand notation for 'the fact indicated by the propositional content of the sentence', 'T(x)' for 'x is a time interval', and 'DUR(x,y)'
'x<y' as before?

\[(3)\ (\exists x) (\exists y) (T(x) \& T(y) \& DUR(p,x) \& DUR(z,y) \& x<y)\]

(3) expresses the temporal relationship between the two situations; It expresses the fact that the nature of the 'matter at hand' is left open by having the variable z in the argument place for the second situation (which we will call the 'reference' situation, borrowing Reichenbach's (1947) term). No explicit mention is made, however, of any other relation holding between p and the unspecified reference situation. But this need not be a deficiency of our representation: the grammatical fact may be that just as the nature of the reference situation is unspecified, so is whatever kind of non-temporal relation holds between p and the reference situation unspecified.

This is not merely a logical possibility; rather, it is an extremely plausible hypothesis. The overlay of non-temporal relations over the grammatically specified temporal one is the rule rather than the exception. This can be illustrated with examples of the type A when B, of which we have given an analysis in temporal terms in 2.3.3.2.

Consider sentences (4) and (5):

(4) When John got that letter, he flew into a rage.

(5) Just before John flew into a rage, he got that letter.

(4) and (5) place the events of John getting a letter and of John getting mad into the same temporal sequence. But (4) accepts quite naturally the overlay of a causal relation, which is impossible for (5). There are
pairs that go the other way: the sentence with before can suggest a non-temporal relation, where the one with when cannot (at least not without some modifications):

(6) Just before Max left for a long vacation, he increased the insurance on his house by 100%.

(7) When John increased the insurance on his house by 100% he left for a long vacation.

Furthermore, the non-temporal relations are of various kinds. While (4) suggests that the earlier event provided the cause of the later one, we see in (8) an example where the earlier state of affairs provided the opportunity for the later one, and in (9) an example where the earlier clause provided a clue for inferences to the later one:

(8) When we were in Paris, we visited the Louvre.

(9) When he is quiet like that, he is liable to explode any minute.

It is obviously missing the point to construe all these possibilities as illustrations of subtle sense differentiations within the meaning of the connective when. The virtue of a form not being specific with respect to the semantic domain of non-temporal relations may, from the speaker's point of view, be a concomitant lack of commitment on his part to any particular position on those relations. He can arrange matters to suggest certain connections, but he is not committed to
having asserted them.

This methodological digression served to support our suggestion that (3) is an adequate expression of the semantic contribution made by the perfect to the interpretation of a sentence as a matter of grammatical fact. The perfect, on this analysis, fixes the temporal relation between two situations, one indicated by the propositional content of the sentence, and the other left unspecified except, as we shall see, for its temporal relation to the time of the speech event. The fact that the reference event is unspecified, however, has consequences of its own. It decreases vastly the specificity of the non-temporal relations that a speaker can use the perfect to suggest. The when-sentences in (4), (8), and (9) can be used to suggest with considerable precision which non-temporal relation the speaker has in mind, mainly because both situations involved are explicitly described, and the hearer can draw on his knowledge of the world in narrowing the number of ways they could be connected.

There is a further, Gricean, point to be made here. If a speaker is presumed to say no less than he knows, what can be concluded about the conversational potential of highly unspecific forms? On the basis of just this one principle one would conclude that the unspecific form is chosen because the speaker's knowledge of the facts is too limited to say anything more specific. But there is a second broad range of situations where unspecific forms will be favored: situations where the specifics are obvious. This is a consequence of a crosscutting principle to the one mentioned above, to the effect that the cooperative
speaker is not overly redundant.

These general considerations concerning lack of specificity in the grammar, and the status of such a lack of specificity in light of rules for good conversational behavior, have direct bearing on broad classes of situations in which the perfect is appropriate. We will refer back to this general discussion in later sections as the need arises.

As far as the place of the perfect in a larger system of verbal categories is concerned, we accept Jakobson's suggestion that there is a broad semantic domain of Taxis, to which the English perfect belongs, but which excludes tense. We give our own informal characterization of the difference (not exactly Jakobson's) as follows: Taxis relates two situations, Tense fixes the time of a situation in relation to the time of the speech event.

3.3 The perfect as a category of Anterior Taxis.

In this section we set out in some detail the consequences of adopting (3) as an adequate analysis of the English perfect. The discussion is generally speaking in a defensive mold: the normal case will be that perfect sentences will appear to say more than our analysis says they do.

3.3.1. The first use of the perfect to be discussed has been described as the 'perfect of persistent situation' (Comrie, 1976), or the 'state-up-to-the-present' sense of the perfect (Leech, 1971). McCawley (1971) says that this use of the perfect indicates "that a state of affairs
prevailed throughout some interval stretching from the past into the present." Examples of this use are (10) – (12).

(10) The Smiths have lived in London since last September.

(11) I've known Max since 1960.

(12) They have been waiting for ages.

Our analysis treats (10), e.g., as follows: Given that the perfect is ordered after the epistemic progressive, the semantic rules involving Epistemic Status have applied at the stage the perfect is interpreted. For the non-progressive form in (10) we get a representation like 
\[ [s(p^*,z)] \], where \( p^* \) incorporates the propositional content of "The Smiths live in London." The temporal adverbial since last September is treated as a restriction on the time variable; if '1' represents 'last September', the restriction is \( \sim(x<1) \), where \( x \) is the time variable. Finally, tense is interpreted on the reference situation; \( t_0 \) stands for the time of the speech event. In sum, (10) is represented by (13):

(13) \( (\exists x)(\exists y) (T(x) \& T(y) \& \text{DUR}([s(p^*,z)],x) \& \sim(x<1) \& \text{DUR}(w,y) \& y=t_0 \& x\leq y) \)

(13) says considerably less than (10) appears to express. It asserts that there is a time interval during which the Smith's living in London is a structural fact, and, further, that this time interval is not earlier than last September, but before the time of the speech act. (10) appears to say, in addition, that the Smiths have lived in London
continuously since last September, and, moreover, still live there.

But, we argue, none of what (10) appears to say in addition to (13) is a question of sentence grammar. Consider what contexts (10) is likely to be uttered in. If it is in response to the question "How long have the Smiths lived in London," the questioner is presumably aware that the London residence is the current one, and has been for some time, continuously. He just wants to know exactly how long this state of affairs has prevailed. Our analysis is entirely compatible with this kind of situation. Nothing in (13) says that the Smiths have not lived in London continuously, nor need (13) be taken to imply that they no longer live in London. One might even propose that in such cases \([s(p^*,z)]\) suggests itself as a replacement for the variable \(w\), which is to say that the fact of the Smiths living in London before the present is related merely to the fact of them living there now.

Similarly, in the case of (11), the information about 'persistence into the present' is readily supplied by normal assumptions about acquaintances. No special reading of the perfect needs to be appealed to.

It is sometimes suggested (e.g. in Comrie (1976), McCawley (1971)) that the status of the 'perfect of persistent situation' as a distinct category becomes especially plausible in view of the fact that languages other than English (e.g. German, French, Russian) use the present tense here, although they have a perfect which, in other respects, behaves rather like the English perfect. The facts of German, French, and Russian are indisputable, but they can hardly be presumed to be directly relevant. Even a cursory examination of English present tense and
German present tense shows that they do not have the same meaning. While the meaning of the former is 'simultaneous with the time of the speech event', the meaning of the latter is something more like 'including the time of the speech event, or later than it'. If situations like those in (10) - (12) are to be described, the English present could only describe what the current facts are. Corresponding to (10) we would only have "The Smiths live in London." This may suggest that the Smiths have been in London for a while, but no assertion about any past time is actually made. If it is important that past time be included in the assertion, perhaps even highlighted, the English present is simply not suited to the purpose. The German present, on the other hand, can extend freely into the past, as long as the time interval in question includes $t_0$.

But neither the German present nor the English present perfect have any distinguishable sense that picks out just the situation diagrammed in (14):

\begin{equation}
I: \text{interval consisting of } n \text{ units of time preceding } t_0 \text{ plus } t_0 \text{ itself}
\end{equation}

\begin{tikzpicture}
\draw[->] (0,0) -- (0,1) node[above] {$\text{t}_0$};
\draw[->] (2,0) -- (2,1) node[above] {$\text{t}_0$};
\draw (0,0) -- (2,0);
\end{tikzpicture}

The situation can be described precisely. But there is no reason to expect that some or other grammaticalized form of a given language must have a meaning that matches just this situation. The present tense in German is the least marked grammaticalized form in the language that is consistent with being used to describe (14). In English, no form is
similarly unspecific. The closest English can come to covering situation (14) is with the present perfect, whose unspecified reference situation provides the necessary degree of interpretive freedom to render (14) as follows (we will call the fully specified situation the 'narrated situation'):

(15)

The entire stretch of n units of time immediately preceding \( t_0 \) plus \( t_0 \) itself, can be taken up by one persistent situation if the reference situation is identical in propositional content to the narrated situation. This is a choice that is available to the speaker, and thus, the present perfect can be given an interpretation that is consistent with describing (14).

The picture we have developed here implies that if a speaker of English wants to make explicit that he intends a contingent statement to be true for all time, a sentence like (16) falls short of being fully explicit, while (17) is adequate in this respect:

(16) Certain approaches to politics have always been morally wrong, and they always will be.

(17) Certain approaches to politics have always been morally wrong, they are wrong now, and they always will be.
The judgment is subtle. But it does not seem that the middle conjunct of (17) is intolerably redundant; pedantic, yes, but that is what being fully explicit is. Contrast (16) and (17) with (18) and (19):

(18) Max and his wife Lisa have six children, and they don't want any more.

(19) Max and his wife Lisa have six children, Lisa has given birth to their sixth child, and they don't want any more children.

The middle conjunct of (19) is not just an example of pedantry, but of sheer redundancy. This difference turns on the difference between an implicature and an entailment: it is pedantic to make explicit the former, but redundant to do so with the latter.

3.3.2. The second use of the perfect that has been taken to reflect a distinct sense of the form, is called the 'indefinite past' by Leech (1971), the 'experiential perfect' by Comrie (1976), and is characterized by McCawley as indicating "the existence of past events." Examples of this use are (20) and (21).

(20) Jim has read several books on semantics.

(21) Mack has won money at the dog track exactly twice in his entire life.

This is a rather puzzling category. The obvious question is how it differs from the simple past. Does the simple past not also indicate the
existence of past events? Surely, in both "I saw him drunk a grand total of three times" and "I have seen him drunk a grand total of three times" the existence of exactly three past situations is 'indicated'. This is one of the points we made in section 3.2. Leech proposes to resolve this puzzle by insisting that the exact formulation of this meaning should be 'at-least-once-in-a-period-leading-up-to-the-present'. But this is of no help. For neither "I saw him drunk" nor "I have seen him drunk" can be asserted truly unless there is at least one occasion of me seeing him drunk in my past experience. The point about inclusion of the present moment that we considered in the preceding section hardly matters here: nobody has ever suggested that the past event whose existence is indicated must continue to the present moment.

Comrie gives what appears to be the correct answer (Comrie, 1976, p.59, fn. 4) in a footnote: "...the requirement of present relevance is important". Our own informal discussion of how (1) and (2) differ clearly points in this direction. We enumerated some circumstances which would tend to make the fact of Bruno having dined at Mama Anna's in the past less relevant to the present. They are all circumstances that discourage use of the present perfect: Bruno's death, the restaurant closing down, Bruno moving to another city.

Does our analysis of the perfect speak to this point? Not explicitly; but it comes close. The unspecified reference situation in our formula (3) has logical consequences, as follows: The variable indicating the position of the reference situation in the logical structure of a sentence must be replaced with some constant (we leave
quantification out of consideration for ease of exposition) when the sentence is to be asserted. Otherwise, the sentence is an open sentence, has no truth-value, and can thus not be asserted. The second point to note is our earlier point about non-temporal relations being regularly overlaid over grammatically specified temporal ones. When we discussed this matter in 3.2 we hypothesized that lack of specificity with regard to the reference situation was related to the degree of precision with which a particular non-temporal overlay relation could be suggested. The less specific the reference situation is, the less precise any suggestions of overlay relations will be. This is where 'relevance' comes in: it is simply the least common denominator for 'non-temporal overlay relations'.

