Chapter 1
Introduction

... to investigate speech as a natural phenomenon, much as a physiologist may study the beating of the heart, or an entomologist the tropisms of an insect, or an ornithologist the nesting habits of a bird.
George Kingsley Zipf, *The Psychobiology of Language*

This book argues that the Surface Syntax of natural language acts as a completely transparent interface between the spoken form of the language, including prosodic structure and intonational phrasing, and a compositional semantics. The latter subsumes quantified predicate-argument structure, or Logical Form, and discourse Information Structure.

That is to say that although surface forms of expression in all languages notoriously disorder elements that belong together at the level of meaning, and although information-structural distinctions like theme and rheme appear somewhat independent of traditional predicate-argument structure, there is a theory of grammatical operations that allows a unified semantic representation to be built directly from surface forms, without the intervention of any intermediate level of representation whatsoever. According to this theory, syntax subsumes Intonation Structure and semantics subsumes Information Structure, the two standing in what Bach (1976) has called a "rule-to-rule" relation. This means that each syntactic rule is paired with a rule of semantic interpretation, such rules being entirely compositional—that is, defined as a function of the interpretation of the constituents to which it applies. This position is closely related to the by now widely accepted requirement for "monotonicity" among modules of grammar—that is, the requirement that no component of grammar should have to modify a structure resulting from an earlier stage.

This is not a particularly startling claim, since some such direct relation between sound and meaning might be expected on evolutionary grounds, or from any other standpoint for which considerations of parsimony and economy are paramount. A similar assumption lies behind programs as apparently different as those of Montague (1970) and Chomsky (1995). However, the nature of the phenomena manifested by the languages of the world has made it extraordinarily difficult to deliver such theories, and in practice most theories of

natural language grammar have drastically restricted their coverage of those phenomena, or have introduced additional levels of representation over and above the one directly related to meaning, or both. Such additional structural representations are often more than a dispensable notational convenience, for they tend to be accompanied by nonmonotonic structure-changing operations of movement, deletion, copying, reanalysis, restructuring, and the like. Insofar as such relations are predicated over structures at more than one level of representation, they compromise the assumption of transparency with which we began. They also tend to bring with them an increase in expressive power, jeopardizing the explanatory force of the theory.

This book argues that the problem with such theories stems from an understandable but mistaken desire to keep the rules of syntax and the Surface Structures that are implicit in syntactic derivation as close as possible to those that are manifested in the meaning representation. In the case of rigidly fixed word order languages like English, this temptation is particularly strong. An example that in fact provided the starting point for the present study is Transformational Grammar of the *Syntactic Structures/Aspects* vintage (Chomsky 1957, 1965). In such grammars, surface forms differ very little from underlying forms, and then only in very specific respects. For example, the dominance relations among the local complements of ditransitive verbs may be rearranged or "wrapped," a relativized *wh*-argument may "move" from the neighborhood of its verb in the underlying representation to the neighborhood of the root node in the Surface Structure of an arbitrarily complex relative clause, or constituents of one conjunct or the other that are represented in the underlying structure of a coordinate sentence may be deleted at Surface Structure. In most cases, however, the Surface Structure residue of wrapping, *wh*-movement, or deletion under coordination preserves as much as possible of the predicate-argument structure that provides its input.

Such analyses have captured profound insights into the nature of such phenomena and have provided the first systematization of the data on which all subsequent formal theories of syntax, including the present one, have built. However, there are two things wrong with this kind of account as a theory. First, as soon as we admit the possibility that the structure implicit in a derivation is not isomorphic to the structure implicit in interpretation, there is no reason why the derivation should bear any simple structural relation to the meaning representation. All that matters is that the derivation yield that interpretation. Indeed, in less rigidly ordered languages than English it is inevitable that the structures implicit in the derivation must be quite distant from those

corresponding to canonical semantic representation.

