PART I

Grammar and Information Structure

BLANK PAGE

Chapter 2

Rules, Constituents, and Fragments

Only connect! That was the whole of her sermon. Only connect the prose and the passion .... Live in fragments no longer.
E. M. Forster, *Howards End*

The artificial languages that we design ourselves, such as logics or programming languages, exhibit a very strong form of the rule-to-rule relation between their semantics and the syntax as it is defined in the textbook or reference manual. This condition in its most general form means simply that there is a functional relation mapping semantic rules and interpretations to syntactic rules and constituents. We will return to the nature of the mapping and its consequences below, but the function of syntax as a vehicle to convey semantics makes the requirement of simplicity in the rule-to-rule mapping seem so reasonable and desirable that it might be expected to transcend all particulars of function and content.

When we finally come to understand the natural system, we must therefore expect to find a similarly direct relation between syntax and semantics, for it is hard to imagine any evolutionary pressure that would force it to be otherwise.

Indeed, there is at least one identifiable force that can be expected to work positively to keep them in line. It arises from the fact that children have to learn languages, apparently on the basis of rather unsystematic presentation of positive instances alone. Since under the simplest assumptions even such trivial classes of grammars as finite-state grammars are not learnable from mere exposure to positive instances of the strings of the language (Gold 1967), and since there appears to be little evidence that any more explicit guidance is provided by adults (Brown and Hanlon 1970 and much subsequent work), some other source of information, "innate" in the weak sense that it is available to children prelinguistically, must guide them. While statistical approximation and information-theoretic analysis using unsupervised machine learning techniques over large volumes of linguistic material remains a theoretically interesting alternative, the most psychologically plausible source for the information children actually use is semantic interpretation or the related conceptual representation.[1]

In the context of modern linguistics, the suggestion that children learn language by hanging language-specific rules and lexical categories on semantic or conceptual representations goes back at least to Chomsky (1965, 29, 59) and Miller (1967). Of course, the idea is much older. See Pinker 1979, 1994 and Fisher et al. 1994 for reviews of some proposed mechanisms, and see Gleitman 1990 for some cogent warnings against the assumption that such semantic representations have their origin solely in the results of direct perception of the material world in any simple sense of that term.

However inadequate our formal (and even informal) grasp on children's prelinguistic conceptualization of their situation, we can be in no doubt that they have one. If so, it is likely that this cognitive representation includes such grammatically relevant prelinguistic notions as actual and potential participants in, properties of, and causal relations among, events; probable attitudes and attentional focus of other conversational participants; and representations of more obviously material aspects of the instantaneously perceivable world.

2.1 Constituents

Three consequences follow from assuming a rule-to-rule relation between syntax and semantics. The first, which follows from the expectation of transparency between syntax and semantics, is so strong and so uncontentious that no theory of grammar has failed to observe it in spirit, though it is probably true to say that none has so far succeeded in following it to the letter. To say that syntax and semantics are related rule-to-rule is to say no more than that every syntactic rule has a semantic interpretation. However, it immediately follows that the syntactic entities that are combined by a syntactic rule must also be semantically interpretable. (Otherwise, they could not be combined by the semantic interpretation of the rule.) It follows that syntactic rules can only combine or yield *constituents*.

This condition, which has been called "The Constituent Condition on Rules," has been a central feature of Generative Grammar from its earliest moments. It frequently surfaces in that literature in the guise of "structure dependency" of grammatical rules. It is also the notion that is embodied in the "proper analysis" condition on transformations proposed in Chomsky 1975a (chapters written in 1955). Perhaps the most illuminating and ambitious early endorsement of this principle is to be found in Chomsky 1975a (210–211, chapters written in 1956), where the following four "criteria" (the scare quotes are Chomsky's) are offered as tests for grammatical constituents and constituent boundaries:

(1) a. The rule for conjunction
    b. Intrusion of parenthetical expressions
    c. Ability to enter into transformations
    d. Certain intonational features

These criteria are very cautiously advanced and carefully surrounded with qualifications, and the subsequent discussion is deliberately designed to demonstrate that some of them raise as many questions as they answer. Nevertheless, there is an implicit claim of great boldness, however programmatically stated. If these operations are tests for constituency, it can only be because they are rules of grammar, subject to the Constituent Condition on Rules. The bulk of Chomsky 1975a, and most work in Generative Grammar since, mainly bears on the claim relative to the third criterion, concerning transformational rules of movement (and their modern equivalents and alternatives), which it has overwhelmingly supported.

It has proved much more difficult to make good on the implicit claim with respect to the remaining three phenomena. Theories of coordination, intonation, and (insofar as there are any) parentheticalization have generally been forced at some point to compromise the Constituent Condition on Rules. The present work should be viewed as an attempt to bring Chomsky's original program nearer completion.

The second consequence of assuming a rule-to-rule relation between syntax and semantics is to imply that the *only* grammatical entities that have interpretations are constituents. This consequence is again entirely uncontentious, and virtually all theories of competence grammar have adhered to it (insofar as they have involved an explicit semantics at all). However, it will be relevant to the discussion of processing in part III.

Finally, the rule-to-rule hypothesis, and its justification in terms of its parsimony with respect to the theory of language learning and evolution, imply that syntactic and semantic rules should have the property of monotonicity. That is, there should be no rules like certain old-style transformations which convert structures that are ill formed (and hence uninterpretable) into structures which are well formed, and vice versa.

To claim that syntax is monotonic is not of course to deny that theories of language need different levels of rules, such as phonology, morphology, syntax, and semantics, or to deny the modular nature of the grammar. However,

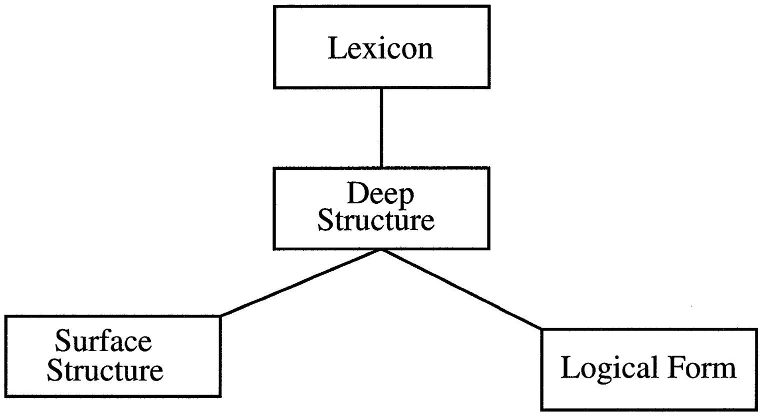


Figure 2.1
Architecture of a standard theory of grammar

It does imply that those levels too should be monotonically related, a point to which I return below.

2.2 Fragments

To what extent do the observed regularities of natural language syntax conform to the expectations set out above? As noted earlier, the generative theoretical tradition has had considerable success in accounting for many constructions involving discontiguities between elements that are semantically dependent upon one another. Many such constructions were originally brought within the fold of the Constituent Condition on Rules by the introduction of transformational rules of "movement" of constituents, relating an underlying level or levels of representation at which predicate-argument relations relevant to semantic interpretation were contiguously related, and from which a Logical Form representing further distinctions such as quantifier scope could be directly computed, to a surface structural level at which the discontiguities were manifest. The introduction of these rules gave rise to the familiar "Y-diagram" architecture of first-generation generative theories, shown in figure 2.1.[2] Theories within the generative tradition differ with respect to how many underlying levels they postulate (one in the case of *Aspects*-style Transformational Grammar "Deep Structures" and Lexical-Functional Grammar (LFG) "f-structure," two in the case of Government-Binding Theory (GB) "S-Structure" and "D-Structure," and even more in some cases). These theories also differ in how they interpret the notion of "movement." Nevertheless, they can all be seen as modifications of this architecture, and the metaphor of movement is so persuasive and revealing that I will use it freely to describe syntactic phenomena, even though the present theory eschews the notion of movement as a syntactic operation.

Movement transformations and the constructions that they capture fall naturally into two groups. The first group includes phenomena that can be accounted for entirely in terms of "bounded" dependencies—roughly, dependencies between items that occur within the same tensed clause, like those between *Dexter* and the verb *win* in the following examples:

(2) a. Dexter expects to win.
 b. Dexter is expected to win by Warren.

As Brame (1976, 1978) and Bresnan (1978) were among the first to point out, the clause-bounded nature of these dependencies means that they can be base-generated or (equivalently) specified in the lexicon, thus bringing them within the domain of the Constituent Condition on Rules without the use of movement as such, and explaining a number of "structure preserving" constraints upon such constructions (Emonds 1976).

This and much subsequent work has shown that the bounded constructions are subject to a number of apparently universal constraints upon such dependencies which reflect dominance and command and an "obliqueness" ordering of arguments of predicates, according to which subjects are ordered with respect to objects and other arguments. An important example for present purposes is afforded by asymmetries in binding possibilities for reflexives and reciprocals like the following, which Keenan (1988) shows to be independent of basic word order across languages:

(3) a. The dogs like each other./*Each other like the dogs.
 b. I showed the dogs to each other/*each other to the dogs.
 c. Sid wants Nancy to like herself/*himself.

I will return to the question of the source of such asymmetries.

The generative approach has also proved extremely successful in accounting for the phenomenon of unbounded dependency exhibited in relative clauses and topicalizations such as the following, again in terms of movement:

(4) a. a book which *I expect I will find*
 b. These articles, *I think that you must have read without understanding.*

In such constructions, elements that are related in the interpretation of the construction, such as the topicalized or relativized NPs and the verb(s) of which they are arguments, can be separated by arbitrarily long substrings and unbounded embedding. Although the residue of topicalization or relativization at first glance looks like a nonconstituent fragment of a sentence, it can be regarded as a constituent of type S, with a special kind of invisible coindexed or "moved" argument, and can thereby be brought under the Constituent Condition on Rules.

Examples like the following suggest that unbounded dependencies are also subject to constraints reflecting conditions of dominance and command involving obliqueness among arguments:

(5) a. *a man whom he thinks that Mary likes
b. *a woman whom I persuaded to like

(The first cannot describe a man who thinks that Mary likes him, and the second cannot describe a woman whom I persuaded to like herself. Again, I will return to this question.)

It has proved much more difficult to account for coordination, parentheticalization, and phrasal intonation within the confines of the Constituent Condition on Rules. It is worth looking at some data in this connection.

2.2.1 Coordination and Parentheticals

At first glance, there is a striking overlap between the kinds of fragments that result from relativization and the related topicalizing construction, and those that can coordinate. In particular, practically everything that can occur as the residue of leftward movement can be coordinated, as in examples like the following:

(6) a. a book which *I expect I will find*, and *I think that you must have read without really understanding*
b. *I expect I will find*, but *I think that you must have read without really understanding*, that novel about the secret life of legumes.

The second, (6b), involves rightward movement (again, the term is used descriptively).

There is a similar (though less complete) conspiracy between the residues of leftward and rightward movement. That is, most residues that arise from leftward movement can also arise from rightward movement. Moreover, both kinds of extraction are subject to very similar "island constraints:"[3]

(7) a. *a book which *I hope that I will meet the woman who wrote*
b. **I hope that I will meet the woman who wrote*, and *you expect to interview the consortium that published*, that novel about the secret life of legumes.

However, the fragments that result from coordination are much more diverse than those that result from (leftward and rightward) movement. For example:

(8) a. I *want to try to write*, and hope to see produced, a musical about the life of Sir Stafford Cripps.
b. Give Deadeye Dick a sugar-stick, and *Mexican Pete, a bun*.
c. I want to try to write a novel, and *you, a screenplay*.

The study of such constructions has revealed that they too are subject to some very strong crosslinguistic constraints, many of which were first discussed by Ross (1970), and which will be discussed at length in chapter 7. These can be summarized as reflecting an "order-preserving" property of coordination, whereby (in configurational languages, at least) if a leftmost constituent is moved, is raised, or otherwise "goes missing" from one conjunct, then it shows up to the left of the entire coordinate structure, whereas a missing rightmost constituent turns up on the right. Thus, in a language like English whose basic word order is Subject-Verb-Object, coordinate sentences like the following are prohibited:[4]

(9) a. *A musical *want to try to write*, and hope to see produced, I.
b. *Deadeye Dick a sugar-stick, and *Mexican Pete, a bun* give.
c. *I a novel, and *you want to try to write a screenplay*

(I will show later in what sense example (9c) is an instance of the same universal.)