Let us illustrate with a simple case. We will let our example sentence be "Jim has read the Pentagon Papers." Drawing on earlier material, we introduce the predicate \( S_R(x) \) for 'x is a subevent of an event of Jim reading the Pentagon Papers'. The rest of the non-logical symbols in (22) below are used as before.

\[
(22) \quad (\exists x)(\exists y)(T(x) \land T(y) \land (z)(S_R(z) \supset DUR(z,x)) \land DUR(w,y) \land y=t_0 \land x<y)
\]

Notice that this sentence cannot have a 'persistent situation' reading. The way we have proposed to analyse such readings is as follows: replace the variable \( w \) with a copy of the narrated event. In this case, we would substitute '(z)(S_R(z) \supset DUR(z,y))' for 'DUR(w,y)'. Aside from the accidental fact that this representation of the reference event is unsatisfiable (since the reference event is present; see section 2.3.3.3),
the narrated event and the reference event must also be **distinct**, since their associated time intervals are non-overlapping, and all subevents of the narrated event take place during one interval, and all subevents of the reference event during the other. Hence, (22) cannot be taken as describing just **one** continuous situation.

Now, whatever is substituted for w must be distinct from the narrated event. The trivial relation of 'continuation of the same situation' is no longer available for the non-temporal overlay relation. A less trivial but very unspecific relation of relevance is, therefore, suggested, as elaborated above.

The thrust of this example is to show how, without any proliferation of senses, the analysis of the perfect advanced here can account naturally for the different uses of the form.

3.3.3. A third use of the perfect that is generally regarded as a distinct sense is the 'perfect of result', which is said to indicate that the direct effect of a past event still continues.

Examples of this use are (23) - (25); the kinds of results that are typically taken to be indicated by the perfect are supplied in parentheses with each example.

(23) He has been given a camera. (**He now has the camera**).

(24) Max has been fired. (**He is currently out of work**).

(25) John has had a bath. (**He is clean now**).

This kind of example can be characterized quite precisely within
our framework. Recall our description, in 2.3.1, of events as consisting in a progression from an initial state to a terminal state. The common property of examples (23) - (25) is that a single event is taken to satisfy the portion of formula (3) that deals with the narrated situation. In addition, the terminal state of that event is taken to satisfy the portion of formula (3) that deals with the reference event.

Again, formula (3) is entirely consistent with the favored interpretation of the set of examples considered, without determining that interpretation. There is an interesting tendency to be observed, nonetheless: the tendency to regularize the choice of a reference situation in a number of rather well-defined cases. This tendency in turn appears to be informed by a principle, which we might call the principle of minimal conceptual distance. That is to say, the general trend seems to be to pick a reference situation that is conceptually as close to the narrated situation as possible. Identity is the extreme case of such a relation of closeness. The relation of an event to the terminal state of that event is very close as well: it is a relation of entailment. This observation will acquire some interest in section 3.4.

3.3.4. The fourth category, the 'perfect of recent past', is not particularly well-defined as a separate category. The property of recentness which is supposed to distinguish the narrated situation here is not excluded from any of the other three categories. That the present perfect is often used in situations where the narrated event is quite
recent, is to be expected. The chances of a narrated event being relevant to the present decrease the more remote in time the narrated event is.

Thus, this category, while generally recognized, appears to add nothing of interest to an analysis of the perfect. And yet, if one attends closely to the claim that the relevance of a narrated event is apt to decrease as time goes by, one can surely not be entirely happy. Take the example of a promise which we will imagine Uncle Bert to have given little Tim to the effect that the uncle would take his nephew to a baseball game. Now as time goes by, and Uncle Bert fails to keep his promise, the promise may lose relevance in terms of young Tim's expectations of seeing a baseball game with his uncle. On the other hand, the more time passes, the more relevant the unkept promise may become in terms of Tim's assessment of Uncle Bert's trustworthiness. Or consider the death of a political leader, which may become more relevant as time goes by, and as more and more of his appointees fail to get reappointed, resign, or die themselves; while for a time the ship of state may persist on the old course, there may be radical changes later. But the conditions directly after the leader's death are just those that his life and his policies are relevant to, and not his death, while the more remote, radically changed conditions are just those that his death is relevant to.

In other words, there is some real danger here that we are using a term, namely 'relevance', which is not merely vague or general, but practically vacuous (its long tradition notwithstanding). The difficulty can be highlighted with an earlier example. When we discussed the sentence Bruno has dined at Mama Anna's many times we observed that it would
be inappropriate to use that sentence if Bruno was dead. The account offered for this observation a few pages later was that Bruno's death makes questions about where he used to dine irrelevant. But we can easily make up scenarios in which those questions become highly relevant, and yet, the appropriateness judgments about the above sentence are exactly as before. Suppose, for example, that the plumbing at Mama Anna's had been defective, and toxic substances had gotten into the water used for cooking. A one-time customer would suffer no adverse consequences, but regular patrons like Bruno would be slowly poisoned. Now given this scenario, Bruno's dining at Mama Anna's many times is extremely relevant to his subsequent death. Nevertheless, a witness in the ensuing investigation could still not volunteer the above perfect sentence as part of his testimony.

This weakness of our analysis should not be downplayed. On the other hand, it does not seem to be a weakness that vitiates the general approach adopted here. Rather, the problem is quite narrowly circumscribed: it lies in the particular choice we have made of a concept to capture the nature of the relation between narrated situation and reference situation in perfect sentences.

The discussion of the last four sections is suggestive of at least one alternative that may, at a purely descriptive level, avoid some of the pitfalls of using the notion (or perhaps non-notion in this case) of relevance. What if the appropriate notion were, instead of relevance something more like maintenance of the status quo? The status quo involved here would be whatever state arises automatically from the narrated event or situation, where 'automatically' just means 'in the absence of any unusual, disturbing, or interfering circumstances'. Maintenance,
the second part of our descriptive phrase, is not to be endowed with any agentive overtones; it is to be understood in the sense of 'persistence through time'.

This idea of 'maintenance of the status quo' fits the case discussed in section 3.3.1 much more naturally than does 'relevance'. The persistence of a state is the paradigm case of maintenance of the status quo. For the case of single events, discussed in 3.3.3, the natural candidate for the 'state to be maintained' is the terminal state. Thus, for the two cases mentioned, a description in the terms suggested here has some immediate appeal.

In the case of the experiential perfect, which we discussed in 3.3.2, the immediate consequence of recasting the description is at least this: different questions must be asked. The question can no longer be "what is there in the current state of the world that some situation in the past is relevant to?"; rather, we must now ask "what is the state that arose automatically from some past situation and that prevails still?". This does not seem to be much of a gain at first. Yet some reflection shows it to be a step in the right direction. Recall our example sentence Jim has read the Pentagon Papers, and consider the following two current situations that the fact expressed by that sentence might be relevant to:

(i) Jim now knows the contents of the Pentagon Papers, and (ii) Jim is in trouble because he read those documents without proper authority, and by so doing committed a federal crime whose investigation has just been intensified. The difference between (i) and (ii) is that the former refers to a state that is a more or less automatic consequence of reading, while the latter refers to a situation that is linked to the reading by a spec-
ial chain of relevance.

Now consider current discourse contexts that focus on (i) and (ii), respectively. In the case involving (i), somebody might say "I don't understand your obsession with those Pentagon Papers you've heard about. But if you're so curious, talk to Jim. He's read them." The passage is intended to highlight the fact that Jim knows what is in those documents; substituting "He read them" for its present perfect counterpart should, and does, yield a less tightly-knit discourse. The reason is that "He read them" does not suggest, as does "He's read them", that Jim now knows the contents of the papers. But only if he did would talking to him about them be apt to satisfy somebody's curiosity in that regard.

In the case involving (ii), on the other hand, an ambitious FBI agent might say "We've got to check out Jim Trim. That guy is overdue. And he read those papers, if anybody did. I'm sure of it." This sinister passage is intended to depict the agent as being out to get Jim, and the mere fact of Jim's reading the papers is enough for his purposes. Whether or not Jim remembers anything about his illicit reading matter is of no consequence to the plot here. The sentence "He has read those papers, if anybody has", if substituted for its simple past counterpart, would accordingly be somewhat out of place. It suggests a preoccupation with what Jim now knows, while the context suggests that the preoccupation is really with what Jim did.

The shift from 'relevance' to 'maintenance of the status quo' thus appears to be a promising one. In the case of the experiential perfect, however, it still suffers from the defect of being too vague. What is the status quo involved in (21), which is repeated below for convenience?
(21) Mack has won money at the dog track exactly twice in his entire life.

Why does (21) suggest not only that Mack is still alive, and that the dog track still stands, but also that Mack is still betting on dogs? Could it be that the state that arises automatically from certain situations is to be characterized by something like 'potential of recurrence'? For certain cases a satisfactory analysis can indeed be obtained by taking the status quo (and, consequently, the reference situation) to be a modal statement of the form "there is an actual potential for p to occur", where p is, roughly, the propositional content of the narrated situation, and 'actual' potentials are those established on the basis of precedent, and not logic alone. A refinement suggested by the facts of (21) would be to substitute a restricted variable for the focus constituent in p, so that the status quo to be extracted from (21) would be "there is an actual potential for Mack to win money at the dog track x many times (in his entire life)".

This is deliberately very rough. But notice that such an analysis would have some desirable properties. It would explain the intuition that for (21) to be uttered felicitously, Mack must be alive and the dog track must be in existence still. The reason is that if either of these conditions were not met, the modal statement we arrived at could not be true. The further (and weaker) intuition that Mack is still betting on dogs depends on the fact that Mack can win only if he bets. There is no entailment involved here, however, because Mack is a potential winner as long as he is at least a potential bettor; he does not currently have to be an actual bettor. Hence the relative weakness of the intuition about his
current betting habits.

The decision to substitute a restricted variable for the focus constituent is supported by the observation that focal stress is known to change the 'presuppositions' of perfect sentences. If we say "The Lewisburg Lambs have won three national titles in succession" with normal intonation, we sanction an inference to the effect that the Lewisburg Lambs are still a team. If we put focal stress on the subject NP in the above sentence, we remain neutral as to the current existence of the team.

We will not pursue these explorations any further here, interesting though they may be. What they have shown us is that the notion of 'current relevance' that has played a role in so many descriptions of the English perfect is in dire need of explication, or, if some of our speculations are on the right track, even in need of replacement. The problem with the proposed substitute, 'maintenance of the status quo', is that it is itself rather versatile at this point. Better than 'relevance', evidently, but not good enough. The second problem, whose solution depends to some extent on a solution to this first one, is to explain how the semantic analysis embodied in (3) above, determines the choice of whatever concept is appropriate to describing the relation holding between narrated situation and reference situation. Neither of these two problems appears to be very close to a satisfactory solution, in spite of some of the attractive steps in the right direction we have been able to suggest. For this reason, we will do without any terminological revisions at this point, and in what follows we will continue to talk about a 'relevance relation'. This will do no harm, as long as we remain aware of the problematic nature of this relation.
3.3.5. There is an interesting fact about the English perfect that is conspicuously missing from our account. This is the fact that the perfect interacts with time adverbs in a rather curious way. If the perfect cooccurs with present tense there is a restriction to the effect that specific past time adverbials can generally not occur in the same sentence. An illustrative paradigm is (26) - (28):

(26) They had packed on Monday, so that they were all set to go on Wednesday morning.

(27) They will have packed on Monday, so that they will be all set to go on Wednesday morning.

(28)* They have packed on Monday, so that they are all set to go this lovely Wednesday morning.

We have a rather speculative solution to the problem posed by (26) - (28). But first, some comments may clarify the nature of the problem. The first observation is negative: there is no incompatibility between the perfect form (have + en) and specific time adverbials indicating the time of the narrated event. This is clearly shown by
(26) and (27).