The second error lies in viewing Surface Structure as a level of representation at all, rather than viewing it (as computational linguists tend to) as no more than a trace of the algorithm that delivers the representation that we are really interested in, namely the interpretation. Here an example may help. The Augmented Transition Network (ATN) of Woods (1970, 1973) was conceived in part as a direct implementation of Transformational Grammar, and the finite state transition networks that constitute the core of an ATN grammar correspond fairly closely to rules of a context-free surface grammar. However, one of the many interesting properties of the program was that it did the work of transformational rules like "passive," "dative movement," "*wh*-movement," and "deletion under coordination," without ever building any representation of surface structure. Instead, certain annotations or "augmentations" to those rules, specifying storage in, retrieval from, or testing of a fixed number of registers or storage locations, allowed the ATN to build the equivalent of Deep Structures incrementally, as a side effect of the analysis. For example, the analysis of a surface subject of a sentence would initially cause a pointer to a representation of that subject to be allocated to a SUBJ register. However, on encountering passive morphology later in the sentence, the program would transfer the pointer to the OBJ register. Even more interestingly, encountering a relative pronoun would cause a pointer to its head noun to be placed in a special relativization register called HOLD. At any point in the subsequent analysis of the relative clause, the contents of this register could be accessed in lieu of analyzing an NP or other argument in situ. The device thereby achieved the same effect as an unbounded movement.[1]

The interest of the ATN for present purposes lies not in providing a particularly constrained alternative to the transformational analysis—it is in fact too faithful a reconstruction of Transformational Grammar for that—but in showing that for both bounded and unbounded constructions, even in an *Aspects*-style theory, Surface Structure need not be viewed as a representational level. The interest of this observation lies in the fact that, if the statement of such analyses does not require transformations as such—that is, rules mapping entire trees onto other trees—then much of the motivation for keeping deep and surface analyses in line vanishes. (Such similarity simplifies transformational rules but is likely to be irrelevant to the more dynamic alternative.)

These two observations suggest that it might be possible to free derivation from predicate-argument structure in a much more radical way than either *Aspects*, the ATN, or most of their successors have done. Such a move has

purely linguistic attractions, because once we move beyond the comparatively restricted bounded constructions like passive and the various sorts of raising, and beyond *wh*-movement among the unbounded constructions, there are in fact many constructions that appear to be much less dependent on the notion of structure that is relevant to predicate-argument structure. Coordination itself, though treated in illuminating ways in both *Aspects* and the ATN, is such a construction, because it can on occasion both delete and leave as residue objects that are not traditional constituents at all, as in the English "gapping" construction (Ross 1967):

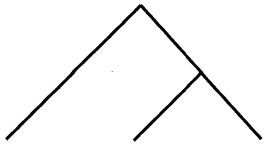
(1) I *want to try to begin to write* a musical, and *you, a play.*

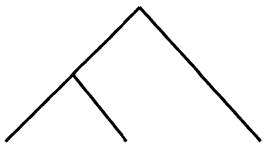
Other grammatical domains in which similarly free notions of structure seem to be needed are intonational phrasing and parentheticalization.

On the basis of evidence from such phenomena, together with other constructions involving unbounded dependencies, a number of recent theories based on Categorial Grammar make the claim that substrings like *Anna married* are possible surface syntactic constituents of sentences, even in simple examples like the following:[2]

(2) Anna married Manny.

According to these theories, even such minimal sentences have two possible Surface Structures:

(3) a. 

b. 

More complex sentences like *Harry says that Anna married Manny* may have many Surface Structures for each reading.

Such Surface Structures do not exhibit traditional notions of constituency, nor do they embody traditional notions of dominance and command. It follows that we must assume that all the grammatical phenomena that are typically explained in terms of such relations—notably, binding and control—are defined at the level of underlying or logical form, rather than at Surface Structure. (Such a proposal is foreshadowed by much work in Montague Grammar since Bach and Partee 1980, and in Government and Binding Theory (GB) since Lasnik and Saito 1984.) By not attempting to define such relations over Surface Structures, we continue to escape the need to represent Surface Struc-

ture explicitly, as a full representational level in addition to Logical Form, and thereby attain a more parsimonious theory, preserving desirable properties such as monotonicity and monostratality (see Steedman 1996b).