Although considerably less attention has been devoted to parenthetical utterances (but see Emonds 1976, II.9, and 1979, McCawley 1982, 1989, Levelt 1989, Espinal 1991, Croft 1995, Taglicht 1998 and Doran 1998), some similarly unconstrained fragments arise from their intrusion, as in (10):

(10) a. *Have you ever been*, they asked, a member of the Friends of the Ukraine Film Society?
b. *You could give*, suggested Dexter, a policeman a flower.

The result has been that, although linguistic theories have had some success in accounting for the relative-clause construction in terms of devices that reinstate the Constituent Condition on Rules by deriving such fragments from traditional constituents such as S via devices like movement (Chomsky 1965), indexed "traces" (Chomsky 1975b), and feature passing (Gazdar 1981), they have been much less successful in showing that the same devices will account for coordination. Instead, coordination has led to the introduction of rules of deletion to supplement rules of movement. Such rules again attempt to reinstate the Constituent Condition over the rule of coordination, by deriving the fragments from underlying constituents of type *S*. However, the deletion rules themselves have generally failed to adhere strictly to that condition. For example, (8b) appears to require either the movement or the deletion of a nonconstituent, and (8c) appears to offer no alternative to the deletion of the nonconstituent *want to try to write*. More worrying still, this fragment looks suspiciously like the kind of fragment that is the surface-structural *result* of deletion or movement, as in (8a).

These complications are surprising, because intuitively all of these constructions appear to be related to the semantic notion of *abstraction*, or definition of a *property*. Most obviously, a restrictive relative clause like (11a) seems to correspond to a predicate or property of *being married by Anna*. Formally, such properties, concepts, or abstractions can be conveniently and transparently represented by terms in the λ-calculus like (11b):

(11) a. (the man that) Anna married
b. $\lambda x.marry'x\ anna'$

Those who are unfamiliar with the λ-calculus are referred to Partee, ter Meulen and Wall 1990, chap. 13, for a full exposition. However, it will be sufficient for present purposes to note that the operator λ declares the symbol x to be a variable local to the expression that follows, distinct from any other variable elsewhere that happens to have the same name. The variable is thus in every way comparable to a parameter or formal variable of a subroutine or function in a computer programming language, which, when instantiated with a value, passes that value to the occurrences of the variable in the expression. The λ-term (11b) can therefore be thought of as defining a function in such a language that maps entities onto truth-values according to whether Anna married them or not. Here as elsewhere in the book, constants like $marry'$, distinguished from variables by primes, are used to identify semantic interpretations whose details are not of immediate interest. Application of a function f to an argument a is simply written fa, and a convention of "left associativity" of function application is observed, so that the above formula is equivalent to the following:

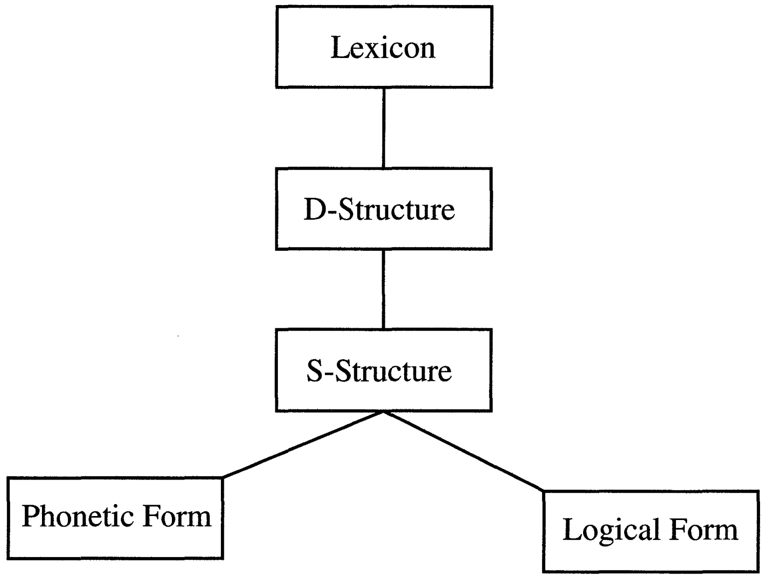


Figure 2.2
Architecture of a government-binding theory of grammar

(12) $\lambda x.(marry'x)anna'$

Most current theories of natural language grammar since "standard" Transformational Grammar (Chomsky 1965), including the present one, more or less explicitly embody the analogy between relativization and abstraction over a variable.

Nevertheless, in the case of coordination and these other unruly phenomena, their apparent freedom from the Constituent Condition on Rules has led to a situation in which, despite the apparently close relation between coordination and relativization, the responsibility for the former phenomenon, together with parentheticalization, and on occasion phenomena like "scrambling" in other languages, has been delegated to a separate, more surface-oriented domain of "stylistic" rules. This led directly to the distinction in GB between the level of S-Structure, at which relativization or *wh*-movement is defined in terms of traces closely resembling syntactic realizations of the bound variables of the λ-calculus, and the level of Phonetic or Phonological Form (PF), at which, contrary to what its name might suggest, significant syntactic work is done. (Chomsky 1986 refers to PF as "Surface Structure.") The result is the theoretical architecture shown in figure 2.2.

2.2.2 Intonation Structure

In a similar apparent contradiction to the Constituent Condition on Rules, some closely related fragments abound in spoken language, arising from phenomena associated with prosody and intonation, as well as less well behaved phenomena like restarts and the parentheticals discussed earlier. For example, one quite normal prosody for an answer to question (13a), involving stress on the word *Anna* and a rise in pitch at the end of the word *married*, imposes an Intonation Structure which is orthogonal to the traditional syntactic structure of the sentence, as informally indicated in (13b) by the parentheses (stress, marked in this case by pitch maxima, is indicated by capitals):

(13) a. I know that Alice married ALAN. But who did ANNA marry?
b. (ANNA married)(MANNY).

We will of course need more empirical evidence and more formal notations to define this phenomenon more precisely in the chapters that follow. But the effect is very strong. (It is ironic that one of the difficulties in teaching introductory syntax is to persuade students that this is *not* the notion of structure that is relevant to the study of syntax. One conclusion that can be drawn from the argument in this book is that the students are, in an important sense, quite right.)

As is often the case with informally controlled contexts like this, other intonation contours are possible. In particular, because the context (13a) is compatible with an interpretation under which *Anna* is the topic or theme of the utterance, a contour with an intonational boundary separating the subject and predicate, *(ANNA)(married MANNY)*, is also possible. (For further discussion, see Jackendoff 1972; Steedman 1991a; and chapter 5 for further discussion.)

Intonation Structure nevertheless remains strongly constrained by meaning. For example, contours imposing bracketings like the following do not seem to be allowed, as Selkirk (1984) has pointed out:

(14) #(Three CATS)(in ten prefer CORDUROY).

Halliday (1967a) observes that this constraint, which Selkirk calls the "Sense Unit Condition," seems to follow from the *function* of phrasal intonation, which in English is in part to convey what Halliday called "Information Structure"—that is, distinctions of focus, presupposition, and propositional attitude toward entities in the discourse model. These discourse entities are more diverse than mere NP or propositional referents, but they do not seem to include such nonconcepts as "in ten prefer corduroy."

The question in (13), *Who did Anna marry?*, appears to introduce a new "theme" or topic of conversation, corresponding like the relative clause in (11a) to the concept of *someone such that Anna married them.* As Jackendoff (1972) points out, it is once again natural to think of this theme as a functional *abstraction* and to express it using exactly the same expression of the λ-calculus as was used in (11b) for the relative clause, repeated here:[5]

(15) $\lambda x.marry'x\ anna'$

When this function or concept is supplied with an argument *manny′*, it *reduces* to give a proposition, with the same function-argument relations as the canonical sentence (Again, function application associates to the left.)

(16) $marry'manny'anna'$

It is the presence of the abstraction (15) rather than some other that makes the intonation contour in (13b) felicitous. (That is not to say that its presence uniquely *determines* this response, or that its explicit mention is necessary for interpreting the response.)

These observations have led linguists such as Selkirk to postulate a level of Intonation Structure, independent of syntactic structure and related to an Information Structural component of LF, implying an architecture something like the one in figure 2.3 for the theory of grammar as a whole (see Selkirk 1984, 205, and cf. Chomsky 1971).[6]

The involvement of two apparently uncoupled levels of underlying structure on the way from sound to meaning in natural language grammar appears to complicate the path from speech to interpretation unreasonably, and to thereby threaten the entire theory of grammar (not to mention its worrying implications for the feasibility of a number of applications in computational speech recognition and speech synthesis).

In the light of the increasing complexity of the mainstream theories of grammar in the face of these less well behaved constructions, it is interesting to observe that the coordinate constructions considered in section 2.2.1, whose semantics also seems to be reminiscent of functional abstraction, are also subject to something like the Sense Unit Condition that limits intonational phrases. For example, strings like *in ten prefer corduroy* seem to be as reluctant to take part in coordination as they are to be treated as intonational phrases:

(17) *Three cats in twenty like velvet, and in ten prefer corduroy.

Parentheticalization is similarly bad at such junctures:

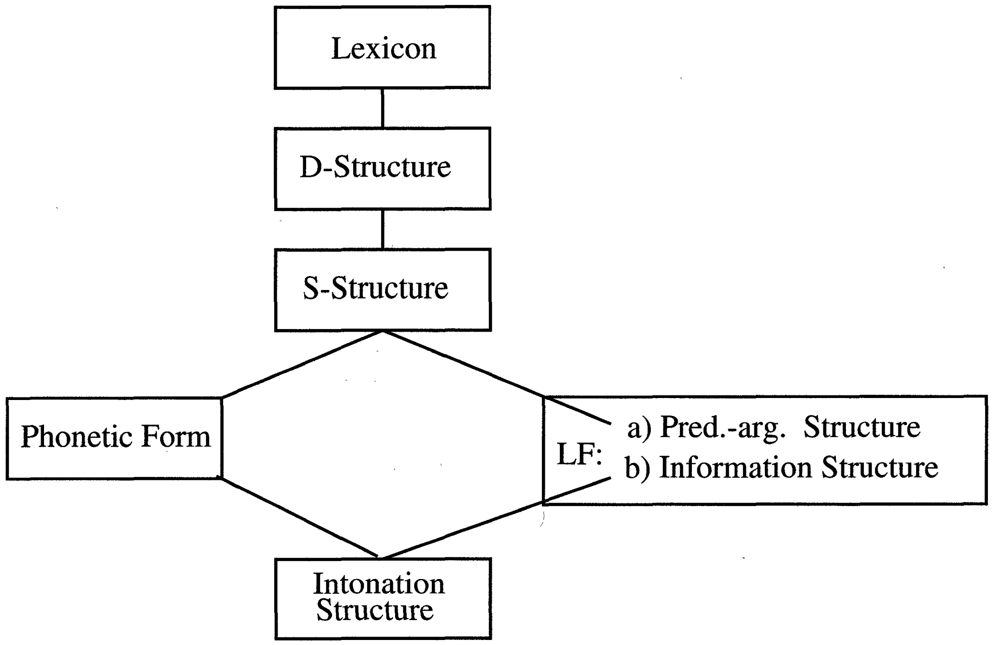


Figure 2.3
Architecture of a government-binding prosody

(18) *"Three cats," ejaculated Tom, prematurely, "in ten prefer corduroy."

Since coordinate constructions and parentheticals constitute major sources of complexity for current theories of natural language grammar and also offer serious obstacles to computational applications, it is tempting to suspect that this conspiracy between syntax and prosody might point to a unified notion of syntactic constituent structure that is somewhat different from the traditional one, but that unifies the GB notions of S-Structure, PF and Intonation Structure under a single notion that I shall call Information Structure, to avoid confusion with other notions of Surface Structure that are in play.

2.3 Issues of Power and Explanation

To cut ourselves adrift from traditional linguistic notions of constituency is a frightening prospect. Fortunately, there are other principles of linguistic investigation to guide the search for an alternative.

The first among these principles is Occam's razor, which says that we should capture the phenomena of language in theories with as few degrees of freedom as possible. The strength of a theory lies in the extent to which (a) it captures all and only the phenomena that can occur, and (b) could not in principle capture phenomena that we have good reason to believe could not occur, and therefore has no need to explicitly exclude them by stipulation.

It follows that a theory should involve as few levels of representation as possible, consistent with the main goal of capturing generalizations. Ideally, following the Montagovian tradition, we would like to assume no more than an interpretation related in a strictly monotonic fashion to syntactic derivation.