The second observation is also negative: the problem with (28) is not that the present tense morpheme cannot cooccur with a past time adverbial in the same sentence. The latter is Comrie's position. "It is not specification of time as such that is excluded, since one can specify the time within which the past situation held, provided the time includes the present." There is a minor qualification that Comrie later adds to the above statement "...the general constraint against combining the Perfect with a specification of time does not hold when the time specification is the adverb recently or one of its close synonyms." (29) below, though well-formed, does not fall under either of these two generalizations; hence, they cannot be correct:

(29) We have never had to bother with this in the past.

The adverbial in the past does not include the present moment, nor is it a close synonym of recently.

This leads to a positive proposal: the proper restriction is that the time adverbial may not specify a past time interval whose right bound is such that there must exist another time interval between it and $t_0$. The diagram for the inadmissible configuration is (30).

(30) 

<table>
<thead>
<tr>
<th>time interval</th>
<th>intervening $t_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>specified by</td>
<td>interval</td>
</tr>
<tr>
<td>adverbial</td>
<td></td>
</tr>
</tbody>
</table>

Adverbials like last year, 5 days ago, a minute ago, yesterday, etc.
fit (30); adverbials like since Christmas, in the past, recently, today, before this, etc. do not.

From this positive proposal we now develop the sketch of a solution to the problem at hand. For the sake of concreteness, we limit our discussion to a case of the 'perfect of result': it has the merit that the reference situation is readily determinable. Our example is (31).

(31) Brenda has been elected to the board. (She is on the board now.)

This is schematically represented in (32), where \( x \) is a time interval during which the narrated event occurs, and \( y \) is a time interval during which the reference situation obtains.

(32) \[
\begin{align*}
&\begin{array}{c}
\text{Brenda is}\\ \text{elected}
\end{array} \\
\hline
&\begin{array}{c}
\text{Brenda is on the board}
\end{array}
\end{align*}
\]

The notion of relevance we are using to relate the narrated event and the current situation is quite weak, as we have suggested before. The examples of the 'perfect of persistent situation' and of the 'perfect of result' can be taken to suggest that the relevance relation requires temporal contiguity: the continuation of a situation is of course contiguous to what went before, and the terminal state of an event is contiguous to the event. Let us now assume that the relevance relation
does require temporal contiguity in order to be sustained (between the narrated situation and the reference situation, as well as between one instant at which the reference situation prevails and the next). The only way the speaker could guarantee that the reference situation is sustained throughout a period of time is by having access to, or control over, that entire period of time. If there are temporal gaps in the representation, the relevance relation no longer suffices to relate two situations to each other.

Now we can see why there is a difference between the present perfect, on the one hand, and the past perfect and the future perfect, on the other hand. The decisive difference between present tense and the other tenses in English is that the former requires the time of the reference event to be simultaneous with $t_0$. This fixes the time of the reference event absolutely, with respect to any utterance of a sentence. But past tense and future tense merely stipulate that the time of the reference event is either some time earlier, or some time later, than the speech event. The temporal representation of a sentence can always be adjusted so as to avoid temporal gaps between narrated event and reference event (but see example (36)).

For the sentence, "He has left on Monday" we get the schematic diagram:

(33) $\begin{array}{c}
\text{Monday} \\
\text{he leave} \\
\text{intervening interval}
\end{array}$

$\begin{array}{c}
t_0 \\
\end{array}$
For "He had left on Monday" we get either (34) or (35).

(34)

```
he leave
---------
interval of reference situation
```

(35)

```
Monday

he leave
---------
interval of reference situation
```

(35) is the interpretation of "He had left on Monday" which is brought out unambiguously by preposing the adverb as in "On Monday, he had left."

But we do not get a well-formed diagram for "On Wednesday, he had left on Monday" or "He had left on Monday on Wednesday."

(36)

```
Monday

[-------------]

Wednesday

[-------------]

intervening interval
```

This account of the ill-formedness of sentences like (28) rests squarely on the claim that the relevance relation has the peculiar weakness of not being able to bridge a temporal gap. Going back to Brenda's election to the board (example (31)) for a moment, what exactly
is the reason for this weakness? By asserting that an event in its entirety took place, a speaker commits himself to the terminal state of that event having been reached, but only, of course, for the moment of time at the very end of the event. The terminal state will be maintained, however, as long as nothing intervenes to change it. From Brenda's election to the board, her being on the board follows. But, does it still follow two weeks later? If the perfect is to suggest that it does, then the interim period of two weeks must be accounted for. The speaker has to assume responsibility for it. The temporal representation must be continuous from narrated event to reference situation. It is important to bear in mind that the reference situation is asserted only to the extent that it flows naturally from the narrated event, and naturally continues in the absence of interference. If the reference situation is independently asserted, then a perfect sentence that merely suggests it cannot follow such an independent assertion in a well-formed discourse. Thus (37) would not be a felicitous opening to a news report.

(37) #Mike Mortimer is dead. He has died at the age of 87...

This is to be expected. A well-formed discourse does not lose information from one discourse element to the next. It only becomes richer in information. Thus it is quite anomalous to introduce a piece of information (Mike's being dead) as an asserted fact, and then to go on and cautiously suggest that the same piece of information is a fact, all else being equal.

This is still roughhewn. But the solution we have intimated has
the following interesting properties:

1. It depends on there being a definite right bound to the time interval the narrated event is located in. This was just the property that distinguished the compatible time adverbials from the incompatible ones.

2. It depends on there being a definite left bound to the time interval the reference situation is located in. This is a property of $t_0$, and it is furthermore the difference between (34), (35) on the one hand, and (36) on the other hand. The joint satisfaction of the conditions 1. and 2. is a precondition for the existence of an overt temporal gap, which is criterial on the present account.

3. It distinguishes sharply between present tense and the other tenses. The former imposes a definite left bound on the time interval covering the reference situation, the latter do not. Therein lies the reason for the asymmetry exhibited in paradigm (26) – (28).

4. It depends on a construal of the notion of 'relevance' as being extremely weak. The account does, therefore, not carry over to cases that involve a stronger relation between narrated event and reference situation. As we will argue in Chapter 5, the paradigm case of such a stronger relation -- deducibility -- is introduced by the interpretation of the modals of English. Thus, our account distinguishes non-modal sentences from modal ones, and the well-formedness of "He may have left yesterday" is predicted.

Importantly, our solution leaves the semantic characterization we have given for the perfect intact. The phenomena we have discussed
throughout section 3.3 all arise from the fact that the meaning of the perfect leaves the reference situation unspecified. The situation is rather reminiscent of the one we found ourselves in when discussing the English progressive. A number of phenomena that appeared, on the face of it, to be semantic properties of the progressive itself were shown to arise from the interaction of the meaning of the progressive with other factors. In this connection, it should be emphasized that once again, the Univocality Assumption has been borne out as it is intended: as a hypothesis about how many different things a grammaticalized form can mean, and not as an hypothesis about the variety of uses a form with a given meaning can be put to.

3.4. The Complement Convention and the perfect.

The perfect gives rise to considerable semantic structure. It introduces a reference situation and relates it to a time interval by means of the predicate $\text{DUR}(x,y)$. In addition, it specifies the relation of that time interval to the time interval associated with the narrated event. We have taken it to be the distinctive characteristic of the domain of Taxis that it relates two situations; the perfect falls under this description since it relates two situations temporally. As we have pointed out before, the perfect is fully grammaticalized. Therefore, the general principles of Chapter 1 apply to it. The question to be addressed now is: What is the unmarked counterpart of the perfect, and what are its semantic effects?

First, what is the unmarked counterpart of the perfect?
Exactly as in the case of the progressive, it is a null form; that is, the unmarked counterpart of the perfect is absence of the overt form $\text{have + en}$. Next, what is its meaning? The Complement Convention says that the meaning of the unmarked form is the complement of the meaning of the marked form. We also know that the meanings of marked and unmarked form together must exhaust the domain. In the domain of Temporal Taxis it can be precisely specified what counts as exhausting the domain: all possible temporal relations between narrated event and reference situation must be covered. Thus the meaning of the complement of the perfect is uniquely determinable: where the perfect specifies a temporal relation $x\lesssim y$, the non-perfect specifies a temporal relation $\sim(x\lesssim y)$. Any temporal relation between $x$ and $y$ is either a special case of the one or the other, and the exhaustiveness criterion is thus met. There is some reason to think, however, that the complement is actually more highly structured. If, as in the case of when-clauses, temporal overlaps can be rendered by other grammatical devices, such as aspectual distinctions, exhausting the possible relations in Temporal Taxis may come down to the following three relations: $x\lesssim y$, $x\approx y$, and $y<x$. This is the position we will adopt here, chiefly because of a gain in clearness of exposition. Finally, there remains one more principle to take note of: the Semantic Closure Principle which ensures that non-perfect sentences will still be marked in some way for Temporal Taxis.

This appears to be a very undesirable consequence of our theory. The above observations all put together imply the following analysis of a simple sentence like "John left" (we will let '1' stand for 'the event of John's leaving'): 
But recall our earlier discussion of the 'perfect of persistent situation' and of the 'perfect of result'. We noted then that the way perfect sentences are actually interpreted is, other things being equal, by making the reference situation as close conceptually to the narrated event as is logically possible. In the case at hand, this strong tendency determines matters uniquely. Identity is the relation of least conceptual distance. Given that for the temporal relation (38) allows \( x = y \), this determines \( l_j \) as the 'conceptually closest' situation to replace the variable \( z \). This entire operation is logically vacuous; (38), on this interpretation, is equivalent to (39) which leaves out all the structure introduced by Taxis:

\[
(39) \ (\exists x)(\exists y)(T(x) \ & T(y) \ & DUR(l_j,x) \ & DUR(z,y) \ & y < t_0 \ & (x=y \lor y < x))
\]

So one might conclude that (39) is preferable to (38) on grounds of simplicity. Moreover, it is absolutely certain that no native speaker's intuitions would favor (38) over (39).

But notice that the trivial interpretation of the tactic elements of (38) is only available in the 'unmarked' condition in which no relation stronger than relevance is assumed to hold between narrated event and reference situation. What if by some grammatical means a stronger relation than relevance were imposed? What if the reference situation received an independent characterization such that identity with the narrated event is precluded? Then we would of course expect
the tactic structure present in (38) (and not in (39)) to have exactly
the semantic consequences inherent in it: the reference situation would
have to be simultaneous with, or later than, the narrated event. But
the abovementioned possibility is exactly what our analysis of the
English modals will posit: an independent - albeit indirect - charac-
terization of the reference situation which makes it distinct from the
narrated event. And the consequences predicted by the present analysis
of Taxis are just what we find there.

Considerations of simplicity thus argue against our analysis
of Taxis, while considerations of generality favor it (as Chapter 5
will show). Inasmuch as our analysis is an attempt to discover the
native speaker's intrinsic competence, the simplicity argument is much
less compelling than the generality argument. The latter draws its
strength from the fact that the present analysis is both simple in its
structure (not necessarily in its outputs, as we have seen), and com-
prehensive. To satisfy simplicity criteria such as might be applied to
the choice between (38) and (39), the present analysis need only allow
that the native speaker, once having acquired a linguistic system, is
still able to trim excess fat from the basic analysis so as to arrive
at a more manageable system for purposes of logical operations, etc.
What systems are predetermined by universal principles, and how these
systems are represented for performance purposes are quite distinct
concerns; the intuitions concerning (38) vs. (39) have no direct bear-
ing on the first of these two concerns.