The above remarks concern the architecture of "competence" grammar, to use Chomsky's (1965) term, as distinct from the psychological "performance" mechanisms that actually deliver grammatical analyses (and may even on occasion systematically fail to deliver them). Although we will assume in what follows that syntax, semantics, and the workings of the processor are very tightly coupled, it does not follow that observations concerning any one of the members of this triad will necessarily be equally suggestive of insights into the nature of the system as a whole. Chomsky has repeatedly insisted that syntactic form, and syntactic form alone, provides an accessible and reliable source of information regarding the system as a whole. Although this is presented as a fact of research methodology, rather than as a necessary truth, the insight has been so productive of generalizations and lasting results that it is inconceivable that we should abandon it now. It is perhaps worth stating explicitly why this should be.

The methodological priority of competence over performance follows from the exceedingly fortunate fact that the computation involved in mapping natural language strings onto their interpretations appears to be one for which a "theory of the computation" can be specified independently of the particular algorithm that is actually used to perform the computation. That is, it is what Marr (1977) calls a "Type I" problem, or in Jerry Fodor's (1983) terms, a "modular" system.[3] We know that even for quite simple classes of language, there are infinitely many processing algorithms. Of course, relatively few of these are even remotely plausible in psychological terms. Nevertheless, in the absence of an adequate formulation of exactly what it is that the corresponding algorithms for natural language compute, we are unlikely to make progress in understanding the system as a whole on the basis of performance data.

Of course, technically speaking, there are also infinitely many grammars that will capture any given language, in the "weak" sense of defining the set of all and only the strings of the language. Here we are helped (up to a point at least) by the fact that very few of those grammars support a semantics that captures our fairly strong intuitions about meaning at the level of words and their predicate-argument relations, propositions, referring expressions, and relations among them, such as coreference and entailment. However, these aspects of meaning seem to capture only part of the semantics, and to arise at one remove from the level that shapes the superficial form of utterance. It has proved to

be quite hard to explain the very diverse ways in which the same propositional meaning can be grammatically realized as an (active, passive, topicalized, cleft, elliptical, intonationally marked, etc.) sentence, in response to the demands of context. For that reason, standard notions of semantics do not seem to tell us enough for us to completely understand how utterances give rise to meanings. This process seems to be very opaque to introspection, and our most reliable source of information about it seems to be syntactic structure itself.

Indeed, one of the implications of the theory presented below is that contemporary theories of syntax have not been nearly ruthless enough in avoiding the temptation to draw strong conclusions from certain compelling but misleading intuitions about sentence meaning that confound two related but distinct components of meaning, only one of which is directly related to surface syntactic structure, properly understood.

Despite the methodological priority of competence grammar, it and the performance mechanism must have evolved as a package, and in the end theories of language must be evaluated as theories of both components. The significance of monotonicity in rules and monostratalism in representation partly lies in performance considerations. Such grammars make it easier to understand how a processor can make direct use of the competence grammar, a position related to the "Competence Hypothesis" widely endorsed within the generative tradition (see Chomsky 1965, 9; Chomsky 1975a, 7).

The book also investigates this relation. Since CCG integrates Intonation Structure and discourse Information Structure into the grammar itself, the claim is that it is directly compatible with parsing techniques that couple the resolution of local ambiguity and nondeterminism during processing to the semantic coherence and contextual appropriateness rival local analyses. There is considerable experimental evidence that this sort of mechanism is involved in human sentence processing. Although all grammatical theories are in principle compatible with such a mechanism, the one proposed here is more directly compatible. In particular, a processor for this kind of grammar does not need to build any structures other than those defined by the competence grammar itself in order to work in this way. The processor is therefore compatible with an extremely strict version of the principle that Bresnan and Kaplan (1982) endorse under the name of the *Strong Competence Hypothesis*. It will be convenient to refer to the present even stricter version as the *Strict Competence Hypothesis*.