We should also try to minimize power in the modules of the theory, consistent with the primary goal of building interpretations that capture predicate-argument relations correctly. If we can do so with grammars that are provably low on the Chomsky hierarchy of formal languages and automata-theoretic power, then we are on stronger theoretical ground than if we adopted theories that achieve the same coverage at the expense of greater power, because greater automata-theoretic power increases the variety of alternative constructions and phenomena that we *could* capture beyond the point where there seems to be any motivation from empirical or imaginable semantic dependencies. Here I follow the tradition of Generalized Phrase Structure Grammar (GPSG; Gazdar 1982) and mildly context-sensitive theories of grammar such as Tree-Adjoining Grammar (TAG; Joshi, Levy and Takahashi 1975).

Of course, to achieve parsimony in automata-theoretic power is possible only to the extent that we have good information about the real space of possible natural languages and grammars. Fortunately there is a certain amount of information available on this question.

In asking what is the least powerful class of grammars on some scale such as the Chomsky hierarchy that includes all natural grammars, we must distinguish between "strong" and "weak" adequacy of grammars or sets of rules for capturing languages in the formal sense of sets of strings of words or other terminal vocabulary symbols. A grammar or set of rules that merely generatively specifies all and only the strings of a language may be only weakly adequate as a grammar for the language. To be strongly adequate, it must also assign a correct structural description to the string, where "correct" structural descriptions are the ones that support the semantics of the language. For example, a language whose semantics demands a context-free grammar whose rules happen to permit embedding only on rightmost elements has a weakly adequate finite-state grammar that generates all and only the same strings. However, the weakly adequate finite-state grammar does not directly express the embedding that supports the semantics. It is only the strongly adequate context-free grammar that does that.

However strong our intuitions may be concerning some aspects of meaning in natural language, we do not have access to natural semantics in the direct sense in which we have access to its strings. It follows that the only formal proofs that we can construct concerning lower bounds on the power implicit in natural languages involve weak adequacy. Of course, if we can show that some level of the automata-theoretic power hierarchy is not even weakly adequate to capture all natural languages, then it immediately follows that the level in question is not strongly adequate either. However, finding a lower bound on the power of strongly adequate grammars via a proof of weak adequacy depends on finding a construction in the language that not only has an intuitive semantics that demands that power to produce correct structural descriptions, but that also happens not to admit a weakly adequate grammar of lower power.

Partly because of the widespread presence of lexical ambiguity in natural languages, the search for a formal proof of a lower bound on weak adequacy was extremely protracted. Chomsky (1957) gave an early argument that nothing lower than context-free grammars—that is, grammars that can be written with production rules expanding a single nonterminal or phrasal type—could be weakly adequate. However, many of the earliest examples of constructions that were claimed to prove that the lower bound on power was strictly greater than context-free were flawed, either because they confounded extragrammatical factors with grammatical ones (like the argument from the English *respectively* construction) or because they depended on intuitive assumptions about strong adequacy and failed to exclude the possibility of a weakly equivalent grammar of lower power. These arguments are helpfully reviewed by Pullum and Gazdar (1982) and by Pullum (1987).

This curious delay in proving any lower bound on power greater than context-free should be kept in perspective. Since the late 1970s, there has been very little doubt that (a) the competence grammars implicated by the semantics or predicate-argument relations for natural languages were strictly greater than context-free in power, and (b) that power was not very much greater than context-free.

One reason for believing natural grammars not to greatly exceed context-free power, and in particular not to come anywhere near the power of context-sensitive grammars, is that most phenomena of natural languages, including those involving unbounded dependencies, can be captured by context-free rules. Although I will not at this point go into the way in which certain kinds of unbounded dependencies can be captured using context-free or near-context-free rules, the possibility is strongly suggested by the observation that the two dependencies in (19a) must nest, rather than intercalate, as they would have to for (19b) to have a meaning to do with playing sonatas on violins (the asterisk here means "not allowed with the intended reading.")

(19) a. a violin which$_i$ [this sonata]$_j$ is hard to play$_j$ upon$_i$

b. *a sonata which$_i$ [this violin]$_j$ is hard to play$_i$ upon$_j$

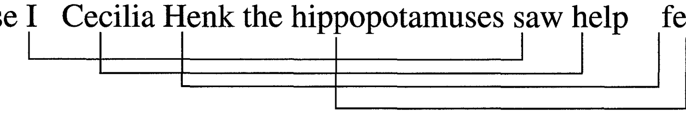
To nest dependencies is of course the characteristic behavior of a pushdown automaton, which is the automaton for recognizing context-free languages, so it seems likely that there is a strongly adequate context-free grammar for this construction. GPSG (Gazdar 1981, 1982; Gazdar et al. 1985) extensively explored the extent to which all grammatical phenomena can be captured in context-free terms.

The reason for believing that natural grammars are of strictly greater than context-free power lies in the fact that, although nonnesting or intercalating dependencies are rare, strong adequacy will undoubtedly require them to be captured in some, and perhaps all, natural languages. The most convincing observation of this kind came from Dutch, in work by Huybregts (1976), although it was some time before a formal proof in the form of a proof of weak inadequacy of context-free grammars was forthcoming on the basis of a similar phenomenon in related Germanic dialects (Huybregts 1984; Shieber 1985).

In Dutch there is a strong tendency for all verbs in subordinate clauses to be clause-final and for matrix verbs to appear to the left of the verbs in their complements. This means that certain sentences with embedded infinitival complements that in English embed right-peripherally involve crossing or intercalating dependencies between predicates and arguments, as indicated by connecting lines in (20):[7]

(20) ... omdat ik Cecilia Henk de nijlpaarden zag helpen voeren.

... because I Cecilia Henk the hippopotamuses saw help feed



'... because I saw Cecilia help Henk feed the hippopotamuses.'

The important property of this construction is that the semantic dependencies between nouns and verbs do not nest in Dutch, as they do in English. They intercalate, or interleave. This means that any strongly adequate grammar for Dutch—that is, one that maintains a rule-to-rule relation between syntax and semantics and captures the dependencies correctly—is likely to be non-context-free (Wall 1972).

The challenge that is offered by the contrast between nesting examples like

(19) and intercalating examples like (20) is to find a strongly adequate class of grammars that is "mildly" context-sensitive, allowing these examples and the previous ones that are characteristic of coordination and prosody, without allowing all sorts of illicit phenomena.

There is a third source of information that we can draw upon in our search for such a class. There are a number of known crosslinguistic constraints on grammar that are so strong as to apparently constitute universal limitations on natural grammars. We will be concerned with many such universal constraints below, but two in particular will play a central role in the argument.

The first is due to Ross (1970), who pointed out that the construction that shows up in English as (medial) gapping, in sentences like (21), shows a strong crosslinguistic regularity concerning which of the two conjuncts undergoes "deletion" or otherwise has the verb go missing:

(21) Dexter likes cats, and Warren, dogs.

The pattern is that in languages whose basic clause constituent order subject-verb-object (SVO), the verb or verb group that goes missing is the one in the right conjunct, and not the one in the left conjunct. The same asymmetry holds for VSO languages like Irish. However, SOV languages like Japanese show the opposite asymmetry: the missing verb is in the *left* conjunct.[8] The pattern can be summarized as follows for the three dominant constituent orders (asterisks indicate the excluded cases):[9]

(22)	SVO:	*SO and SVO	SVO and SO
	VSO:	*SO and VSO	VSO and SO
	SOV:	SO and SOV	*SOV and SO

This observation can be generalized to individual constructions within a language: just about any construction in which an element apparently goes missing preserves canonical word order in an analogous fashion. For example, English ditransitive verbs subcategorize for two complements on their right, like VSO verbs. In the following "argument cluster" coordination, it is indeed in the right conjunct that the verb goes missing:

(23) Give a policeman a flower, and a dog a bone.

The second phenomenon identified above, the crosslinguistic dependency of binding of reflexives and anaphors upon Jackendoff's Jackendoff (1972) obliqueness hierarchy is discussed by Keenan (1988) and Clark (1985, 1991), among others. Regardless of basic word order (here there are data from OS languages and constructions within languages), or indeed of configurationality itself, anaphoric pronouns like *themselves* and *each other* may corefer with another argument of the verb just in case that other argument is less oblique—that is, earlier in a series that places subject first, object next, and more oblique arguments later. In English this shows up in the fact that a subject may bind an object, but not vice versa:

(24) a. Dexter and Warren like each other.
b. *Each other like Dexter and Warren.

In the case of the ditransitive verbs, it shows up in the fact that the first object can bind the second, but not vice versa (see Barss and Lasnik 1986 for discussion):

(25) a. I introduced Dexter and Warren to each other.
b. *I introduced each other to Dexter and Warren.

This is not an idiosyncratic fact about English or SVO languages. Keenan shows that in VSO and even VOS and mixed VSO/VOS languages, less oblique arguments such as subjects bind more oblique arguments such as objects, and not vice versa.

2.4 Grammar as an Applicative System

The two universals singled out in the last section, both of which can conveniently be illustrated using the same English ditransitive construction, induce opposite tensions in the theory of grammar. For reasons that will be developed below, the dependency of gapping upon canonical word order suggests that directionality under concatenation is directly projected from the lexicon to the string by strictly concatenative rules. On the other hand, binding suggests the existence of a level of representation at which obliqueness is represented independently of surface order, or that projection is not strictly concatenative, or both.

It is interesting in this connection to note that there are alternative systems to the λ-calculus for capturing the notion of abstraction, and that these systems entirely avoid the use of bound variables. They are the combinatory systems invented by Schönfinkel (1924) and Curry and Feys (1958) as a formal foundation for the semantics of the λ-calculus. They are entertainingly discussed in Smullyan 1985, where the combinatory operators take the form of birds, and from which a number of the epigraphs to the present chapters are taken. In such systems, which I discuss in detail in chapter 8, terms equivalent to abstractions are built up using a few simple operations for combining functions, such as functional composition. Systems using quite small numbers of combinators can be shown to be equivalent in expressive power to the λ-calculi. The existence of these systems raises the possibility that alternative theories of grammar can be developed based as directly upon the combinatory applicative systems as the traditional ones implicitly are upon the λ-calculus. The significance of this possibility is that the different form that syntactic rules take in combinatory systems may lead us to look at the kinds of phenomena discussed above in a new way.

Because combinators are operations that map *functions* onto other functions, and because the categorial grammars that were originally developed in their "pure" context-free form by Ajdukiewicz (1935) and Bar-Hillel (1953) provide a notation in which functional type is made explicit, this insight has led in recent years to an explosion of research in frameworks that are collectively known as "flexible" categorial grammars (see, e.g., Lambek 1958; Montague 1970; Geach 1972; Cresswell 1973; Karlgren 1974; Bach 1976, 1979, 1980; Shaumyan 1977; Keenan and Faltz 1978; von von Stechow 1979; Levin 1982; Ades and Steedman 1982; Dowty 1982; Hausser 1984; van Benthem 1986; Flynn 1983; Huck 1985; Zwarts 1986; Uszkoreit 1986; Wittenburg 1986; Desclés, Guentchéva and Shaumyan 1986; Oehrle 1987, 1988; Zeevat, Klein and Calder 1987; Bouma 1987; Szabolcsi 1989; Moortgat 1988a; Hoeksema 1989; Carpenter 1989; Hepple 1990; Jacobson 1990; Segond 1990; Karttunen 1989; Hepple 1990; Jowsey 1989; Steele 1990; Reape 1996; Wood 1993; van der Linden 1993; Potts 1994; Houtman 1994; Ranta 1994; Morrill 1994; Hendriks 1995; and Aarts 1995.[10] One of the interesting properties of combinatory applicative systems is that in general they offer many equivalent combinatory expressions for a given normalized λ-term. This property is reflected in another distinguishing feature of certain of the above theories. The use of rules related to combinators encourages a property of "associativity" in linguistic derivations. That is, for any unambiguous sentence, there are typically several distinct categorial derivations, all of which are semantically equivalent in the sense that they deliver the same function-argument relations. The notion "constituent of a derivation" is correspondingly generalized to cover many of the puzzling fragments discussed above.