This concludes our analysis of the perfect/non-perfect alternation.
A number of problems still remain; the most serious of these is the lack of any firm structure in the domain of discourse strategies or discourse principles. We have simply made some assumptions that appeared to be reasonable, clustering around the notion of a non-temporal 'overlay' over semantically determined temporal relations, and around the question of what the discourse consequences are of introducing, but not specifying, an argument in such a temporal relation. This entire area surely requires more serious attention. As far as our main objective here is concerned, however, - exploring the structure of the TMA component - the facts of the English perfect have been argued to be consistent with the principles proposed in Chapter 1.
4.1. **The syntactic analysis of Tense in English.**

We propose that English tense is syntactically a verbal specifier. This fixes, in the framework assumed here, the triple-bar level as the level of embedding. Also fixed is the fact that there is only one such syntactic position in every sentence. We assume two overt elements of tense in English, one an affix (past tense), and one a stem (future tense). These overt forms contrast with an unmarked form (present tense). These syntactic assumptions are summarized in (1):
The syntax of the English modals is most naturally dealt with in conjunction with that of tense. The reason for this is that we analyze the English modals exactly the same way we do future will. We will concentrate on a set of eight modals, and their entries will be analogous to that for will, as follows:

<table>
<thead>
<tr>
<th>Stem</th>
<th>Affix</th>
<th>Stem</th>
<th>Affix</th>
</tr>
</thead>
<tbody>
<tr>
<td>must</td>
<td>¤</td>
<td>can</td>
<td>¤</td>
</tr>
<tr>
<td>should</td>
<td>¤</td>
<td>could</td>
<td>¤</td>
</tr>
<tr>
<td>will</td>
<td>¤</td>
<td>may</td>
<td>¤</td>
</tr>
<tr>
<td>would</td>
<td>¤</td>
<td>might</td>
<td>¤</td>
</tr>
</tbody>
</table>

This is an extremely weak array of forms from the point of view of applying the TMA Ordering Principle to it. That principle says that those syntactic formatives representing a verbal category that are closer to the verb stem get interpreted first. The structural criterion for closeness in terms of number of intervening syntactic nodes fails to decide. The supplementary criterion for closeness in terms of linear order fails to yield a 'clean' decision: it says to interpret past tense before anything else, and it does not decide about future tense. Nevertheless, we will take our cue from the direction in which the 'unclean' decision points, and interpret tense before the modals. There is one point in the analysis where this decision becomes interesting, and we will further comment when we address the question of whether any tense is assigned to modal sentences.

4.2. The Semantics of English tense in detail.
We have already anticipated our semantic analysis of English tense in previous chapters. There is a clear exposition of the analysis in section 2.3.3.3, and there is little to add to that exposition. The one point that is worth making here is this: the time interval which tense sets into a relation to \( t_0 \), is consistently the time interval associated with the reference situation. In other words, a general characterization of tense in English is as follows: tense specifies which temporal relation obtains between \( t_0 \) and the time interval during which the reference situation prevails. For non-modal perfect sentences this is clearly the only sensible thing to do; in the case of non-modal non-perfect sentences it appears to be counterintuitive, but is arguably without damaging consequences; and we will argue in Chapter 5 that it is just the right thing to do in the case of all modal sentences.

Aside from this qualification, which makes the position outlined earlier more precise, we subscribe to the tense system as presented before. The make-up of that system leads to a crucial question rather directly. Given that future tense and past tense are the marked forms, and present tense is the unmarked form, how does the Complement Convention decide that the correct meaning to assign to the unmarked form is 'simultaneity with \( t_0 \)'? Why not 'inclusion of \( t_0 \)'? Is this decision based entirely on facts internal to the tense system? Or do facts about which other verbal categories are represented in the system, and in which form, have a bearing on this issue? If the latter, we might ask if such dependencies are structured in any way or not. It seems reasonable that only categories that are within the scope of a given
other category would be relevant to the internal constitution of that category. There is little doubt but that there are dependencies of the kind outlined here. The fact that English grammaticalizes Aspect, and German does not, may very well be directly relevant to the fact that English has a much more specific interpretation of present tense than does German.

We can only raise these questions here, not answer them. But the point needs to be made that the TMA component is almost certainly not just a collection of elements with certain external relations holding among them (say, the TMA Ordering Principle), but a tightly-knit system with rich internal connections as well. The conclusion to be drawn from this state of affairs is that the study of a single verbal category in isolation will, in general, not yield all the kinds of data necessary to understand that category. In the remainder of this section we will examine a number of cases where tense interacts with some other category in an interesting way.

4.2.1. Consider the following passages from Leech (1971):

(3) "The UNRESTRICTIVE use of the Simple Perfect is found with verbs expressing states. It is so called because it places no limitation on the extension of the state into past and future time."

and

"The Simple Present is notable for employment in the expression of 'eternal truths', and so is found in scientific, mathematical and other statements made 'for all time'."

Relevant examples are (4) (for unrestricted state) and (5)
(for eternal truth).


(5) Two and three makes five.

The question that arises is, in its most pointed form, as follows: Is the analysis of present tense offered here not hopelessly inadequate? Its claim is that the assertions of (4) and (5) are valid for a time simultaneous with $t_0$, i.e. for a time of minimal extension—in short, for a point in time. This is the polar opposite of the traditional claim, Leech's version of which we have cited above, which says that present tense assertions must be able to encompass 'all time'.

But let us consider carefully what the present analysis does say. For (4) it yields the expression in (6) (in this, as in all other examples, redundant taxis is omitted):

(6) $(\exists x)(T(x) \& DUR([s(p^*,y)],x) \& x= t_0)$

We let $[s(p^*,y)]$ stand for 'the fact that the fact of the Smiths living in London is structural with respect to them', with the subject of the sentence replacing the variable $y$. So (6) says that the fact that the Smiths live in London is a structural one, at present. What more is there to say about (4)? Any assumptions one might make about the permanence of the situation are bound up in real-world assessments of structures of different types. If some structures, such as that of mathematics, are taken to be 'eternal', so be it. It may well have occasionally been thought that the phrase 'true for all time' is rather
apt in capturing the ordinary conception of such structures; that is perhaps interesting psychologically or philosophically, but not linguistically. What a person assumes about the nature of mathematics need not be made explicit in the linguistic expressions he uses to convey mathematical truths.

It should be noted that we are not forgetting about a Gricean point here that we rely on elsewhere, namely that a speaker is to say no less than he knows. This principle makes sense only where there is a choice of saying more or less. If (6) is the most we can say (without pedantry) about any truth in English, then it is irrelevant that in the case of mathematics we really know more than a sentence like (5) says.

4.2.2. We now turn to a more interesting range of facts. In 2.4.3.3 we discussed the use of the non-progressive often called 'programmed future', and also the 'futurate' use of the progressive ("Bill leaves tomorrow at noon.", and "Bill is leaving tomorrow at noon"). Notice first that if the time adverbial tomorrow at noon and the present tense morpheme are both taken to provide information about the time of Bill's departure, then the present analysis is immediately and irremediably falsified. On our interpretation of present tense Bill's departure would be at \( t_0 \), while the time adverbial fixes the departure for at least 12 hours later than \( t_0 \). Plainly, we ought to have a contradiction, as in "Jim will go to school three years ago." But in fact, in the futurate uses mentioned above, no contradiction arises. This has sometimes been taken as evidence that the meaning of 'present tense' is not 'simultan-
eous with $t_0$', but rather 'simultaneous with, or later than, $t_0$'. This notion is reinforced by an analysis of the English tense system as consisting of just two affixed morphemes (past and present). Given that some version of the principle of 'exhausting the domain' in the sense we have discussed before, is generally recognized (perhaps only tacitly) and given the fact that past tense clearly doesn't cover future time, the above solution is inescapable, on the assumption of a two-tense system.

We will now show that the entire conception of tense and time adverbials (in non-perfect sentences) as invariably supplying information about the same time interval is simply wrong, and then go on to argue that a more adequate conception is implicit in the analysis of the epistemic progressive presented in 2.4.

The kind of sentence that shows conclusively that the solution to the descriptive problem raised by the futurate uses of present tense forms cannot be found by stretching the meaning of the present, is illustrated by (7).

(7) I was taking Mary out for dinner tomorrow night, but now she's come down with the flu.

Here the descriptive problem is insurmountable if it is conceived of in terms of having to construe the meaning of 'past tense' as somehow being compatible with the time specification of a future-denoting temporal adverb.

The analysis developed of the epistemic progressive/non-progressive
distinction yields a very natural account of sentences like (7).

We will represent the first conjunct of (7) in (8).

\[(8) \ (\exists x)(T(x) \ & \ DUR([n(p^*,y)],x) \ & \ x_{<t_0})\]

The key to the account is to be sought in the content of \(p^*\), or more particularly, the propositional content incorporated in it. We propose that this propositional content is 'I take Mary out for dinner tomorrow night'. This propositional content characterizes a phenomenal fact about the speaker of the sentence. We have argued that 'phenomenal facts' about people can be intentions. Thus, in ordinary prose 'taking Mary out for dinner tomorrow night' is described as the contents of an intention of the speaker. But it is 'having an intention' that is asserted by (7), as (8) clearly shows: the propositional content is deeply embedded; the predication of non-structuralness ('phenomenality') with respect to the propositional content is what enters in the form of a fact-nominal into the assertion predicate \(DUR(x,y)\). It is the intention, and not its content, that is given a temporal location as being in the past.

This is one of the most persuasive examples of the semantic domains we are dealing with taking scope over one another. In particular, Tense must take scope over Epistemic Status. Epistemic Status introduces an extra level of predication into the semantic structure, removing the predication constitutive of the propositional content from the direct scope of Tense.

Having analysed (7) successfully, there is no point in going
over futurate examples in the present tense; they are analysed in a parallel fashion. The entities whose properties such futurate sentences describe are, as we have said before, schedules and people with regard to their intentions. To leave no residue of open problems here, we should make quite clear a simple point about both structural and phenomenal descriptions: their validity is homogeneous over whatever period of time they are truly asserted of, so that they can be truly asserted of any portion of that period, down to atomic portions, of which $t_0$ is an example.

There is one interesting problem in this connection, but it is perhaps a problem of conceptualization more than one of linguistic analysis. This problem is the fact that our analysis of (7), which is in terms of a past intention, might be expected to carry over quite naturally to 'past schedules'. But a sentence like (9) is just not English, although it might be thought to have an analysis as describing a former schedule which called for the speaker to leave on the day after the sentence was uttered.

(9)#I left tomorrow, but the flight's been cancelled.

There are two broad possibilities here: either our account of the other futurate sentences is flawed, and (9) proves that it is; or our account is really right, and the problem with (9) is peculiar to (9) only (and any sentence of this general type, of course). There is at least some reason to believe that the latter alternative is the right one. The locus of the problem might be the status of the variable $y$ in fact-
nominals of the type \('[s(p^*,y)]\'. We have said previously that there are no restrictions as to which class of entities (animates, concrete objects, abstract concepts, etc.) can replace \(y\). We still stand by that. The question raised by (9), however, is a different one: Irrespective of questions of class membership, what are the criteria for something being a conceptually respectable entity? In particular, is an aborted schedule a respectable kind of entity? If the answer is yes, the problem with (9) persists. Otherwise, it disappears of course, since the aborted schedule could not then be the referent of an expression that closes the proposition \(s(p^*,y)\) at \(y\).

This resolution of (9) depends on there being a difference in conception between unfulfilled intentions and aborted schedules. It also depends on being able to claim that sentences like (9) are indeed always about aborted schedules. But the latter claim can be sustained quite easily. If (9) is not about an aborted schedule but about a schedule that still stands, then it violates conversational rules of the familiar type: the speaker, to be maximally informative, would have to report on that schedule in the present tense.

We conclude from the discussion of the 'futurate' sentences covered in this section that the preponderance of the evidence strongly supports our analysis, thus enabling us to maintain our semantic description of the present tense as being simultaneous with \(t_0\). It must be emphasized here that the possibilities inherent in the domain of Epistemic Status for mediating reference to future situations are quite limited: schedules, programs, itineraries, time-tables, etc., are a very limited class of objects that deal with fixing intentions; on the other hand, intentions are,
as we have observed before, limited to human beings. The major consequence of this state of affairs is that outside of this very special class of entities, any other reference to the future cannot be effected by means of a present tense form. We saw examples of such ill-formed sentences in Chapter 2, and we give another pair to make this point quite emphatically:

(10) *John is in bad shape tomorrow.

(11) *Ralph does better in next year's tournament.

This kind of example is, at the same time, a direct argument for our claim that will is the future tense morpheme. Sentences (12) and (13) are unexceptionable.

(12) John will be in bad shape tomorrow.

(13) Ralph will do better in next year's tournament.

4.2.3. In this section we return briefly to a class of sentences which was the main topic of 2.3.3.2: sentences with when-clauses. So far, we have only considered them from the point of view of studying the interaction of the temporal connective when with aspectual distinctions. But since aspectual distinctions are limited to event descriptions, there is an open question as to whether the analysis of when offered in 2.3.3.2 carries over to non-events in a satisfactory fashion.