The enduring methodological priority of the study of competence in this endeavor means that much of the material considered below must be presented in purely linguistic terms, to stand or fall on the usual criteria of that discipline.

Nevertheless, it will be clear throughout that the investigation is driven by questions about the language-processing system as a computational whole.

The main body of the book is therefore divided into three parts, the first two of which are entirely concerned with competence, and the last of which returns to questions of performance mechanisms and computational issues.

Part I, which can be read as a self-contained introduction to CCG, is entitled "Grammar and Information Structure." It argues that that surface syntactic derivation in CCG not only specifies Logical Form compositionally and monotonically, but also directly specifies the Intonation Structure of spoken English, and the partition of utterance content according to discourse-semantic relations that intonation conveys, and which is often referred to as Information Structure. The claim is that the traditional notion of Surface Structure can be entirely replaced by a freer notion of surface constituency corresponding to Information Structure, and that this is the only derivational notion of syntactic structure that is linguistically necessary. The argument in part I goes as follows: Chapter 2 provides some more detailed motivation for a radical rethinking of the nature of Surface Structure from coordination, parentheticalization, and intonation. Chapter 3 then outlines CCG in terms of simple examples that motivate the individual rule types. Chapter 4 defines the space of possible CCGs more exactly, and it briefly summarizes the ways in which some basic constraints on bounded and unbounded constructions, including some apparently "nonconstituent" coordinations, emerge as inevitable consequences of the theory. It also briefly considers the way in which scope ambiguities for natural language quantifiers can be handled within a strictly monotonic grammatical framework. Chapter 5 concludes part I by showing that the level of Intonation Structure identified by phonologists such as Selkirk (1984) can be directly subsumed under surface syntax as it is viewed here. It also shows that the level of Information Structure identified by discourse semanticists such as Halliday (1967b) and Prince (1986) is captured in the semantic interpretations that the grammar compositionally assigns to the nonstandard constituents of Surface Structure/Intonation Structure in this sense of the term. This chapter completely revises and extends my first attempt to solve this problem in Steedman 1991a to cover a wider variety of tunes and informational constituents including "discontinuous" themes and rhemes. It includes some discussion of the relation of the present proposals to other contemporary linguistic theories.

Part II continues the development of the grammatical theory in a more technical direction and consists of two connected case-studies. The first, in chapter 6, concerns the notorious cross-serial multiple dependencies characteristic

of the "verb-raising" construction in Dutch (Huybregts 1984; Shieber 1985), a construction whose analysis is known to require an expressive power in the grammar greater than (but apparently not *much* greater than) context-free. This phenomenon is examined in interaction with the earlier account of extraction and coordination, revising Steedman 1985 (where I first looked at these questions in an earlier formalism) and taking account of more recent developments in the theory, extending the analysis to Germanic main-clause order, building on work by Hepple (1990). Chapter 7 looks at gapping in English and Dutch, revising and extending an account first sketched in Steedman 1990 via a new categorial theory of medial gapping and its relation to verb-initial coordination, to cover a wider range of constructions and the full range of language universals relating the direction of gapping to basic word order proposed by Ross (1970), including certain supposed counterexamples to Ross's generalization, Zapotec and Dutch itself among them. Taking the introductory chapters 3 and 4 as read, these chapters constitute a second self-contained (and more technical and purely linguistic) monograph, and they may well be skipped on a first reading. Nevertheless, I make no apology for their inclusion, for it is as a theory of multiple unbounded dependency and coordinate structure that I believe CCG has most to offer as a theory of grammar.

Part III turns to issues of computation and human grammatical performance. Chapter 8 examines questions of expressive and automata-theoretic power, the question of why natural grammars should take this combinatory form, and the general nature of the parsing problem for CCGs. Chapter 9 then discusses a specific architecture for a parser, including the role of semantic interpretation. It is at this point that the expectation of a direct relation between grammar and processor that was announced at the start of this introduction is shown to be well-founded. Chapter 10 summarizes the architecture of the theory as a whole, its role in acquisition and performance, and its relation to other theories of grammar.

PART I

Grammar and Information Structure