I will not attempt to review this very diverse literature here.[11] However, many of the ideas explored below draw upon these works, and I will try to make their parentage clear as we go. This book is most closely related to the subgroup of these theories developed by Ades, Oehrle, Jacobson, Szabolcsi, and Hepple, among other cited above, although they should not be assumed to endorse the present theory in all (or even many) respects. In these theories, syntactic rules corresponding to simple individual combinators such as functional composition are used to lend such fragments as *want to try to write* and even *a policeman a flower* the status of constituents, without the use of movement or deletion. Such grammars will be shown to provide a unified treatment of a wide variety of syntactic phenomena in natural language and to explain phenomena of long-distance dependency (including relativization), coordination, parentheticalization, and intonation, within the confines of the Constituent Condition on Rules and in terms of a single principle. That principle is that the predicate-argument relations that hold in sentences of natural languages are projected onto long-range syntactic dependencies from the relations defined locally in the lexicon by syntactic operations corresponding to combinators, rather than by syntactic operations involving empty categories or traces corresponding to syntactically realized bound variables.

In order to demonstrate that these novel grammars deliver the correct interpretations, we will need a semantic notation. Although we could use combinators for the purpose, λ-calculus is far more readable and in every way equivalent. The appearance of variables in the semantic notation might give the impression that traces have been reintroduced. However, these variables are no more than a readable notation for a uniform mechanism whereby *all* arguments, whether "extracted" or "in situ," get semantically bound to predicates.

BLANK PAGE

Chapter 3

Intuitive Basis of Combinatory Categorial Grammars

Given any birds A, B, and C, the bird C is said to compose *A with B if for every bird x the following condition holds:* $Cx = A(Bx)$
Raymond Smullyan, *To Mock a Mockingbird*

Because combinators are operations on functions, we will need a notation for grammars that makes prominent the functional type or "category" of linguistic entities—that is, a notation that specifies the kinds of things that a linguistic entity combines with and the kind of thing that results. Categorial Grammar (CG), invented by Ajdukiewicz, Bar-Hillel, and others, which in its pure form is (weakly) equivalent to other context-free grammars, provides such a notation, and it is there that we will begin.[1]

3.1 Pure Categorial Grammar

Categorial grammars put into the lexicon most of the information that is standardly captured in context-free phrase structure (PS) rules. For example, consider the following PS rules, which capture some basic syntactic facts concerning English transitive sentences:

(1) $S \rightarrow NP\ \ VP$
$VP \rightarrow TV\ \ NP$
$TV \rightarrow \{married, finds,\ \ldots\}$

In a categorial grammar, all constituents—and in particular the lexical elements, such as verbs and nouns—are associated with a syntactic "category" that identifies them as either *functions* or *arguments*. In the case of functions, the category specifies the type and directionality of their argument(s) and the type of their result. The present work uses a notation in which the argument or domain category always appears to the right of the slash, and the result or range category to the left. A forward slash / means that an argument of the appropriate type must appear to the right of the functor; a backward slash \ means that the argument must appear to the left.[2] All functions are "Curried," so that a function of n arguments becomes a unary function onto a Curried function of $n-1$ arguments. The category of a simple transitive tensed verb is therefore written as follows, capturing the same facts about English transitive sentences as the PS rules in (1):

(2) married := $(S\backslash NP)/NP$

Curried functions, which are so-named after Haskell Curry, are equivalent to the corresponding "flat" functions of n arguments (Schönfinkel 1924), and it will be convenient to refer to S in the above function as its range, and the two NPs as its domain or arguments.

The class of all possible categories is recursively defined as an infinite set of terms including the basic categories S, NP, and so on, and all terms of the form α/β and $\alpha\backslash\beta$, where α and β are categories.

The lexicon of a given language is a finite subset of the set of all categories subject to quite narrow restriction that ultimately stem from limitations on the variety of semantic types with which the syntactic categories are paired in the lexicon. In particular, we can assume that lexical function categories are limited to finite—in fact, very small—numbers of arguments. (For English at least, the maximum appears to be four, required for a small number of verbs like *bet*, as in *I bet you five dollars he's a good dog.*)

The most basic assumption of the present approach is that the responsibility for specifying all dependencies, whether long-range or local, resides in the lexical specifications of syntactic categories for the "heads" of those dependencies—that is, the words corresponding to predicate-argument structural functors, such as verbs. This principle, which is related to the Projection Principle of GB, can be more formally stated as follows:[3]

(3) *The Principle of Lexical Head Government*
Both bounded and unbounded syntactic dependencies are specified by the lexical syntactic type of their head.

This is simply to say that the present theory of grammar is "lexicalized," a property that makes it akin to LFG, TAG, Head-Driven Phrase Structure Grammar (HPSG; Pollard and Sag 1994), and certain recent versions of GB (see Hale and Keyser 1993; Brody 1995).

Lexicalized grammars make the lexical entries for words do most of the grammatical work of mapping the strings of the language to their interpretations. The size of the lexicon involved is therefore an important measure of a grammar's complexity. Other things being equal, one lexical grammar is simpler than another if it captures the same pairing of strings and interpretations using a smaller lexicon.

An important property of CCG, which it shares with LFG and GB, and which sets it apart from TAG, GPSG, and HPSG (which in other respects are more closely related), is that it attempts to minimize the size of the lexicon by adhering as closely as possible to the following stronger principle:

(4) *The Principle of Head Categorial Uniqueness*
A single nondisjunctive lexical category for the head of a given construction specifies both the bounded dependencies that arise when its complements are in canonical position and the unbounded dependencies that arise when those complements are displaced under relativization, coordination, and the like.

That is not to say that a given word may not be the head of more than one construction and hence be associated with more than one category. Nor does it exclude the possibility that a given word-sense pair may permit more than one canonical order, and hence have more than one category per sense. For example, in chapter 6 Dutch and German word order is captured by assuming that verbs in these languages systematically have two categories, one determining main-clause order and the other subordinate clause orders. Baldridge (1999) suggests that languages with freer word order such as Tagalog may have even more categories for verbs. The claim is simply that each of these categories specifies both canonical order and all varieties of extraction for the clause type in question. For example, a single lexical syntactic category (2) for the word *married*, which does not distinguish between "antecedent," "θ," or any other variety of government, is involved in all of the dependencies illustrated in (5):

(5) a. Anna *married Manny*.
b. the man *that* I believe that Anna *married*
c. I believe that Anna *married* and you believe that she dislikes, the man in the grey flannel suit.

In both TAG and GPSG these dependencies are mediated by different initial trees or categories, and in HPSG they are mediated by a disjunctive category.

We will on occasion be forced to allow exceptions to the Principle of Head Categorial Uniqueness. However, each such exception complicates the grammar and makes it compare less favorably with an otherwise equivalently valued grammar that requires no such exceptions.

Such functions can combine with arguments of the appropriate type and position by rules of functional application, written as follows:

(6) *The functional application rules*

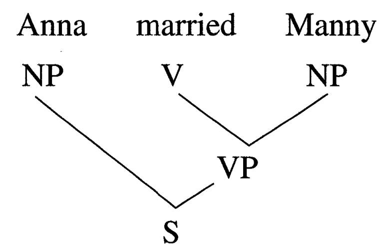
a. $X/Y \quad Y \quad \Rightarrow \quad X$ $(>)$

b. $Y \quad X\backslash Y \quad \Rightarrow \quad X$ $(<)$

These rules have the form of very general binary PS rule schemata in which X, Y, and the like, are variables ranging over categories. If the grammar is limited to "pure" categorial grammar involving these schemata alone, then it is nothing more than a context-free grammar that happens to be written in the accepting, rather than the producing, direction, and in which the major burden of specifying particular grammars has been transferred from the PS rules to the lexicon. Although it is now convenient to write derivations as in (7a), below, they are clearly equivalent to the familiar trees, (7b):

(7) a.

Anna	married	Manny
NP	$(S\backslash NP)/NP$	NP
	$S\backslash NP$ (>)	
S (<)		

b. 

(Underlines in the categorial derivation indicate combination via the application rules, and the left/right arrows mnemonically indicate which rule has applied.)[4]

The basic categories like S and NP can, and in fact must, be regarded as complex objects that include both major syntactic features, of the kind used in $\bar{\text{X}}$-bar theory, and minor syntactic features like number, gender, and person agreement. Such syntactic feature bundles will for present purposes be abbreviated as S, NP_{3s}, $S\backslash NP_{3s}$, and so on, since the particular feature set is not at issue here, and the precise implementation of minor features or feature bundles like agreement is of no immediate relevance. Thus, we will from time to time want to write the category of *married* as follows:

(8) married $:= (S\backslash NP_{3s})/NP$

(9)

Anna	married	Manny
NP_{3sf}	$(S\backslash NP_{3s})/NP$	NP
	$S\backslash NP_{3s}$ (>)	
S (<)		

The derivation in (9) illustrates the way in which the syntactic features are used to capture agreement in categorial grammars, following Bach (1983).[5]

In such a framework, *3s*, *3sf*, and the like, can be regarded as abbreviations for particular, possibly underspecified values or sets of values for an attribute or set of attributes *AGR*. Variables such as *agr*, agr_1, and the like, can range over such values. Arguments that are unspecified for agreement (such as the object of *married*, (8)) by convention have such a variable as their value on this feature bundle or attribute, allowing them to "unify" with any more specified value in the sense of the term current in programming languages like Prolog. (See Sag et al. 1986 for the basic approach, and Bayer and Johnson 1995 for discussion of some complexities that we will also pass over here. Unification as a basis for combinatory rules is discussed in Steedman 1996b)

3.2 Interpretation and Predicate-Argument Structure

Although for many purposes we will continue to be able to ignore the details of semantics in derivations, it will from time to time be important to remember that categories also include semantic interpretations. Interpretations can be thought of either as purely model-theoretic objects or as predicate-argument structures (i.e., as Logical Forms in the logician's sense of the term). For present purposes, it will be helpful to think of them as predicate-argument structures.

One convenient way to capture the interpretation of a verb like (8) is to associate a λ-term with the syntactic category, via a colon operator, as follows:

(10) married := $(S\backslash NP_{3s})/NP : \lambda x.\lambda y.marry'xy$

Constants in such interpretations are again distinguished from variables by primes. The λ-operators define the way arguments are bound into lexical Logical Forms. This particular Logical Form could be further reduced by η reduction to the term *marry'*, and I will freely use such reduced forms where they save space.[6]

The syntactic and semantic components of such categories are related by the following principle, which is the categorial equivalent of a principle of "type driven translation" in rules (Klein and Sag 1985), and related to the $\bar{X}$ theory (Chomsky 1970; Jackendoff 1977).

(11) *The Principle of Categorial Type Transparency*
For a given language, the semantic type of the interpretation together with a number of language-specific directional parameter settings uniquely determines the syntactic category of a category.

For example, a verbal function from subjects to propositions in English may be of type $S\backslash NP$ whereas a function from objects to predicates must be S/NP.

The present use of the untyped λ-calculus as a simplified notation for Logical Forms does not help us to straightforwardly define this principle formally. Even the typed version would require some extensions to distinguish nominal and verbal functions of type $e \rightarrow t$, and to distinguish subject arguments from objects. Providing those extensions would in turn require considerable argumentation, in view of the problematic status of notions like subject and object across languages. Since this problem is essentially the same within all grammatical frameworks, and the concerns of this book lie elsewhere, a formal definition will not be attempted here. However, the Principle has a number of corollaries which limit the degrees of freedom that are available within the CCG framework.

One such corollary of the Principle of Categorial Type Transparency is its inverse: the semantic type of an interpretation is entirely determined by the syntactic type on the left of the colon, under a function $\mathcal{T}$ (see Montague 1974, who defines a similar mapping):

(12) *The Inverse of the Principle of Categorial Type Transparency*
For any category $\Sigma : \Lambda$, Λ is of type $\mathcal{T}\Sigma$

$\mathcal{T}$ is recursively defined as follows:

(13) If Σ is a basic syntactic type NP, S, N, etc. then $\mathcal{T}\Sigma$ is a corresponding fixed semantic type such as e (entity), t (truth-value), $e \rightarrow t$ (property), etc. If Σ is a syntactic functor category α/β or $\alpha\backslash\beta$ then $\mathcal{T}\Sigma$ is a corresponding semantic functor of type $\mathcal{T}\beta \rightarrow \mathcal{T}\alpha$.