Consider sentence (14).

(14) When Charles lived in Paris, he was ready for any kind of adventure.
Using \([l_{cp}]\) for 'the structural fact of Charles' living in Paris', and \([r_{ca}]\) for 'the phenomenal fact of Charles' being ready for adventure', we arrive at (15) as the semantic analysis of (14) (omitting tense).

\[
(15) \ (\exists x)(\exists y)(T(x) \& T(y) \& DUR([l_{cp}], x) \& JR([r_{ca}], y) \&
\]

(15) says that at least one time interval during which the fact of Charles' living in Paris obtained precedes at least one time interval during which the fact of Charles' readiness for adventure obtained.

How adequate is this as a representation of the temporal facts expressed by when-sentences? The only fact it captures directly is that the state described in the main clause cannot wholly precede the state described in the when-clause (the final stretch of the main clause and the initial state of the when-clause can also not exactly coincide, but we will leave that aside). And indeed, it is quite true that one situation (14) cannot be used to describe is one where Charles' readiness for adventure has vanished by the time he takes up residence in Paris. But otherwise, (14) appears to be rather indifferent as to how the two facts are temporally related. We diagram the abovementioned inadmissible situation in (16), some admissible ones in (17), and, finally, a problematic but interesting one in (18).

\[
(16) \ "\text{When } A, B \ \\
\hspace{1cm} A \ 'Charles in Paris' \ \\
\hspace{1.5cm} B \ 'Charles adventuresome' \ 
\]
(17) When A, B

a. \[ \text{A 'Charles in Paris'} \]
\[ \text{B 'Charles adventuresome'} \]
b. \[ \text{A 'Charles in Paris'} \]
\[ \text{B 'Charles adventuresome'} \]
c. \[ \text{A 'Charles in Paris'} \]
\[ \text{B 'Charles adventuresome'} \]

(18) ?When A, B

\[ \text{A 'Charles in Paris'} \]
\[ \text{B 'Charles adventuresome'} \]

(16) is properly excluded from being in the scope of (15), and all of (17) are properly included as being situations (14) could be used to describe. The case in (18) is not excluded by (15); but this seems to be quite wrong.

But the problematic case of (18) does not, in fact, undermine the analysis given in (15). The resolution of the problem lies in the fact that when is itself highly grammaticalized, which is to say, it is an element of a small closed set of temporal connectives. This leads us back to the Gricean point about the cooperative speaker not being unnecessarily uninformative. We gave an example earlier where an existentially quantified expression was inappropriate in a situation where the facts warranted use of a universally quantified expression. Just this situation is what we encounter here. The set of temporal connectives that when is a member of also includes after. For after we propose the following
semantic analysis:

(19) \[ A \text{ after } B =_{df.} (3x)(y) (T(x) \& \text{DUR}(A,x) \& ((T(y) \& \text{DUR}(B,y)))x < y) \]

This formula is minimally different from the one we gave for \( A \) \text{ when } \( B \) -- Chapter 2, #(36) -- in that (36) involved existential quantification over time intervals during which \( B \) is true, while (19) involves universal quantification over such time intervals.

By the Gricean criteria mentioned above, then, (19) is more informative than (36), which means that (36) cannot be used with any communicative success to convey the meaning of (19). Thus the reason (14) cannot be used to express the situation in (18) is not that (18) is inconsistent with the meaning of (14), but that -- with no extra effort -- (20) expresses far more specifically than does (14) the situation sketched in (18).

(20) After Charles lived in Paris, he was ready for any kind of adventure.

There is an interesting and important twist to this account. It will be recalled -- from 2.3.3.2 -- that the general configuration of (18) is expressible by \text{when}-sentences whose \text{component clauses are interpreted aspectually}. But there it is a difference in quantification introduced by the scopally inferior domain of Aspect that is sufficient to make the distinction. Both \text{when} and \text{after} are elements of what we might call Bound Taxis, a domain related to that of Free Taxis (which the perfect/non-perfect
alternation is mapped into). Taxis is scopally superior to Aspect. The semantically less marked *when* can thus be used in place of the semantically more highly marked *after*, since the difference between them has already been fixed by a mechanism of a scopally subordinate domain.

The preceding discussion has been of some intrinsic interest. But it is included in this chapter mainly for the bearing it has on the analysis of tense. The most vulnerable claim made in this work with respect to tense is that the present tense of English indicates simultaneity with \( t_0 \), an atomic time interval. This hypothesis about the English present has some interesting consequences. It leads to an explanation for the asymmetry in aspectually interpreted progressive/non-progressive pairs that are in the present tense, rather than in the past or future tense (cf. section 2.3.3.3). It also leads to an intriguing account of a lack of parallelism between present and non-present perfect sentences that include certain temporal adverbials (cf. section 3.3.5). As we have seen in the preceding section, it strongly supports our analysis of certain 'futurate' sentences that are not marked for future tense.

The point that has to be granted, however, in order for this analysis to succeed, is one we have made explicitly, but not very conspicuously, in the preceding section: All entities referred to by propositions that are not analyzed as events -- whether they be called states, conditions, or facts -- differ crucially from events, in that they do not have any internal temporal structure. They are homogeneous through time.

The meaning components of events need not all be satisfied simultaneously. The meaning components of contact verbs like *kick* may be motion
and impact (cf. Fiengo (1974)), but they are realized successively. To represent an event, the proper domain of quantification is not time intervals during which 'eventing' goes on, so that, when no more 'eventing' takes place, the event can be said to be a complete entity of its type; rather, the proper domain of quantification is subevents.

The meaning components of non-events, on the other hand, all have to be satisfied simultaneously. The state of bachelorhood combines the states of being male, unmarried, and adult, and anytime one of these is absent, the entire state no longer holds. While events as abstract entities cannot thus be characterized merely by reference to 'times of being realized', states as abstract entities can be. A state cannot be 'incompletely characterized', no matter how small (even to the point of being atomic) the time interval is during which it is asserted to obtain.

This is why the different status in point of semantic well-formedness that we encounter with aspectually interpreted progressive/non-progressive pairs in the present tense does not carry over to non-aspectually interpreted sentences. This case (discussed in 2.3.3.3) is the most obvious example where the difference between events and states is criterial. The discussion of when-clauses in this section, in conjunction with the material discussed in 2.3.3.2, provides another, more complex, illustration of the same point, but it makes the point without direct reference to tense. Thus we conclude that we have invoked the event/state distinction in the sense discussed not in an ad hoc manner to account for certain tense-related phenomena only. The distinction is independently motivated.
4.3. **Semantically exceptional occurrences of tense.**

There is an important respect in which all our discussion of Tense and its semantics has to be qualified: The analysis we have given here applies to free occurrences of the tense morphemes only. This qualification needs to be made clear: we will call an occurrence of a tense morpheme **bound** if its form is determined by a grammatical rule of agreement, or harmony. All other occurrences are **free** occurrences.

We will give an example of such an agreeing tense morpheme which is not drawn from the usual paradigms illustrating the Sequence of Tense phenomenon. The example is provided by the category of Bound Taxis, more traditionally called temporal adverbial clauses. The agreement facts are as follows: if the matrix is in the past tense, so it the adverbial clause. Elsewhere, the adverbial clause is in the present tense. Examples are (21)-(23).

(21) a. When Harry arrived at the airport, he was met by his fans.
    b. When Harry arrives at the airport, he is met by his fans.
    c. When Harry arrives at the airport, he will be met by his fans.
    d. #When Harry will arrive at the airport, he will be met by his fans.

(22) a. Bill hit it until it shattered.
    b. Bill hits it until it shatters.
    c. Bill will hit it until it shatters.
    d. #Bill will hit it until it will shatter.

(23) a. Jack helped before he was asked to.
b. Jack helps before he is asked to.

c. Jack will help before he is asked to.

d. Jack will help before he will be asked to.

Especially telling are the examples in (22), since the semantics of the connective until place the state of affairs represented in the adverbial clause after the state of affairs represented in the main clause. Thus, if the main clause state is in the future, the until-clause state is too, a fortiori. If tense in the until-clause were free, (22d) would have to be the correct form. But (22d) is anomalous. The conclusion that follows is that tense in until-clauses is not free; in particular, the tense form in the until-clause of (22c) is not free, but bound. Therefore, the semantic rules for interpreting free tense forms do not apply. No contradiction arises, as there would if the present tense in (22c) were interpreted to indicate simultaneity with \( t_0 \), while the meaning of until in conjunction with the tense of the matrix would indicate some time later than \( t_0 \).

The fact that tense in adverbial clauses is bound is probably no accident. Tense and Taxis are very likely complementary categories. The former relates a proposition temporally to the time of the speech event directly, the latter only indirectly, by relating a proposition \( p \) temporally to another proposition \( p' \). If the tense status of \( p' \) is determined, then the temporal relation of \( p \) to the time of the speech event is determined automatically, since the often-noted restriction (cf. Heinämäki (1974)) that \( t_0 \) must not fall between the time of \( p \) and the time of \( p' \) is a general condition on the well-formedness of such temporal chains (i.e.,
sentences like "After Harry came to Rome, he will soon learn to drive like a maniac" are ruled ill-formed by an independent principle, deriving, as Heinämäki speculates, most probably from conversational principles of the Gricean kind.

Aside from bound tense, there are two further small classes of cases where the tense rules here developed do not apply. There is a small set of environments in which there is a past tense/present tense contrast possible that receives no temporal interpretation, and another set of environments in which past tense is the only form permitted, likewise without temporal interpretation. The first set of environments are conditional clauses, in which the present/past contrast can be used to indicate that the condition is rather hypothetical (past tense form), or not so hypothetical (present tense form). An example is given in (24).

(24) a. If he comes, we will put him up in the guest room.
    b. If he came, we would put him up in the guest room.
    c. If he will come, we will put him up in the guest room.

The second set of environments is a heterogeneous collection of matrix frames with subjunctive force (i.e., the state of affairs in the embedded clause is presented as desired). Examples of this type are given in (25).

(25) a. It's about time we got going again.
    *get going again.
    b. I'd rather you didn't do that.
    *don't do that.
c. I wish you were more careful.

*are more careful.

In this section we have presented a number of different cases of forms which fall within the scope of this chapter as far as their syntactic analysis is concerned. They are affixal verbal specifiers. Yet the thrust of the presentation in this section has not been to try and subsume these cases under the analysis previously developed. Rather we have suggested that they should be treated apart. The question to be answered in our framework is whether or not this is consistent with the Univocality Assumption.

The answer is straightforward. Splitting up the meaning of the tense forms of English into a variety of different meanings is inconsistent with the Univocality Assumption as it is presently formulated.

At an impasse of this kind, there are two courses of action open to us: we can try to reanalyze the data and thereby restore consistency with a general hypothesis under fire, or we can attempt to modify the theoretical hypothesis. The alternative of simply abandoning the theoretical hypothesis is quite unattractive in this case; not only does it have a great deal of initial plausibility, but it has also received rather strong support in our analyses of Aspect, Epistemic Status, and Taxis.

Let us now turn to the first plausible response to the difficulty encountered vis-a-vis the Univocality Assumption: that of reinterpreting the data. A solution along precisely these lines has been proposed by Joos (1964). Joos, first of all, claims that tense is the category "in which a finite verb ... is either marked with -D or lacks that marker."
Then by definition there can only be two tenses. He then introduces labels for the two tenses that suggest their range of meaning. "The unmarked tense will be called actual and the marked one remote. The latter name fits the meaning precisely. The modern English remote tense has the categorical meaning that the referent (what is specified by the subject-verb partnership) is absent from that part of the real world where the verb is being spoken." But even though this definition is extremely general, it does not even mean what it seems to mean, for we are told, a little later, that "English treats future time as not remote from the present occasion, and remoteness in time in English is always categorically past time" (Joos, 1964, p.121). We learn further that "the modern English remote tense has exactly the same form, no matter whether the meaning is unreality or past reality, with a single exception: replacement of was by were when the meaning is unreality." Concerning the latter category of unreality, Joos says that "unreality specified by using the remote tense consists in a posited substitute for accepted reality being essentially in conflict with the reality. The conflict is mutual -- each contradicts the other absolutely."