For example, the semantic type of a transitive verb like *married* (10) has to be of type $e \rightarrow (e \rightarrow t)$ under this definition—a function from entities to functions-from-entities-to-truth-values. We are not free to assume that it is t, or $e \rightarrow t$ just because we feel like it. A fuliy explicit Logical Form for the English SVO transitive verb in (10) would further specify that the first argument is the object of the proposition and the second the subject, where those roles are defined by a universal predicate-argument structure in which the subject commands the object. This particular Logical Form could therefore be simplified by "η-normalization" as $marry'$. However, the above definition of $\mathcal{T}$ would also permit an OVS language in which this category had the translation $\lambda y.\lambda x.marry'xy$, in which the subject combines first, followed by the object. Such a Logical Form cannot be simplified, and makes an essential, though entirely local, use of λ-abstraction. In Steedman 1996b, I argue that this degree of freedom in the lexical specification of Logical Forms is required to account for binding phenomena in English, as well as in VSO and OSV languages. In later sections we will see that it is also implicated by coordination phenomena in verb-initial constructions in English, and in other SVO and VSO languages.[7]

When interpretations are made explicit in this particular notation, the application rules must be similarly expanded to associate predicate-argument structures with (variables over) syntactic categories, as follows:

(14) *Functional application*

a. $X/Y : f \quad Y : a \quad \Rightarrow \quad X : f\,a$ $\quad (>)$

b. $Y : a \quad X\backslash Y : f \quad \Rightarrow \quad X : f\,a$ $\quad (<)$

It is important to notice that these two rules apply an identical compositional-semantic operation—functional application—in both syntax and semantics. This particular notation might make it appear that the theory allows the freedom to involve an arbitrary semantic operation associating some interpretation other than fa with the result X of the rule—say, faa or af. However, the Principle of Categorial Type Transparency means that such interpretations are incompatible with the semantic types that are involved. And in fact, the syntactic rule is simply the translation of a semantic rule of functional application. In general, all combinatory rules that are permitted in CCG are subject to the following principle, whose definition corrects an earlier version in Steedman 1996b:

(15) *The Principle of Combinatory Type Transparency*

All syntactic combinatory rules are type-transparent versions of one of a small number of simple semantic operations over functions.

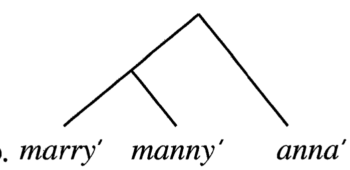
We have yet to see what is meant by "simple" and "small number" here, and we have already noted in the introduction that linguists have more secure access to the syntactic end of rules than to the semantic one. However, we may note that functional operations do not come much simpler than functional application, and that if the semantic types corresponding to X, Y, and X/Y, respectively, are x, y, and $y \rightarrow x$, then the term fa is the *only* normalized λ-term of type x that can be formed from f and a. This is typical of all the combinatory rules we will consider.

The earlier derivation appears as follows in this notation:

$$
\text{(16)}\quad
\dfrac{\dfrac{\text{Anna}}{NP_{3sf} : anna'} \quad \dfrac{\dfrac{\text{married}}{(S\backslash NP_{3s})/NP : \lambda x.\lambda y.marry'xy} \quad \dfrac{\text{Manny}}{NP : manny'}}{S\backslash NP_{3s} : \lambda y.marry'manny'y}{>}}{S : marry'manny'anna'}{<}
$$

(The rules in (14) include higher-order variables like f ranging over functions like $\lambda x.\lambda y.marry'xy$.)

The example illustrates the way a derivation synchronously builds a predicate-argument structure for the sentence, via the interpretations of constituents. Interpretations obey a convention under which the application of a function (like *marry'*) to an argument (like *manny*) is represented by concatenation (as in *marry'manny'*), where such expressions "associate to the left." This means that the interpretation *marry'manny'anna'* of the S result in derivation (16) is equivalent to expression (17a), the brackets being suppressed by this convention:

(17) a. $(marry'manny')anna'$ b. 

Predicate-argument structures like (17a) are therefore equivalent to binary trees, like (17b), which preserve traditional notions of dominance and command, but do not preserve the linear word order of the string. Word order is defined solely by the directional slashes in the syntactic category. CCG categories are therefore reminiscent of the lexical items of Zubizarreta (1987), Grimshaw (1990), or Williams (1994), and (less directly) those in HPSG (see Pollard and Sag 1987, 1994), or a restricted version of elementary trees in Lexicalized Tree-Adjoining Grammars (LTAG; see Joshi and Schabes 1992). CCG differs only in the way in which these predicate-argument relations are "projected" by the combinatory rules of syntax.

It is important to realize that the λ-notation is simply a convenient device for binding arguments into predicate-argument structures during a derivation like (10). The potentially powerful mechanism of λ-abstraction does no real work outside the domain of an individual lexical item, and its potential for application over unboundedly large structures is never invoked. It could therefore be replaced with less powerful finite-state devices or (at the cost of certain complications that we will come to) be eliminated entirely, as in the related Montagovian categorial tradition of Bach, Dowty and others.

3.3 Coordination

In earlier papers on the present theory, coordination was introduced via the following schema, which goes back at least as far as Chomsky 1957, 36, (26), and can be paraphrased as "Conjoin like categories":[8]

(18) *Simplified coordination rule* ($<\Phi>$)
$X \quad CONJ \quad X' \Rightarrow \quad X''$

X, X', and X'' are categories of the same type but different interpretations. Using such a rule schema, transitive verbs (for example) can coordinate as follows:

(19)
```
Anna   met           and    married        Manny
----   -----------   ----   -----------    -----
NP     (S\NP)/NP     CONJ   (S\NP)/NP      NP
       --------------------------------<Φ>
              (S\NP)/NP
              ---------------------------------->
                            S\NP
----------------------------------------<
                  S
```

Such a rule is an oversimplification, because it runs counter Ross 's (1967, 90; 1986, 99) observation that in English (as opposed to other languages—see Schachter 1985, 47), conjunctions are "prepositional"—that is, they associate structurally with the right conjunct, not the left.[9] I will return to the question later, continuing to use the simplified coordination rule (18) in derivations.

The rule (18) is a schema over types, and semantically its instances must be distinguished, to differentiate coordination of nonfunctions, unary functions, binary functions, and so on. We can represent this by the following schema (which we will later need to restrict further):

(20) *Coordination* ($<\Phi^n>$)
$X : g \quad CONJ : b \quad X : f \quad \Rightarrow_{\Phi^n} \quad X : \lambda \ldots b(f \ldots)(g \ldots)$

Apart from the simplification already mentioned, this is everyone's coordination rule (see Gazdar 1981). It captures the ancient intuition that *coordination is an operation that maps two constituents of like type onto a constituent of the same type*. Because X may be a functor category of any valency or number of arguments, the rule is formulated as a schema over such types.[10]

Given such a rule or rule schema, derivations like the following are permitted, and yield semantically correct results (agreement is omitted, and functions like $\lambda x.\lambda y.marry'xy$ are abbreviated as $marry'$):

(21)
$$
\begin{array}{ccccc}
\text{Anna} & \text{met} & \text{and} & \text{married} & \text{Manny} \\
\hline
NP & (S\backslash NP)/NP & CONJ & (S\backslash NP)/NP & NP \\
: anna' & : meet' & : and' & : marry' & : manny'
\end{array}
$$
$$
\frac{(S\backslash NP)/NP : meet' \quad CONJ : and' \quad (S\backslash NP)/NP : marry'}{(S\backslash NP)/NP : \lambda x.\lambda y.and'(meet'xy)(marry'xy)} <\Phi>
$$
$$
\frac{(S\backslash NP)/NP : \lambda x.\lambda y.and'(meet'xy)(marry'xy) \quad NP : manny'}{S\backslash NP : \lambda y.and'(meet'manny'y)(marry'manny'y)} >
$$
$$
\frac{NP : anna' \quad S\backslash NP : \lambda y.and'(meet'manny'y)(marry'manny'y)}{S : and'(meet'manny'anna')(marry'manny'anna')} <
$$

The interpretation may give the impression that we have introduced rules of copying or deletion into the grammar. Any Logical Form of this kind must of course express the fact that the arguments appear in two predications. However, this is not the same as introducing *surface syntactic* operations of this kind. By operating in syntax entirely on types, we automatically ensure the equivalent of an "across-the-board" or "parallelism" condition on such deletions, excluding examples like the following without stipulation (see Williams 1978; Goodall 1987; Moltmann 1992):

(22) *Anna [met Manny]$_{S\backslash NP}$ and [married.]$_{(S\backslash NP)/NP}$

3.4 The Bluebird

In order to allow coordination of contiguous strings that do not constitute traditional constituents, CCG generalizes pure Categorial Grammar, to include certain further operations on functions related to Curry's combinators (Curry and Feys 1958). For example, functions may *compose*, as well as apply, under the following rule:

(23) *Forward composition* ($>\mathbf{B}$)
$$X/Y \quad Y/Z \quad \Rightarrow_{\mathbf{B}} \quad X/Z$$

The most important single property of combinatory rules like this is that they have an invariant type-driven semantics, as required by the Principles of Categorial and Combinatory Type Transparency, (11) and (15). The semantics of this rule is almost as simple as functional application. It is in fact functional *composition*. The combinator that composes two functions f and g is called $\mathbf{B}$ by Curry, and is the Bluebird in Smullyan's (1985) combinatory fable.[11] It can be defined by the following equivalence:

(24) $\mathbf{B}fgx \equiv f(gx)$

A convention that application associates to the left is again followed, so that the left-hand side is equivalent to $((\mathbf{B}f)g)x$. It follows that we can consider the application of $\mathbf{B}$ to f and g as producing a new function equivalent to abstracting on x in the above expression, thus:[12]

(25) $\mathbf{B}fg \equiv \lambda x f(gx)$

The new rule (23) is semantically a typed version of this combinator. Hence, the arrow in the rule is subscripted $\Rightarrow_{\mathbf{B}}$, and the application of the rule in derivations is indexed $>\mathbf{B}$.

Using this rule, sentences like *Anna met, and might marry, Manny* can be accepted syntactically as follows:

(26) Anna / met / and / might / marry / Manny

$NP \quad (S\backslash NP)/NP \quad CONJ \quad (S\backslash NP)/VP \quad VP/NP \quad NP$

might marry $\xrightarrow{>\mathbf{B}}$ $(S\backslash NP)/NP$

met and might marry $\xrightarrow{<\Phi>}$ $(S\backslash NP)/NP$

met and might marry Manny $\xrightarrow{>}$ $S\backslash NP$

Anna met and might marry Manny $\xrightarrow{<}$ S

In semantic terms the rule can be written in full as follows:[13]

(27) *Forward composition* ($>\mathbf{B}$)
$X/Y : f \quad Y/Z : g \quad \Rightarrow_{\mathbf{B}} \quad X/Z : \lambda x.f(gx)$

Derivation (26) then appears as follows:[14]

(28) Anna / met / and / might / marry / Manny

$NP : anna' \quad (S\backslash NP)/NP : meet' \quad CONJ : and' \quad (S\backslash NP)/VP : might' \quad VP/NP : marry' \quad NP : manny'$

might marry $\xrightarrow{>\mathbf{B}}$ $(S\backslash NP)/NP : \lambda x.\lambda y.might'(marry'x)y$

met and might marry $\xrightarrow{<\Phi>}$ $(S\backslash NP)/NP : \lambda x.\lambda y.and'(might'(marry'x)y)(meet'xy)$

met and might marry Manny $\xrightarrow{>}$ $S\backslash NP : \lambda y.and'(might'(marry'manny')y)(meet'manny'y)$

Anna met and might marry Manny $\xrightarrow{<}$ $S : and'(might'(marry'manny')anna')(meet'manny'anna')$

The formalism immediately guarantees without further stipulation that this operation will compose the interpretations, as well as the syntactic functional types.[15]

The result of the composition has the same syntactic type $(S\backslash NP)/NP$ as a transitive verb, and readers may easily satisfy themselves that its translation is such that, if applied to an object and a subject, it is guaranteed to yield exactly the same predicate-argument structure for the sentence *Anna might marry Manny* as would have been obtained without the introduction of this rule.

Of course, the grammar continues correctly to exclude examples like the following, because only *adjacent* like categories can coordinate:

(29) *Anna [met Manny]$_{S\backslash NP}$, and [might marry]$_{(S\backslash NP)/NP}$.

A generalization of composition is required for sentences like the following:[16]

(30)
$$\frac{\text{I}}{NP}\quad \frac{\text{offered,}}{((S\backslash NP)/PP)/NP}\quad \frac{\text{and}}{CONJ}\quad \cfrac{\frac{\text{may}}{(S\backslash NP)/VP}\quad \frac{\text{give,}}{(VP/PP)/NP}}{((S\backslash NP)/PP)/NP}{>}\mathbf{B}^2\quad \frac{\text{a flower}}{NP}\quad \frac{\text{to a policeman}}{PP}$$

The generalization simply allows composition into certain functions of more than one argument. The requisite composition rule is stated as a schema over functions of varying numbers of arguments, as follows:

(31) *Generalized forward composition* ($>\mathbf{B}^n$)
$$X/Y : f \quad (Y/Z)/\$_1 : ...\lambda z.gz... \quad \Rightarrow_{\mathbf{B}^n} \quad (X/Z)/\$_1 : ...\lambda z.f(gz...)$$

The rule uses a notation introduced in a more general form in Ades and Steedman 1982 to schematize over verbs with different numbers of arguments, which I will call the "\$ convention." It can be defined recursively as follows:

(32) *The \$ convention*
For a category α, $\{\alpha\$\}$, (respectively, $\{\alpha/\$\}$, $\{\alpha\backslash\$\}$) denotes the set containing α and all functions (respectively, leftward functions, rightward functions) into a category in $\{\alpha\$\}$ (respectively, $\{\alpha/\$\}$, $\{\alpha\backslash\$\}$).

I will use unbracketed $\alpha\$$, $\alpha/\$$, and $\alpha\backslash\$$ to schematize over the members of the sets $\{\alpha/\$\}$, $\{\alpha/\$\}$, and $\{\alpha\backslash\$\}$ respectively, using subscripts as necessary to distinguish distinct schematizations. For example, $\{S/\$\}$ is the set $\{S, S/NP, (S/NP)/NP, \ldots\}$ and $S/\$$, $S/\$_1$, and so on, are schemata over that set. The use of $\$_1$ in rule (31) indicates that these are occurrences of the *same* schema.

$(Y/Z)/\$$ is thereby defined as a schema over functions yielding Y combining with n arguments to the right, the last or "innermost" of which is of type Z, where we can assume for English that $n \leq 3$. In essence this makes the rule equivalent to a finite set of rules specified for all the verbal categories of the English lexicon.[17] The semantics of each instance depends on the value of n and is one of the series of combinators called $\mathbf{B}$, $\mathbf{B}^2, \mathbf{B}^3$. It is represented by Curry's own schematization of these composition combinators as $\mathbf{B}^n$, as the annotation on the arrow indicates.[18]

3.5 The Thrush

Combinatory grammars also include type-raising rules, which turn arguments into functions over functions-over-such-arguments. Since these rules allow arguments to become functions, they may by that token compose with *other* functions, and thereby take part in coordinations like *Anna married, and I detest, Manny*. Like composition, the type-raising rules have a simple and invariant semantics, as required by the Principle of Combinatory Type Transparency. The semantics corresponds to another of Curry's basic combinators, which he called $\mathbf{C}_*$ but which I will here call $\mathbf{T}$ for type-raising, following Rosser (1935) and to Smullyan (1985), in whose book it appears in the guise of the Thrush.[19] It is defined by the following equivalence:

(33) $\mathbf{T}xf \equiv fx$

It follows that $\mathbf{T}$ applied to an argument creates the following abstraction over the function:

(34) $\mathbf{T}x \equiv \lambda f\, fx$

For example, the following syntactic rule, indexed $>\mathbf{T}$, is needed for coordinate sentences like *Anna married, and I detest, Manny*:

(35) *Subject type-raising* ($>\mathbf{T}$)
$NP \quad \Rightarrow_{\mathbf{T}} \quad S/(S\backslash NP)$

Derivations like the following are therefore allowed, and deliver appropriate interpretations:

(36)
$$
\begin{array}{cccccc}
\text{Anna} & \text{married} & \text{and} & \text{I} & \text{detest} & \text{Manny} \\
\hline
NP & (S\backslash NP)/NP & CONJ & NP & (S\backslash NP)/NP & NP \\
\hline{}^{>\mathbf{T}} & & & \hline{}^{>\mathbf{T}} & & \\
S/(S\backslash NP) & & & S/(S\backslash NP) & & \\
\multicolumn{2}{c}{\overline{\quad\quad\quad\quad\quad\quad}{}^{>\mathbf{B}}} & & \multicolumn{2}{c}{\overline{\quad\quad\quad\quad\quad\quad}{}^{>\mathbf{B}}} & \\
\multicolumn{2}{c}{S/NP} & & \multicolumn{2}{c}{S/NP} & \\
\multicolumn{5}{c}{\overline{\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad}{}^{<\Phi>}} & \\
\multicolumn{5}{c}{S/NP} & NP \\
\multicolumn{6}{c}{\overline{\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad}{}^{>}} \\
\multicolumn{6}{c}{S}
\end{array}
$$

Of course, the following example is excluded, because, once again, only *adjacent* categories can coordinate:

(37) *[Anna married]$_{S/NP}$ [Manny]$_{NP}$ and [I detest.]$_{S/NP}$

This example illustrates an important general property of CCGs: even in coordinate constructions, directionality and word order are projected consistently from the lexicon by the combinatory rules.

The combinatory rules of which (35) is an example can be captured in the following two rule-schemata, complete with their interpretation:

(38) *Type-raising*

a. $X : a \Rightarrow_{\mathbf{T}} \mathrm{T}/(\mathrm{T}\backslash X) : \lambda f.f\,a$ $(>\mathbf{T})$

where $\mathrm{T}\backslash X$ is a parametrically licensed category for the language

b. $X : a \Rightarrow_{\mathbf{T}} \mathrm{T}\backslash(\mathrm{T}/X) : \lambda f.f\,a$ $(<\mathbf{T})$

where T/X is a parametrically licensed category for the language

T is a variable over categories, ranging over the result types of functions over X. (It is distinguished by roman typeface, because strictly speaking it is a variable of the metalanguage. Each type-raised category has its own unique variable of this kind, and when the corresponding variables need to be distinguished at the object level we will distinguish them as T_i, T_j, and so on. Otherwise, the unadorned metavariable T is used whenever possible, to reduce notational clutter.)

The restriction limits $\mathrm{T}\backslash X$ and T/X to types that are permitted under the (informally defined) Principle of Categorial Type Transparency. Among other things, I will assume it prevents infinite recursion of type-raising (because the parametric specification of legal categories must presumably make reference to a fixed set of basic types). The restriction, which is discussed further in chapters 4 (10) and 7 (64), also means that, for example, in English, as opposed to German, $\mathrm{T}\backslash X$ cannot be instantiated as $(S\backslash NP)\backslash NP$. At least in English, and possibly in all languages, we can assume that this restriction limits $\mathrm{T}\backslash X$ to a finite set of categories,

The rules as stated also only turn an argument category such as *NP* into either a rightward-looking functor over leftward-looking functors over *NP*, or a leftward-looking functor over rightward-looking functors over *NP*. They are therefore "order-preserving" with respect to the linear order of function and argument as defined in the lexicon. This restriction is discussed further in chapter 8.

Derivation (36) then appears as follows:

(39)
$$\begin{array}{cccccc}
\text{Anna} & \text{married} & \text{and} & \text{I} & \text{detest} & \text{Manny} \\
\overline{\mathrm{T}/(\mathrm{T}\backslash NP_{3sf})} & \overline{(S\backslash NP_{3s})/NP} & \overline{CONJ} & \overline{\mathrm{T}/(\mathrm{T}\backslash NP_{3s})} & \overline{(S\backslash NP_{3s})/NP} & \overline{\mathrm{T}\backslash(\mathrm{T}/NP_{3p})} \\
:\lambda f.f\ anna & :\lambda x.\lambda y.marry'xy & :and' & :\lambda f.f\ i' & :\lambda x.\lambda y.detest'xy & :\lambda f.f\ manny' \\
\multicolumn{2}{c}{\overline{S/NP : \lambda x.marry'x\ anna'}^{\;>\mathbf{B}}} & & \multicolumn{2}{c}{\overline{S/NP : \lambda x.detest'x\ i'}^{\;>\mathbf{B}}} & \\
\multicolumn{5}{c}{\overline{S/NP : \lambda x.and'(detest'x\ i')(marry'x\ anna')}^{\;<\Phi>}} & \\
\multicolumn{6}{c}{\overline{S : and'(detest'manny'i')(marry'manny'anna')}^{\;<}}
\end{array}$$

The composition of substrings like *Anna married* yields semantically interpreted functions that, if reduced with an object *Manny*, yield the same result that we would have obtained from the traditional derivation shown in (16), namely, $S : marry'manny'anna'$, thus guaranteeing that the coordination will deliver the correct predicate-argument structure, preserving traditional relations of dominance and c-command. This is an important observation, since as far as Surface Structure goes, we have now compromised both those traditional relations.

Type-raising was first used by Lewis and Montague as a semantic device to capture the type of generalized quantifiers. However, the present syntactic use is distinct, and the intuition behind cases like rule (35) is more reminiscent of the linguists' ancient notion of "nominative case."[20] In a language like Latin, nominative case determines an NP argument like *Balbus* to be something that must combine with a predicate, like *ambulat* or *murum ædificat*, to yield a proposition such as $walk'balbus'$ or $build'wall'balbus'$. In categorial terms, nominative case turns *Balbus* into a function whose interpretation is precisely $\mathbf{T}balbus'$—that is, $\lambda p.p\ balbus'$, a function over functions-over-subjects, or predicates. Similarly, accusative case turns NPs into functions over a different type of functions, functions over objects, with a semantics that is again defined in terms of $\mathbf{T}$. The restriction of the general form (38) of type-raising to categories that are arguments of verbs is in fact a natural consequence of its relation to a notion of case generalized to non-nominal arguments.

Thus, the only cause for surprise at this ingredient of CCG is that English behaves like a cased language without in general marking even nominal case morphologically.

Evidence that type-raising is as generally applicable to arguments as case is in other languages is available from the following observation, which originates with Dowty (1988) and a related analysis of Dutch coordination in Steedman 1985.[21] On the assumption that all English NPs can freely type-raise, together with that of a further backward rule of function composition corresponding to the mirror image of rule (23), the existence of the following construction in English is immediately predicted. (The variable T in the raised categories is here instantiated by the actual categories involved. The symbol *DTV* abbreviates $(VP/NP)/NP$, the category of a ditransitive verb, and *TV* abbreviates VP/NP, that of a transitive verb.)

(40)

$$\begin{array}{cccccc}
\text{give} & \text{a teacher} & \text{an apple} & \text{and} & \text{a policeman} & \text{a flower} \\
\overline{DTV} & \overline{TV\backslash DTV}^{<\mathbf{T}} & \overline{VP\backslash TV}^{<\mathbf{T}} & \overline{CONJ} & \overline{TV\backslash DTV}^{<\mathbf{T}} & \overline{VP\backslash TV}^{<\mathbf{T}} \\
 & \multicolumn{2}{c}{\overline{VP\backslash DTV}^{<\mathbf{B}}} & & \multicolumn{2}{c}{\overline{VP\backslash DTV}^{<\mathbf{B}}} \\
 & \multicolumn{5}{c}{\overline{VP\backslash DTV}^{<\Phi>}} \\
\multicolumn{6}{c}{\overline{VP}^{<}}
\end{array}$$

This construction is often referred to as "left-peripheral deletion" or "nonconstituent coordination." However, neither name sits well with the present theory, according to which there is no such thing as deletion, and no such thing as coordination of nonconstituents. I have on occasion followed Schachter and Mordechai (1983), in referring to such sentences as "left node raised," but I will here refer to these constructions as "argument cluster coordination."

The important fact to note about this derivation is that the type-raised categories of the indirect and direct objects are simply those that are allowed by the order-preserving backward type-raising rule (38b), given the category of the English verb. The *only* rule of composition that will permit these categories and adjuncts to combine is the following mirror image of the earlier composition rule (23):

(41) *Backward composition* ($<\mathbf{B}$)
$$Y\backslash Z \quad X\backslash Y \quad \Rightarrow_{\mathbf{B}} \quad X\backslash Z$$

The restriction of type-raising to raising over categorially type-transparent function types means that compositions like the following are excluded in English:

(42)
$$\frac{S/(S\backslash NP) \quad \mathrm{T}/(\mathrm{T}\backslash NP)}{S/((S\backslash NP)\backslash NP)}{>}\mathbf{B}$$

The combination would require the second type-raised category to be instantiated as as function $(S\backslash NP)\backslash NP$, which is not a parametrically licensed category of English.