We have quoted at length from Joos in order to underscore the tremendous conceptual difficulties of positing some supercategory of meaning that each and every occurrence of a given tense form is an instance of. There are serious questions as to whether Joos has succeeded in his enterprise. It is quite unclear, e.g., how the two claims that remoteness is "absence from the part of the real world where the verb is being spoken", and that future time is treated as not remote from the present
occasion, can be reconciled with each other. And there is further the serious question whether we ultimately arrive at a deeper understanding of English tense, or whether we reach only the goal of satisfying a methodological principle.

To sum up: the first way to meet the dilemma that English tense forms and the Univocality Assumption jointly create, seems inordinately costly. We lose all the precision we had attained in characterizing English tense in simple sentences, and it is extremely difficult to assess what we gain in exchange.

So let us turn to the second alternative: that of modifying the theoretical principle. In considering this option, we will try to capitalize on the obvious fact that the non-temporal interpretations of the tense morphemes are dependent on the immediate linguistic context. Moreover, the non-temporal interpretations are either based on a two-form distinction present/past (rather than on the temporal three-way distinction), or on an obligatory past tense form (in the subjunctive environments).

The minimal modification of the Univocality Assumption necessary in order to separate the regular cases from the restricted ones is as in (26).

(26) Given a freely inserted syntactic constituent, there is exactly one semantic expression that specifies its meaning.

The conception implicit in this revision is this: special context phenomena receive special treatment. The complications involved are negligible. The special contexts of English that require such special treatment can
be enumerated; the lists are short and exhaustive. The striking fact that must be born in mind is that an ordinary sentence like "Mary was a fine teacher" never bears a trace of the concept of 'unreality' in its meaning.

Once having made the decision to adopt (26), we may approach the question of the 'semantic unity' of the syntactic tense morpheme from a rather different perspective. Rather than lumping the temporal category of Tense, and the category dealing with degrees of hypotheticalness, e.g., together, we may ask if there are any interesting conceptual parallels between the internal structure of one and the internal structure of the other. If we find such parallels, we can then attempt to construct hypotheses about the possible ways syntactic forms can be related to domains with similar internal structure.

In summary, we adopt here the modification of (26), clearly separating the fully productive system mapping into semantic Tense from the limited systems dealing with the semantic domains of hypothetical meaning and subjunctive mood. Similarly separated are to be all occurrences of tense forms that arise as a result of agreement rules. Given these rather natural distinctions, our analysis of Tense in English is of a piece with the analyses of Aspect, Epistemic Status, and Taxis.
CHAPTER 5.

THE MODALS

5.1. Introductory Remarks.

In this chapter, we will present a rather selective treatment of the modals of English. We will be chiefly concerned with those aspects of the modal system that shed light on the semantic principles that define the structure of the TMA component. Beyond that, we will work towards a general characterization of the domain of Modality in English that has considerable theoretical appeal. This general characterization will fall short of comprehensive coverage in two ways: of all the modals, only must and should will be examined closely, and secondly, a number of modals will remain entirely outside of the scope of this material. The latter group includes some true modals like shall, dare, need, ought to, can (in its ability sense), as well as such semi-modals as be to and be going to.

In spite of these limitations, we will introduce all of what we take to be the central conceptual apparatus of Modality in English, and project a system from it that has some plausibility, but make no attempt to provide any detailed justification for the proposal made. The latter undertaking would take us into a number of areas -- such as the problem of the interaction of the modals with negation -- which are not, strictly speaking, within the scope of this thesis. Many of these problems are self-contained, in the sense of having no direct relevance to the theoretical framework pursued here. Some interconnections of the modals with the
other verbal categories, however, are relevant in just that sense, and they will receive close attention.

5.2. A study in contrasts: MUST and SHOULD.

We introduce the contrast between must and should with a number of closely matched pairs of sentences, first merely trying to determine the difference between the two members of each pair in an informal fashion as we go along.

Our first scenario concerns Jack, who sells nuts and fruit juices in a health food store. Now let's assume that Jack gets caught drunk at his job one day. What's the difference between the personnel manager saying (1) to Jack, and him saying (2) to Jack?

(1) Jack, I must dismiss you, I'm afraid ...

(2) Jack, I should dismiss you, you know ...

What would we say if we were simply describing (1) and (2)? Well, in the case of (1) we would say that Jack is told that he's being fired. But in the case of (2) we would say that there is a strong suggestion that Jack will, in fact, escape dismissal.

The next scenario is about Mary, Jack's wife, who doesn't know about Jack's troubles, and has splurged on a roast. We hear her saying (3), or (4).

(3) The roast must be done now.

(4) The roast should be done now.

Mary has studied the cookbook which says that three hours is the normal
cooking time. Three hours after she put the roast in the oven the timer goes off, and Mary says (4), not (3). Mary then finds the roast not just quite done yet, but close. To be on the safe side, she lets another 30 minutes tick off on the timer. When it goes off she says (3), not (4).

Or let us grant Mary some more experience with roasts. And as a certain distinctive smell spreads from the kitchen into the living room, Mary says (3), not (4).

Next, we write an instruction manual on how to defuse a time bomb. At the end, we say (5), not (6).

(5) The steps of the defusing procedure must be executed in exactly the order indicated.

(6) The steps of the defusing procedure should be executed in exactly the order indicated.

(5) suggests that we mean business -- one wrong move, and things could get very nasty indeed! If we said (6), we would be taken to dispense friendly advice. Somebody may try a different way of doing it, but experience has taught us that our way is dependable and safe, so why experiment?

If we are indeed dispensing just fatherly advice, should is much more appropriate than must.

(7) I think you work too hard. You should take a day off now and again.

(3) I think you work too hard. You must take a day off now and again.
The question "Who says?" is more natural with (8). In the case of (8) the speaker appears to make an appeal to some authority, perhaps a medical column he has read about the incidence of heart attacks among middle-echelon executives. In the case of (7) he draws on his own wisdom.

The setting for our next scenario is a bilingual region. Now we imagine that truly bilingual individuals are always the product of a couple of whom one member is a monolingual speaker of language A, and the other a monolingual speaker of language B. We'll call such parents heterolingual.

Let's assume now we have been introduced to a girl who, as it turns out, is bright and witty, and moreover fully bilingual. Talking about her later, we comment idly,

(9) Her parents must be heterolingual.

and not

(10) Her parents should be heterolingual.

It seems that, given all our assumptions about the situation, we know too much to say merely (10). When could we use (10) at all? Perhaps if we change our assumptions slightly, and not all, but most really accomplished bilinguals have heterolingual parents, and if we had some cause for making an educated guess. If the heterolingual marriages were only legalized in 1949, we would continue (9) by saying "... which means that they didn't get married before 1949," while we would continue (10) by saying "... which would mean that they didn't get married before 1949." In
other words, (9) says we're sure, while (10) stands on shakier grounds.

Using the same bilingual region as our background, suppose we met a couple who have two children 1 and 2 years old, and we want to make a prognosis about the linguistic skills the two little ones can be expected to acquire. We could say

(11) Their children should grow up to be fully bilingual.

but not

(12) Their children must grow up to be fully bilingual.

But if our goal is not just to make a prognosis but to give strong reassurance, for example to a friend who has just opened a business, then we would make our friend feel better if we said

(13) Your business has such a terrific location, it (simply) must succeed.

rather than

(14) Your business has such a terrific location, it (*simply) should succeed.

One last example. If our friend's business crashes, then his uncle, who had stubbornly refused to invest a penny in the venture, might say either (15) or (16).

(15) You must have realized from the beginning that it wouldn't work out.
(16) You should have realized from the beginning that it wouldn't work out.

But the force of the uncle's comment would be rather different. (15) is a way of ascribing to the nephew economically suicidal tendencies, a certain perverse pigheadedness in going through with the entire business scheme. (16), on the other hand, suggests an error of judgment, something that could have been avoided with a little more experience, or a little less optimism.

We have now looked at the must/should contrast in a variety of settings, over a range of conversational uses (guessing, advising, reassuring, criticizing, etc.), and in both the senses generally distinguished ((1)/(2), (5)/(6), (7)/(8) are presumably 'deontic', while (3)/(4), (9)/(10), (11)/(12), (13)/(14), (15)/(16) are most likely 'epistemic'). In spite of all this variation, there is a strong impression of a common thread running through all the examples. In all cases, without exception, the sentences with must convey something 'stronger' than those with should. Must marks the safer guess, the stronger advice, the more comforting reassurance, the harsher criticism. This clear intuition may or may not stem from a unified semantic analysis that underlies all these uses of must and should. The situation is reminiscent of the one we encountered with Tense. There is considerable intuitive support for the claim made explicitly by Joos' analysis of tense that past time is 'remote' from present reality, and that unreality is also 'remote' from present reality. We would not expect to find a language that uses two
suffixes that are in complementary distribution in such a way that one of them marks present tense and counterfactual conditions, and the other past tense and real ('open') conditions. But we argued in that case that a semantic analysis that attempts to express this intuition directly suffers adverse consequences by losing precision and clarity.

So there are no firm conclusions to be drawn from the intuition of a constant semantic difference between must and should. Nonetheless, the fact that this intuition exists is not neutral; it weighs in favor of a unified semantics.

5.2.1. Let us then go back to a particular case illustrated above, and probe more deeply, in order to see what semantic facts can be extracted from it. Let this case be (3).

We will now look at (3) not from the point of view of how it compares to (4), but rather, how it compares to (17)

(17) The roast is done now.

The basic difference is this: A situationally fully appropriate utterance of (17) commits the speaker to having at his command all the facts relevant to determining the truth of (17). The state of affairs represented by the propositional content "The roast be done now" is addressed directly.

A situationally fully appropriate utterance of (3), on the other hand, indicates, first of all, that the speaker is not in possession of the facts that would warrant his saying (17). Rather, he is in possession of some other facts. In the scenario sketched above, the fact that Mary is in
possession of is that there is a certain characteristic smell coming out of the kitchen. That fact is not the same fact as (17). But it is related to the fact in (17) in a certain way. The situation as described makes it plain that there would not be that distinctive smell unless the roast was done. Thus, in addition to being in possession of the fact that there is that smell in the room now, Mary also knows what inferences can be drawn from that fact. And the propositional content of (3) is thus not a fact known, but a fact inferred. The inference is based on an item of general knowledge -- about the relation between a certain kind of smell and the contemporaneous state of a piece of meat in the oven --, and on the apprehension of a specific fact -- that a smell of just the right kind is now in the room.

Let us rest content with these observations for the moment, and examine in the same detailed fashion a rather different case.

This new scenario finds the boss sitting in his office, smoking a cigar, when two men in overalls come in, walk over to a filing cabinet and make preparations to move it. Hollers the boss:

(18) What the hell do you think you are doing? That cabinet stays right where it is!

He does not holler:

(19) What's the matter with you guys? That cabinet must stay right where it is!

The first thing to note about (18) and (19) is that they do not convey
any truth claims. They are instructions to the men in overalls. If Mary of the previous example had said (3), and then found out that the roast was only charred on the outside and quite raw inside, she would have been wrong. Either the smell she noticed wasn't quite the right smell in the first place, or the general rule that the smell of the roast tells you something about how far along it is just isn't a very good rule. If Mary says (17), and the meat is raw inside, then what she said is directly falsified. She is wrong not about her assumptions or her evidence relative to a fact, but about that fact itself.

The case of (18)/(19) is not like that at all. If the boss says (18), and the men calmly proceed to pick up the cabinet and carry it off, the boss hasn't uttered a falsehood. The men who paid no attention to him most probably made a very wrong move; the boss will be furious and make them pay dearly. As we said, not because he can't stand being caught making a false statement, but because he can't stand not having his wishes respected. This is true, though to a lesser degree, in the case of (19).

But there is an important difference between (18) and (19), nevertheless! By saying (18), the boss simply lays down the law. He gives direct expression to his presumed power. Not so with (19). The situation is subtly different. The sentence still gives an unconditional instruction, but it suggests that there might have been alternatives, while stressing that there turned out not to be any in actual fact. (19) makes an implicit appeal to some higher reason for the command. Given that overarching reason, and the facts being what they are, the
boss feels fully justified in saying what he does. Thus (19) is not an expression of raw power as is (18). The boss is still giving the orders with (19), but he is an enlightened individual, not a tyrant.