The possibility of argument cluster coordinations like (40) is therefore predicted by exactly the same ingredients of the theory that were introduced to explain right node raising—namely, order-preserving type-raising and composition. The existence of argument cluster coordination in SVO languages, together with the related dependency of so-called forward and backward gapping on VSO and SOV word order discussed in chapter 7, is in fact one of the strongest pieces of confirmatory evidence in favor of the grammars that admit these two combinatory operations. It is with respect to these constructions that the theory should be contrasted with other closely related function-oriented and unification-based theories, such as those advanced by Joshi, Levy and Takahashi (1975), Bresnan (1982), Karttunen (1989), Uszkoreit (1986), Wittenburg (1986), Zeevat, Klein and Calder (1987), Pickering and Barry (1993), and Pollard and Sag (1994).

The earlier analogy between type-raising and case also suggests an option in the theory. Should we consider type-raising to be an operation of active syntax, like composition, or to be a rule of the lexicon or of morphology? In the latter case, of course, not only nominative NPs like *I*, but also uninflected NPs like *Anna*, and even articles like *the* would have to bear a number of additional categories such as $S/(S\backslash NP)$ and $(S/(S\backslash NP))/N$, and so on, like *Balbus* and *ille* in Latin.

I will return to this point in chapter 4. But one (case-marked) category related to type-raising that certainly must be lexically assigned is the category of the relative pronoun, for the addition of type-raising and composition to the theory of grammar provides everything needed in order to account for leftward extractions in relative clauses. I will assume as in Steedman (1996b) and other earlier work that relative pronouns bear the categories in (43) (Many North American dialects do not allow extraction of "internal" arguments, as in *a policeman whom*$_i$ *I gave* t_i *a flower*, and *a man whom I bet* t_i *\$10 that Ipswich would win the Cup*, and therefore only require the first two of these.)[22]

(43) a. *who/that/which* := $(N\backslash N)/(S\backslash NP)$
b. *who(m)/that/which* := $(N\backslash N)/(S/NP)$
c. *who(m)/that/which* := $((N\backslash N)/X)/((S/X)/NP)$
d. *who(m)/that/which* := $(((N\backslash N)/X)/Y)/(((S/X)/Y)/NP)$

The second of these is a function from fragments like *Anna married* to noun modifiers. The following relative clause is therefore accepted:

(44)

(the man)	that	Anna	married
	$(N\backslash N)/(S/NP)$	$S/(S\backslash NP)$ (>T)	$(S\backslash NP)/NP$
		S/NP (>B)	
	$N\backslash N$ (>)		

It should be obvious that the theory immediately predicts that leftward and rightward extraction will be unbounded, since embedded subjects can have the raised category, and composition can apply several times, as in the following cases:

(45) a. Anna married, but I doubt whether she can afford, Manny.
b. the man who Anna married, but I doubt whether she can afford

In offering a common origin for phenomena of coordinate structure and relativization, the present theory has some affinity with GPSG (Gazdar 1981; cf. Gazdar et al. 1985). It differs from GPSG and its descendants in doing so entirely on the basis of projection from the lexical subcategorization implicit in the category, as is required by the Principle of Lexical Head Government, rather than via a SLASH mechanism distinct from subcategorization.

The advantage of this tactic is that the generalization (31) of composition to functions with more than one argument allows multiple extractions combining rightward and leftward extraction (see (30)), which Maling and Zaenen (1978) noted present problems for GPSG:[23]

(46) a. the policeman to whom I offered, and may give, a flower
b. the policeman to whom I offered, and you may give, a flower

As in GPSG, leftward and rightward extraction are forced without further stipulation to obey the "across-the-board" condition of Williams (1978) on extraction out of coordinates, including the "same case" exceptions, because the grammar does not yield categories of like type for the conjuncts in examples like the following:

(47) a. *(A man who) [Anna married]$_{S/NP}$ but [I dislike him]$_{S}$
b. *(A man who) [Anna married]$_{S/NP}$ and [irritates me]$_{S\backslash NP}$

We will see in chapter 4 that many more of the notorious constraints on extraction that have been identified by the transformationalists follow from the combinatory theory in a similarly elegant way, without stipulation. But it is appropriate to briefly consider island constraints here, since they constitute a notorious limitation on the unboundedness of rightward and leftward movement that we have just captured.

The fact that adjuncts are in general islands might seem to be a natural consequence of the assumption that they are backward modifiers, as can be seen from the categories in the following unacceptable example:

(48) * a book [which]$_{(N\backslash N)/(S/NP)}$ [I will]$_{S/VP}$ [walk]$_{VP}$ [without reading]$_{(VP\backslash VP)/NP}$

However, this leaves unexplained the fact that they are equally strong islands in languages that have preverbal modifiers, such as Dutch and German. Even in English, it follows only on the assumption that verbs like *walk* cannot in general be type-raised over adjuncts, to become $VP/(VP\backslash VP)$. Since at least some VPs are arguments, the definition (38) allows them to acquire this category and compose into the adjunct, allowing the extraction.

The possibility of exceptions to the island status of NPs and adjuncts, and their equally notorious dependence on lexical content and such semantically related properties as definiteness and quantification, can be explained on the assumption that the results of composition are subject to a performance-related pragmatic condition requiring their interpretation to be a "natural kind" giving rise to useful inference in the knowledge base. Concepts like *painting a picture of something* and even *dying without finishing something* can reasonably be supposed to be natural kinds in this sense, but the concept of *walking without reading something* presumably is not, except possibly among the more morbid variety of linguist.[24]

(49) a. the man who Mary painted a picture of
b. a symphony which he died without finishing
c. #a book which he walked without reading

Since type-raised categories are a little unreadable, it will often be convenient to abbreviate them as $NP^{\uparrow}$, $PP^{\uparrow}$, and so on, where the context makes it obvious which specific instance is involved.

3.6 The Starling

The theory requires one further type of combinatory rule. Example (50), which is of a kind first noticed by Ross (1967), and discussed by Engdahl (1983) and Taraldsen (1979) under the name of the "parasitic gap" construction, is of interest both because it involves the extracted item in more than one dependency and because one of those dependencies is upon an item inside an island which is not normally accessible to extraction, as shown in (51).

(50) articles which_i I will file_i without reading_i

(51) a. articles which_i I will $\textit{file}_i$ before reading your instructions
b. #articles which_i I will read your instructions $\textit{before filing}_i$

Parasitic gaps are therefore unlike the multiple dependencies that are permitted "across the board" in the coordinate structures considered earlier, where neither extraction is possible alone:

(52) a. *(articles) which I will file them and you will *forget*
b. *(articles) which I will *file* and you will forget them

They are extensively discussed in categorial terms in Steedman 1996b and can be briefly summarized for present purposes as follows.

The lexical categories for (50) are as follows:

(53) (articles) which I will file without reading

$$\begin{array}{lllll} \textit{which} & \textit{I will} & \textit{file} & \textit{without} & \textit{reading} \\ \overline{(N\backslash N)/(S/NP)} & \overline{S/VP} & \overline{VP/NP} & \overline{(VP\backslash VP)/VPing} & \overline{VPing/NP} \\ & & & \multicolumn{2}{c}{\overline{\hspace{8em}}{>}\mathbf{B}} \\ & & & \multicolumn{2}{c}{(VP\backslash VP)/NP} \end{array}$$

The combinatory rules introduced so far allow us to compose *without* and *reading*, but there the analysis blocks. Composition will not help, nor will the coordination rule (since the categories of *file* and *without reading* are not the same). The introduction of some further operation or operations appears to be inevitable.

The intuition that sequences like *file without reading* constitute a semantically coherent entity of some kind in this construction is very strong. The fact that such sequences can occur in isolation in instructions like *shake before opening* and that they can coordinate with transitive verbs in phrases like *file without reading and forget* suggests that they are predicates of some kind—more specifically, that they bear the category of a transitive verb, VP/NP.[25] Szabolcsi (1983, 1989) proposed a combinatory rule to combine the VP/NP and the $(VP\backslash VP)/NP$ to yield this VP/NP. The rule was a special case of the following one:

(54) *Backward crossed substitution* ($<\mathbf{S}_\times$)
$Y/Z \quad (X\backslash Y)/Z \quad \Rightarrow_\mathbf{S} \quad X/Z$

This rule, which unlike composition and type-raising is not a theorem of the Lambek calculus, is the *only* further rule type that will be needed. Since it provides a strong clue to the nature of the entire space of rules from which we are choosing instances, it is worth examining at some length.

The rule (54) (whose index $<\mathbf{S}_\times$ will be explained later) allows derivations like the following:

(55) (articles) which I will file without reading

$(N\backslash N)/(S/NP)$ S/VP VP/NP $(VP\backslash VP)/VPing$ $VPing/NP$

$\overline{\quad (VP\backslash VP)/NP \quad} >\mathbf{B}$

$\overline{\quad VP/NP \quad} <\mathbf{S}_{\times}$

$\overline{\quad S/NP \quad} >\mathbf{B}$

$\overline{\quad N\backslash N \quad} >$

As usual, the parallel rightward extraction is correctly predicted to be possible:[26]

(56) Mary will $[\text{copy}]_{VP/NP}$, and $[\text{file without reading}]_{VP/NP}$, any article longer than ten thousand words.

As usual, the rule has a simple and invariant semantics. It is a close relative of functional composition and corresponds to a third very basic combinator in Curry's system, called **S**. It is called here "functional substitution" and is the Starling in Smullyan's (1985) fable. It is defined by the following equivalence:

(57) $\mathbf{S}fgx \equiv fx(gx)$

It follows that the application of the combinator to two functions is equivalent to the following abstraction:

(58) $\mathbf{S}fg \equiv \lambda x\, fx(gx)$

Again we must assume that the variable Y in the substitution rule (54) is restricted to predicate categories like VP, in a way I will spell out in detail later.[27] Note that the rule will not allow arbitrary double dependencies, preventing the following from meaning "a man such that I showed him himself" without further stipulation:

(59) *(a man) who(m) I showed

$(N\backslash N)/(S/NP)$ $(S/NP)/NP$

$\overline{\quad\quad} {*}{<}\mathbf{S}_{\times}$

We will see in chapter 4 that extraction from the first site alone, as in (51a), is still allowed, and extraction from the second site alone, as in example (51b), is still excluded.

Rule (54) is therefore written in full as follows:

(60) *Backward crossed substitution* $(<\mathbf{S}_{\times})$

$Y/Z : g \quad (X\backslash Y)/Z : f \quad \Rightarrow_{\mathbf{S}} \quad X/Z : \lambda x.fx(gx)$

where $Y = S\backslash\$$

The rule permits the following derivation for sentences with the structure of (56):[28]

(61) Mary will copy and file without reading these articles

Mary will	copy	and	file	without	reading	these articles
S/VP $:\lambda p.will'\ p\ mary'$	VP/NP $:copy'$	$CONJ$ $:and'$	VP/NP $:file'$	$(VP\backslash VP)/VPing$ $:\lambda p.\lambda q.without'pq$	$VPing/NP$ $:read'$	$NP^{\uparrow}$ $:articles'$

$$\text{without reading} \Rightarrow_{>\mathbf{B}} (VP\backslash VP)/NP : \lambda x.\lambda q.without'(read'x)q$$

$$\text{file without reading} \Rightarrow_{<\mathbf{S}_{\times}} VP/NP : \lambda x.without'(read'x)(file'x)$$

$$\text{copy and file without reading} \Rightarrow_{<\Phi>} VP/NP : \lambda x.and'(without'(read'x)(file'x))(copy'x)$$

$$\text{copy and file without reading these articles} \Rightarrow_{<} VP : and'(without'(read'articles')(file'articles'))(copy'articles')$$

$$\text{Mary will copy and file without reading these articles} \Rightarrow_{>} S : will'(and'(without'(read'articles')(file'articles'))(copy'articles'))mary'$$

The restriction on rule (60) uses the \$ convention (32) to permit only categories of the form (tensed, untensed, participial, etc.) S, $S\backslash NP$, and so on, to unify with the variable X. It excludes the analogous derivation for the following phrase:

(62) *a [picture of]$_{N/NP}$ [by]$_{(N\backslash N)/NP}$ [Rembrandt]$_{NP}$

Further cases of parasitic gapping are discussed in Steedman 1996b.