The parallels are beginning to emerge more clearly. Corresponding to the "brute fact" in the assertion case (i.e., (17)) we have the "raw command" in the command case (i.e., (18)). Corresponding to the "well-supported inference" in the case of (3) we have the "solidly justified direction" in the case of (19). The parallels go further. Corresponding to the "educated guess" of (4) we have the "reasonable course of action" force of (20) below.

(20) I don't know about you guys. That cabinet should stay right where it is.

There is a striking fact about all the sentences considered so far in this chapter: just to look at them, as well-formed strings of English, we cannot tell the difference between the one kind and the other. Consider, in this connection, (21) below, and how it differs from (8), which we repeat here. The strings compared are exactly the same. But their senses differ along the epistemic/root dimension.

(21) You must take a day off now and again. Otherwise you couldn't have a rich tan like that so early in the season.

(8) I think you work too hard. You must take a day off now and again. Otherwise you'll age prematurely.

But notice that this kind of phenomenon is not limited to the modals. When we see a sentence like (22) out of context, we have at least two ways
of viewing it.

(22) Boy scouts are brave, helpful, and good.

(22) could be said by a scoutmaster to a group of boy scouts to give them an idea of what to strive for. It can serve as a guide for good behavior. If one of the little boys makes a rude gesture, the boy is in the wrong, not the scoutmaster ill-informed.

But (22) could also be used by a student of youth organizations who wants to record the facts he has observed. Here, (22) is a generalization over actual behavior, and if lots of boy scouts are cowards, it's a bad generalization.

All that our semantic analysis of non-progressive sentences in Chapter 2 tells us about (22) is that it represents a structural fact. In particular, it does not tell us whether it is 'stipulated structure' or 'empirically established structure' that a given sentence refers to. But that is as it should be. Our goal was to discover the general meaning of a free syntactic alternation. The analysis we arrived at permitted maximal freedom in one respect: the entity whose structure the analysis referred to was left unspecified. At only one point -- when we discussed 'aborted schedules' -- did we find any reason to think that there are at least some criteria for what can be an entity. But we found no reason to discriminate between 'ideal boy scouts' and 'real boy scouts'.

5.2.2. Let us take the case of (22) as an important precedent of a difference in the use of an expression which does not rest on, or correspond to, any distinction in the grammatical form employed, no matter how deeply
analyzed. Keeping this in mind, we will construct a semantic analysis for (3), and then ask whether or not it is also adequate for (19), the grammar perhaps simply being unspecific on the obvious points of difference between the two. If this approach works for must, we will finally return to the must/should contrast and ask what its status in the emerging analysis is.

As far as the analysis of (3) is concerned, we have already noted all the salient facts, albeit in an informal manner and in terms adapted to the particular case. The task before us now is to propose a more general formulation, and to show how the semantic contribution made by the modals is to be integrated into the meaning representations we have developed in Chapters 2-4.

We propose that the following general picture captures the main points of our detailed description of (3). The central notion is that of a derivation. Its conclusion is indicated by the propositional content of the modal sentence, and the premises are all unspecified. One of the premises is a general statement, and one of them is a particular fact true at the time the modal sentence is uttered. This exhausts all the pieces of logical structure we detected in (3).

We now make this logical structure more explicit in order to permit a more perspicuous exposition.

We will call a law any general proposition to the effect that given a true proposition expressing a certain state of affairs, it follows, or we can conclude, that a second proposition, expressing some other state of affairs, is also true. We will symbolize the generality of such a law in the usual way, by means of universal quantification over whatever class
of entities it is that the proposition is held to be generally true of.
A law will thus be an instance of the formula in (23), where \( F \) and \( G \) are monadic predicate letters.

\[
(23) \quad (x)(F(x) \Rightarrow G(x))
\]

That is to say, a particular law will look like (24).

\[
(24) \quad (x)(F_1(x) \Rightarrow G_1(x))
\]

where the predicates \( F_1 \) and \( G_1 \) have replaced the predicate letters \( F \) and \( G \).

We will say that a derivation is a set of propositions \( P \), which we will call premises, plus an additional proposition \( C \), which we will call the conclusion, such that \( C \) follows from some subset of \( P \) by a rule of inference, and such that \( P \) includes a law. We use this last definition to introduce three predicates, as follows: '\( D(x) \)' for '\( x \) is a derivation', '\( P(x,y) \)' for '\( x \) is a premise in \( y \)', and '\( C(x,y) \)' for '\( x \) is the conclusion of \( y \)'.

Armed with all this apparatus, we return to (3), which we repeat for convenience.

(3) The roast must be done now.

Taking '\([d_r]\)\) to represent 'the fact of the roast being done now', we give the preliminary analysis of (3) as (25), including Taxis, but excluding Tense and, of course, Modality.
(25) \((\exists x)(\exists y)(T(x) \& T(y) \& DUR([d],x) \& DUR(z,y) \& x\leq y)\)

The addition of the semantic contents of **must** yields (26):

(26) \((\exists x)(\exists y)(\exists w)(T(x) \& T(y) \& DUR([d],x) \& DUR(z,y) \& x\leq y \& D(w) \& C([d],w) \& P(z,w))\)

Even as rough as this is, it contains a glaring inadequacy: it treats the narrated situation and the reference situation as though they were on a par. But we have insisted that only the reference situation is actually asserted. So, without rewriting (26), we consider an extra conjunct '\(\downarrow\) DUR(z,y)' to have been added to it.

5.2.3. There is an apparently unmotivated claim included in (26). This claim is that the reference situation (which was introduced into the semantic representation of all sentences as a consequence of there being a grammaticalized element of Taxis in English; cf. Chapter 3) is designated by the analysis of **must** as a premise in a derivation whose conclusion is the narrated situation. There is, however, strong support for this claim in the form of three arguments which we will take up in the next three subsections.

5.2.3.1. It is a fact about English that a sentence like (27)

(27) That must be Harry at the door.

reports an inference based on **presently available evidence**.

It is also a fact that (28), while different with regard to the temporal placement of the inferred state of affairs, is analogous to (27), in
that the inference it reports is also based on presently available evidence.

(28) It must have been Harry at the door.

If a speaker wants to convey the information that the evidence on which the inference depends on was available in the past, he must resort to a periphrastic construction, as in (29).

(29) It had to be Harry at the door.

But setting the availability of periphrastic constructions aside, the facts are these: modal sentences allow the inferred event or state to be in the present (cf. (27)); they allow it to be in the past (cf. (28)); and they allow it to be in the future (cf. (13)). But as far as the time of the evidential fact is concerned, there is only one possibility: it must be in the present. This fact is as yet unexplained.

Recall now the syntactic analysis of tense and of the modals sketched briefly in 4.1. The tenses and the modals are represented as mutually exclusive forms. In the case of the future tense vis-à-vis all the modals this follows from the theoretical assumption that there can only be one category marked [-Comp, +Det] embedded under any node of the form $X^3$. In the case of the affixal past tense and the modals, no such general considerations apply; the facts are simply expressed by introducing a null morpheme in stem position with past tense, and in affix position with the modals.

Consider now what the consequences of this syntactic state of affairs
are for the interpretation of tense in modal sentences. The Semantic Closure Principle together with the Minimal Satisfaction Principle guarantees that such sentences must be assigned a specific interpretation with respect to the domain of Tense. The sentences in question are unmarked for tense. Thus the Complement Convention determines the complement of the marked forms as the meaning to be assigned. The marked forms are past and future, and their complement is the present. This determines present tense as the appropriate tense specification for modal sentences.

In section 4.2, tense was defined as the verbal category that specifies which temporal relation obtains between \( t_0 \) and the time interval during which the reference situation prevails. This means that the reference situation in all modal sentences will be present.

Two distinct facts have been established in this section so far: (i) the evidence that the inferences reported in modal sentences are based on is always present evidence, never past or future/ and (ii) the tense assigned to the reference situation in all modal sentences is present, never past or future. By assuming that the rules interpreting the modals designate the reference situation as a premise (i.e., as evidence), the two facts described above are related to each other: (i) is now entailed by (ii) plus the operation of the modal rules.

5.2.3.2. We have argued in the preceding subsection that modal sentences are assigned present tense. In section 3.3.5 we gave an explanation for why perfect sentences in the present tense with specific past time adverbs like yesterday occurring in them are ill-formed. An example of such a
sentence is (30).

(30) Brian has gone to New York yesterday.

This explanation was based on a construal of the notion of 'relevance' (of the narrated event to the reference situation) as so weak that it could not be sustained across temporal gaps. The offending configuration was as in (31).

(31) yesterday reference
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
    |
formedness of (30), is now no longer baffling; it is in fact predicted, since the logical relation of deducibility is independent of temporal considerations of the kind involved with relevance.

5.2.3.3. The third argument for treating the reference situation as a premise comes from considerations involving the non-perfect, the unmarked element of Taxis. In 3.4 we showed what kind of interpretation for the non-perfect our theory led to. This interpretation was that the reference situation was either simultaneous with, or earlier than, the narrated event. This interpretive result turned out to be without any semantic consequences in non-modal sentences. Appealing to a minimality principle as a criterion for the choice of the reference situation, we postulated that the reference situation was simply identical to the narrated event in all such cases. The logical consequence was that the formulas in which the non-perfect was not included and those in which it was were equivalent.

If, however, the reference situation is a premise in a deduction that the narrated event is the conclusion of, they cannot be identical (leaving aside trivial cases that we have not bothered to exclude here, where p is deduced from itself). Thus the semantic specification given for the non-perfect cannot be trivialized. This means that in non-perfect modal sentences the narrated event must be either in the present or in the future, and cannot be in the past, given that the reference situation is invariably in the present.

We illustrate with should that this is exactly what we find.

(33) They should be just about ready now.
(34) They should be ready by next Friday, at the latest.

(35) #They should be ready yesterday, but something went wrong.

These arguments suggest very strongly that the decision to designate the reference situation as a premise is correct. They also serve to strengthen the overall analysis of the English TMA component advanced in these pages.

5.2.4. The next question to be taken up is this: Assuming that (26) is an insightful semantic analysis of (3), does any of it carry over to (19)? That is, can one go from epistemic must to root must, or must the analysis of root must start from scratch?

To approach this matter let us consider an example using root must that is as close to a clearly structured paradigm case as we can come. Our scenario posits the popular stereotype of a Sicilian code of honor. This honor code, let us say, has a simple rule of vendetta by which any grievous harm inflicted on a member of a Sicilian family must be avenged by the closest male relative of the victim. In this less than idyllic setting, disaster strikes. Giorgio kills Felice's father. Grimfacedly Bruno, Felice's uncle, tell him:

(36) Felice, my dear nephew, you are your father's closest living relative; you must go out and avenge him. There must be no other thought in your mind until justice is done.

Melodramatic as this is, what is its analysis? Certain aspects of the situation are reminiscent of Mary's roast. Here, as there, the must draws its
force from a covering law. In the present case it is the honor code. In both cases something happens that makes the covering law applicable. Here, the triggering fact is the murder. And in both cases, the covering law determines the consequences of the triggering fact.

But the obvious difference is this: in the case of Mary's roast the consequences determined by the covering law are real-world facts; in the case of the Sicilian vendetta, the consequence is not a real-world fact, but an ideal-world fact.

This last remark needs to be clarified (and in trying to do so, we will draw the same distinction that is discussed in Wertheimer (1972) in terms of Ideal Systems vs. Actual Systems, in much the same connection). Recall our observation that sentence (22) above could be a claim either about 'ideal boy scouts' or about 'real boy scouts'. We suggest that this distinction carries over to laws in general. There are ideal laws, and there are actual ('real') ones. The difference is that ideal laws describe conventions, while actual laws describe the real world. Only actual laws are open to empirical challenge. Conventions, by definition, do not purport to reflect empirical data, so they cannot be tested against them; they are neither true nor false, but they can be respected or violated. If the antecedent of an actual law is satisfied, and the consequent fails to hold, then we conclude that the law is wrong and must be discarded or revised. This conception is the standard one in scientific endeavor, in linguistic theory for example. If, on the other hand, the antecedent of an ideal law is satisfied, and the consequent fails to hold, then we do not conclude that the law is wrong, but rather that the state of affairs that is not in conformance with it is wrong. This is typical of
conventional systems, e.g. moral systems, etiquette, games, etc.

We appeal to actual laws in order to infer what the real state of affairs is, and to ideal laws in order to infer what the proper state of affairs is. And to subscribe to an ideal law just means to make every effort to ensure that the real state of affairs does not deviate from the proper state of affairs.

This, then, is our reconstruction of the epistemic sense vs. root sense distinction: the former sense is associated with actual laws, and the latter sense with ideal laws. There is no question but that the distinction is solid, real; but there appears to be no reason to say that the English modals make the distinction.

Our analysis of must in (26) gives a rich semantic characterization of this morpheme of English. The only point at which it would have to be modified to encompass both epistemic and root uses is in the definition of a law, which is now tailored to the case of actual laws. This modification would in turn have consequences for the exact definition of a derivation, and would also have ramifications for the interpretation of the conclusion, i.e. whether it is to be construed as reflecting the actual state of affairs, or the proper state of affairs. But the structure of the entire analysis would remain exactly as before.

On the basis of these considerations it seems that a uniform semantic analysis of the English modals suggests itself. This analysis refers to laws unspecified for whether they are actual or ideal, and leaves this matter of interpretation open to be determined one way or the other within the discourse. Note that this is not at all parallel to the case of
tense. With tense, the cases we argued should be separated out for special treatment were distinguished by (i) having a defective paradigm, and (ii) being dependent for their correct interpretation not on discourse facts, but on intrasentential facts about cooccurring morphemes. The parallels are much closer with the case of Epistemic Status where what appeared to be clearly different senses (e.g. the Instantaneous Present of sports commentaries vs. the Programmed Future of schedules and itineraries) were shown to fall together from the point of view of that semantic category; the fact that, outside of that particular perspective, a sports commentary and a schedule are really very different kinds of entities was not even addressed by the analysis, since it leaves the entities it applies to unspecified. Those entities must be fixed for each normal utterance of a sentence, to be sure. But they are supplied from the discourse context, and not by the sentence grammar.

Thus we conclude this section not with a sense of having demonstrated the univocality of must, but of having made it a sufficiently attractive hypothesis by showing exactly what univocality would consist in, that on this matter of whether there is a sense distinction to be made here as a matter of the analysis of the modal morphemes, the burden of proof has been reversed.

5.2.5. We are now ready to return to the original concern of section 5.2: What is the difference between must and should?

We propose that the contrast arises from a binary judgment speakers of English must make when they wish to present a state of affairs as an
inference deduced from a covering law and some particular asserted fact, by means of a modal. They have to decide whether the law is one that applies unconditionally, or whether it is applicable 'as long as normal circumstances prevail'. A law of logic would in terms of this binary distinction always be of the former type -- there are no circumstances that make such a law inapplicable. From $a \& b$ we can always infer $a$, no matter what. The generalizations we find in cookbooks, on the other hand, e.g. about cooking times for various dishes, belong fairly clearly to the other type. The decision whether a given law is unconditional or not is obviously not a grammatical one. People may differ in how they classify laws. The rules of etiquette, e.g., may be taken by some to be absolutely and slavishly obeyed, while others may treat them as general guidelines to be used with some mature judgment.

**Must** and **should** reflect this difference directly: **must** is based on an unconditional law, and **should** is based on a contingent one. This can be made more precise. We will say that all laws have one of two canonical forms. The first is illustrated by the general schema in (23). The second is given in (37) below.

(37) $CM(x) (F(x) \& G(x))$

$C$ stands for any conjunction of sentence letters; but the point about it is hardly its formal definition. Of interest is its substantive characterization: it represents 'normal circumstances'. This is not a very precise notion, but an intelligible one nonetheless. The statement of a law will generally limit itself to the salient aspects of the situation. The normal circumstances will be implicit. So for example in our case of
the cookbook: what will be specified for each recipe is whether it takes
35 minutes in the oven or 60. That these values are for the average oven
under average operating conditions, and that there might be some variat-
ion in actual cooking times depending on a variety of circumstances, is
not all spelled out there. Nevertheless, the cooking time indicated is
not fixed by convention; it is intended as an empirically determined value.

So there are then 'reasonable' empirical generalizations, with room
for error, which correspond to rough-and-ready conventions, with some lee-
way for discretion: these are instances of schema (37). And then there
are rigorous scientific generalizations and highly specific conventions,
without any margin for error: these are instances of schema (23).

Let us -- purely for convenience -- call the quantified part of (37),
'(x)(F(x)\rightarrow G(x))', the salient part of a circumstantially contingent law.
Then the precise meaning of should is that it asserts an instantiation of
the antecedent of the salient part of such a law. About the conditioning
circumstances it says nothing.

Now we are in a position to go back to the examples we gave at the
beginning of this section, and check to see whether this analysis capt-
ures insightfully how must and should differ. We will be content with do-
ing this for just one example, (1)/(2).

Suppose for both (1) and (2) the covering law is a rule to the effect
that employees that are drunk while on duty are to be dismissed. If the
personnel manager uses must in the situation described, this conveys that
he is satisfied to apply the rule as it stands. No ifs and buts. That's
why we feel sure that (1) is to be construed as a dismissal.
If, on the other hand, the manager uses \textit{should}, he signals immediately that there are discretionary elements in the situation. Given normal circumstances, dismissal would still be the outcome. But we have said that with \textit{should} circumstances are not asserted to be normal. There are two possible reasons for this, in a cooperative conversation: (i) the circumstances are not asserted because they are just not known -- in this kind of case an eventual dismissal is still entirely possible --; or (ii) the circumstances are not asserted because it is known that they do not obtain. But then the rule is not applicable and there will be no dismissal. This latter case is why we felt that in the case of (2) Jack might well escape being dismissed.

5.3. A Projection from the results.

This section will be unabashedly speculative. We will project from the result we have culled from the data about \textit{must} and \textit{should} to the full system implicit in this result. We will suggest which modals belong where in this system, and leave it to future research to bear us out on every point.

The structure of the system we propose is very simple, and dictated by what conceptual building blocks we have proposed here. The point of variation beyond the \textit{must}/\textit{should} contrast is this: If \textit{should} can assert the antecedent of the salient part of a law, and be silent about circumstances, then there is the further possibility (i) that there might be a modal that does the opposite: assert the circumstances, but say nothing about the salient part; (ii) that there might be a modal that asserts both. Without bringing negation into the picture, this exhausts the
possibilities implicit in our analysis of should.

The second building block we propose to use is based on negation in a very specific way. We postulate the equivalence that interdefines necessity and possibility in most modal logics as a linguistic redundancy rule relating pairs of modals, as in (38).

\[(38) M_i(p) \rightarrow \sim M_i^*(\sim p)\]

The structure of the system that results is diagramed in (39). At each terminal we will write the pair of related forms as '\(M_i/M_i^*\)'. Two very clear exceptions to univocality should also be pointed out: the will of (39) is not the will of the future tense, but the modal will, as in "That'll be the mailman"; and the can of (39) is not the can in the sense of be able to, which still leaves a variety of uses such as the permission use in "You can do what you like", and an epistemic use as in "It can get very hot here".

\[(39)\]

\[
\begin{array}{c}
\text{applicability of law} \\
\text{unconditional} & \text{contingent} \\
\text{must/can} & \text{which antecedent conditions} \\
& \text{are satisfied} \\
& \text{one} & \text{both} \\
& \text{circumstantial} & \text{salient} \\
\text{would/might} & \text{should/could} \\
\end{array}
\]

This then is what we bequeath to future research on the English modals. And on this forward-looking note we end.
FOOTNOTES

1. The hypothesis that no more than one specifier can occur in any full phrase (a phrase whose most inclusive node is of the form \(X^3\)), and that it is attached at the \(X^3\)-level is bound up with the question of how to constrain local transformations, for example. One of a small number of alternatives Emonds (1970) considers for a general characterization of what were then call 'minor movement rules' was that they should be restricted to moving only "function words". The natural equivalent of "function words" in Jackendoff's theory of the base are the categories marked \([-\text{Comp, +Det}]\). But there are no clear criteria offered in Jackendoff's theory for membership in such a category. The restriction suggested above provides at least some criteria. Aside from local transformations, specifiers also appear to play a significant role in characterizing a class of syntactic rules that apply across lexical categories, such as \(\text{wh-Movement}\); a number of rules of this type appear to be triggered by, or dependent on the presence of certain specifiers.

2. This differs from Jackendoff's own analysis, where the recategorization is effected at the double-bar level. The considerations that motivate recategorization at the single-bar level are as follows: (i) the absence of \(\text{have+en}\) and \(\text{be+ing}\) in participial complements becomes automatic, and (ii) the fact that such sentences have only one set of \(V^2\)-complements, instead of two (one for the higher \(V^2\), and one for the lower \(V^2\)) is also an automatic consequence of the phrase structure posited.

3. The interested reader will find a proposal for a formal explication of the notion 'event' in Katz (1972); Katz uses the label 'process', and
defines 'process'-markers in terms of a sequence of (at least two) 'condition'-markers, where 'conditions' are what is generally referred to as 'states'. This formalization has not been used in the text since it adds nothing of substance to the discussion.

4. The formula in (21) is not completely defined, since $\text{DUR}(x, y)$ was introduced with a meaning specific to events, and has no meaning for pauses. But we will generalize the meaning of $\text{DUR}(x, y)$ later in the text anyway, and will simply assume that the extension to states proposed there is sufficient to cover the case of pauses as well.

5. It appears that writers on Aspect agree that the term is used fully appropriately to refer to at least one other range of cases which is roughly characterized by quantification over whole events. Labels like habitual, usitative, iterative, and generic, are often used to suggest the meaning of categories of this type. The present work, with its explicitly restricted goal of analyzing the English auxiliary system, has no bearing on this matter, except in a curiously negative way. In section 4 of Chapter 2, an analysis of the 'epistemic' non-progressive is given which implies, among other things, that at least in English 'habitual aspect' or 'generic aspect' play no role as semantic categories, the traditional lore notwithstanding.

6. The inclusion of $N^3$ in this pattern is prompted by the fact that with begin, start, finish, NP-complements are the more plausible candidates for alternation with the participial complements. He began preaching
to the crowd is more closely related to He began his sermon than to He began with a call for generous donations. Similarly, He finished the installation of the airconditioner before lunch is the counterpart to He finished installing the airconditioner before lunch. In any event, allowing N³ to participate in the recategorization pattern makes an obvious problem quite apparent. The desirable property of a recategorization of Xᵢ to Xᵢ is that it explains directly why the phrases associated with Xⱼ and Xⱼ are invariably absent. But now, what needs to be explained is why in some cases the phrases associated with Xᵢ and Xᵢ are also invariably absent. So while (25) makes plain why the lower verb has no tense, no perfect, no progressive, etc., it does not explain why the phrase structure possibilities inherent in N³ and P³, or in N² and P², are not realizable. The usual pairs like He kept at it and He kept turning the crank raise no question on that point. But a pair like He kept right at it and He kept right turning the crank does, and suggests that more needs to be said than our exposition would lead one to believe.
BIBLIOGRAPHY

Bierwisch, M. (1967) "Some semantic universals of German adjectives", Foundations of Language, 3.1


Ross, J.R. (1972) "Doubl-ing", *Linguistic Inquiry*, 3:1

BIOGRAPHICAL NOTE

The author was born, raised, and educated in Austria, the entire process starting in 1942. He left college as a 'certified merchant', and, from 1967 to 1969, interfered with South Africa's economic development by working there as an 'investment analyst' (a euphemistic job description neither native nor exclusive to South Africa). Lucky enough to stumble across a copy of Whorf's Language, Thought, and Reality, he decided he wanted to become a linguist. After three years as an undergraduate student at Central Michigan University, and four years as a graduate student at MIT, he is well on his way towards that goal.