Chapter 4

Explaining Constraints on Natural Grammar

"A Starling," said Bravura, "is a bird **S** *satisfying the following condition:* **S***xyz* = *xz*(*yz*)."
"Why is that bird so important?" asked Craig.
"You will find that out when you reach the Master Forest," replied Bravura.
Raymond Smullyan, *To Mock a Mockingbird*

One might ask at this point what degrees of freedom are implicit in the choice of the rules proposed in chapter 3 in order to account for the facts of English, and from what space of possible alternatives we have been selecting the rules that happen to suit us. For in choosing those rules for English, we necessarily commit ourselves to the claim that other possible human languages might exercise the same degrees of freedom in other ways. If descriptive generalizations give reason to believe that human languages do not in fact vary in the predicted ways, then we have some further explaining to do.[1]

4.1 Intrinsic Constraints Limiting the Set of Possible Rules

It is interesting in this regard to examine the rule of functional substitution introduced in section 3.6, for it happens to conspicuously exploit one degree of freedom that we might not necessarily have expected to need, but that will be claimed here to be widespread in natural grammars. It equally conspicuously *fails* to exploit a number of further degrees of freedom that do not appear to be needed in natural grammars, and that, if exploited, would weaken the theory considerably. Here is the rule again:

(1) *Backward crossed substitution* ($<\mathbf{S}_{\times}$)
$Y/Z \quad (X\backslash Y)/Z \quad \Rightarrow_{\mathbf{S}} \quad X/Z$
where $X = S\backslash\$$

It will be useful in contemplating such rules to define the term "principal function" to refer to that function among the inputs to the rule which determines the range of the result—which in present notation is always X. The first thing to notice is that rule (1) combines a principal function that is looking *leftward* for an argument of type Y with a *rightward*-looking function into that category Y.

The effect of allowing such "slash-crossing" rules in the theory is likely to be far-reaching, because if they are allowed for substitution rules, then slash-crossing versions of composition rules are predicted as well. Since such rules are not theorems of the Lambek calculus, which is weakly context-free (Pentus 1993), it is likely that they will induce greater expressive power than context-free grammar. Nevertheless, derivation (55) in chapter 3 suggests rule (1) must be included, for as Szabolcsi (1983) points out, there does not seem to be any question about the choice of categories for the verb group and the adverbial modifier.

The second thing to notice about rule (1) is that it appears to conform in every other respect to the directionality that is implicit in the categories that it combines. The principal function over Y in the rule does indeed combine with something *to its left*. And the directionality of the Z argument in its result is the same as the directionality of the Z arguments in its inputs. In fact, *all* of the combinatory rules exemplified above conform to the directionality of their inputs in these respects, and we can characterize them all by the following three principles:[2]

(2) *The Principle of Adjacency:*
Combinatory rules may only apply to finitely many phonologically realized and string-adjacent entities.

(3) *The Principle of Consistency:*
All syntactic combinatory rules must be consistent with the directionality of the principal function.

(4) *The Principle of Inheritance:*
If the category that results from the application of a combinatory rule is a function category, then the slash defining directionality for a given argument in that category will be the same as the one(s) defining directionality for the corresponding argument(s) in the input function(s).

The first of these principles simply embodies the assumption that some set of combinatory rules will do the job. That is, it says that rules can only apply to finitely many contiguous elements.[3]

I have suggested in earlier papers that these principles are universal, and that they delimit the space of possible combinatory rules in all human languages. The Principle of Consistency excludes the following kind of rule:

(5) $X\backslash Y \quad Y \quad \not\Rightarrow \quad X$

The Principle of Inheritance excludes rules like the following instance of composition:

(6) $X/Y \quad Y/Z \quad \not\Rightarrow \quad X\backslash Z$

It also prohibits analogues of the coordination rule (18) of chapter 3 such as the following:

(7) $X/Y \quad CONJ \quad X\backslash Y \quad \not\Rightarrow \quad X/Y$

Together the principles amount to a simple statement that *combinatory rules may not contradict the directionality specified in the lexicon.* In Steedman 1987, 1991c, I argued that this in turn reflects the fact that directionality is a property of *arguments*—in other words, that these principles are corollaries of the Principles of Categorial and Combinatory Type Transparency, whose close relation to the Projection Principle of government-binding theory was noted in chapter 3.[4] The argument is somewhat technical and is deferred until chapter 8.

The principles permit the following instances of the two syntactic combinatory rule types, in which the \$ generalization under the convention (32) of the last chapter can apply to both sets of rules, replacing Y/Z and $Y\backslash Z$ by $(Y/Z)/\$$, $(Y\backslash Z)\$$, etc. It is again assumed that such schemata are limited to a bounded number of arguments:[5]

(8) *Functional composition*

a. $X/Y \quad Y/Z \quad \Rightarrow_{\mathbf{B}} \quad X/Z$ $\qquad (>\mathbf{B})$
b. $X/Y \quad Y\backslash Z \quad \Rightarrow_{\mathbf{B}} \quad X\backslash Z$ $\qquad (>\mathbf{B}_\times)$
c. $Y\backslash Z \quad X\backslash Y \quad \Rightarrow_{\mathbf{B}} \quad X\backslash Z$ $\qquad (<\mathbf{B})$
d. $Y/Z \quad X\backslash Y \quad \Rightarrow_{\mathbf{B}} \quad X/Z$ $\qquad (<\mathbf{B}_\times)$

(9) *Functional substitution*

a. $(X/Y)/Z \quad Y/Z \quad \Rightarrow_{\mathbf{S}} \quad X/Z$ $\qquad (>\mathbf{S})$
b. $(X/Y)\backslash Z \quad Y\backslash Z \quad \Rightarrow_{\mathbf{S}} \quad X\backslash Z$ $\qquad (>\mathbf{S}_\times)$
c. $Y\backslash Z \quad (X\backslash Y)\backslash Z \quad \Rightarrow_{\mathbf{S}} \quad X\backslash Z$ $\qquad (<\mathbf{S})$
d. $Y/Z \quad (X\backslash Y)/Z \quad \Rightarrow_{\mathbf{S}} \quad X/Z$ $\qquad (<\mathbf{S}_\times)$

Any language is free to restrict these rules to certain categories, or to entirely exclude a given rule type. But the above is the entire catalogue of types.

Some of these rules—namely, $>\mathbf{B}_\times$ and $<\mathbf{B}_\times$, as well as all four rules related to the combinator **S**—are *not* theorems of the Lambek calculus. Their inclusion represents a point of divergence between the present theory and those derived from the Lambek calculus (see van Benthem 1986, chap. 7; Moortgat

1988a, 1997; Hepple 1990; Morrill 1994. The significance of this departure is as follows.

The composition rules $>\mathbf{B}$ and $<\mathbf{B}$ are order-preserving, in the restricted sense that their addition to a pure categorial grammar that does not include higher-order functor categories—that is, ones that take functions as arguments—introduces only new derivations, not new word orders.[6]

On the other hand, the rules $>\mathbf{B}_{\times}$, $<\mathbf{B}_{\times}$, $>\mathbf{S}_{\times}$, and $<\mathbf{S}_{\times}$ that combine functions of different directionality have a *permutation* property. That is, they have the effect of reordering arguments, even for first-order grammar fragments. Indeed, Moortgat (1988a), following van Benthem (1986), shows that merely adding non-order-preserving composition to the axioms of the Lambek calculus causes the system to collapse, generating the permutation closure on the context-free language defined by the lexicon.

It does not of course follow that adding such rules to other kinds of categorial grammar engenders the same collapse. We will see in part III some results due to Weir (1988) and Rambow (1994a,b), which show that a CCG of the present form is not permutation-complete and is in fact under certain assumptions weakly equivalent to TAG (Joshi, Levy and Takahashi 1975).[7]

However, any grammar for a configurational language that includes any of the non-order-preserving rules may have to restrict their application to certain types. (We have already seen one such restriction, in the case of the restriction of the variable *X* in the backward crossed substitution rule (1) to categories such as *VP*.) I will continue to defer discussion of how these type restrictions are imposed until a later chapter.

The existence of extremely nonconfigurational languages suggests that much of the freedom allowed by the three principles via the non-order-preserving rules may be exploited in other languages (see van der Zee 1982; Steedman 1985; Zwarts 1986; Bouma 1985). In particular, we will see that the combinatory grammars of English and Dutch between them require *all* of the above composition rules, both order-preserving and non-order-preserving.

The way in which the principles restrict the rules of type-raising (whether considered as lexical rules or rules of active syntax) is less obvious. This is dealt with in detail in chapter 8. For present purposes we can assume that type-raising is restricted to the following pair of rules:[8]

(10) *The Order-preserving type-raising rules*

a. $X : a \quad \Rightarrow_{\mathbf{T}} \quad T/(T\backslash X) : \lambda f.f\ a \qquad (>\mathbf{T})$

where $T\backslash X$ is a parametrically licensed category for the language

b. $X : a \quad \Rightarrow_{\mathbf{T}} \quad T\backslash(T/X) : \lambda f.f\ a \qquad (<\mathbf{T})$

where T/X is a parametrically licensed category for the language

The restriction, which was discussed in chapter 3, only allows type-raising over categories that are permitted for the language under the Principle of Categorial Type Transparency, defined informally in Chapter 3. It embodies the idea that raised categories are limited to a fixed set of functions sanctioned by the $\bar{\text{X}}$ theory and language-specific word order parameters. Without some such restriction, type-raising can be used to create types that will allow any two arbitrary adjacent categories to compose, causing overgeneralization, as Houtman points out (1994, 63-85).[9]

However, the way in which type-raising and composition interact to allow the equivalent of unbounded extraction still has the potential to create some non-nominal objects with the same parametrically licensed category as a type-raised subject, as in the following partial derivation:

(11)

$$\begin{array}{llllll}
\text{a man whom} & \text{I} & \text{think} & \text{that} & \text{Dexter} & \text{likes} \\
 & \overline{S/(S\backslash NP)} & \overline{(S\backslash NP)/S'} & \overline{S'/S} & \overline{S/(S\backslash NP)} & \overline{(S\backslash NP)/NP} \\
 & \multicolumn{2}{c}{\underline{\qquad\qquad\qquad\qquad}{>}\mathbf{B}} & & & \\
 & \multicolumn{2}{c}{S/S'} & & & \\
 & \multicolumn{3}{c}{\underline{\qquad\qquad\qquad\qquad\qquad\qquad}{>}\mathbf{B}} & & \\
 & \multicolumn{3}{c}{S/S} & & \\
 & \multicolumn{4}{c}{\underline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad}{>}\mathbf{B}} & \\
 & \multicolumn{4}{c}{S/(S\backslash NP)} &
\end{array}$$

Whether the formation of such constituents should be welcomed depends upon whether right node raising of embedded tensed VPs is permitted, as in (12a), in a manner parallel to subject coordination apart from the details of agreement, as in (12b):

(12) a. ?[You doubt that Dexter,]$_{S/(S\backslash NP)}$ but [I wonder whether Warren,]$_{S/(S\backslash NP)}$ is a genius.

b. [Dexter]$_{S/(S\backslash NP)}$ and [Warren]$_{S/(S\backslash NP)}$ are geniuses.

The fact that such nonstandard constituents would have the same syntactic type as a type-raised subject threatens to allow illegal coordinations like the following (discussed in Steedman 1987, Wood 1988, and Henderson 1992, among other papers), which are much worse:[10]

(13) *[Dexter,]$_{S/(S\backslash NP)}$ and [I wonder whether Warren,]$_{S/(S\backslash NP)}$ is a genius/are geniuses.

It is important to realize that this problem is not restricted to CCG. Any theory that treats sentences like (12a) as arising from the equivalent of right node raising the finite VP will overgeneralize in the same way. Elsewhere, (Steedman 1990, 222–223), I have proposed as a technical solution to exclude altogether the formation of "pseudosubjects" like *you doubt that Dexter*, via a restriction on forward composition, excluding composition with Z bound to the tensed predicate category. However, the marginal acceptability of (12a), coupled with the sensitivity of agreement to the distinction between (12a) and (12b), suggests that this proposal cannot be maintained—see Houtman (1994) for further discussion.[11]

The property of all nominal type-raised arguments including subjects that distinguishes them from propositional pseudosubjects is that semantically they are generalized quantifiers and/or referential expressions, headed by nouns. Such expressions have a number of distinctive semantic properties, such as "conservativity" (Keenan and Faltz 1985, 16-17) that are not shared by entities like *You think (that) Dexter*, which are headed by verbs. Conservativity is the property of a function f which makes the following equation true:

(14) f(students) are vegetarians $\Longleftrightarrow$ f(students) are both students and vegetarians. (Keenan and Faltz 1985, 17)

Clearly, this is the defining property of determiners like *every*. Equally clearly, it does not even begin to apply to "pseudodeterminers" like *I think that*, which do not generate referential expressions at all.

Henderson (1992) shows how to use syntactic indices to distinguish the two types of category. However, it is also possible to argue that the anomaly in (13) is purely semantic, a variety of zeugma or equivalently syllepsis, arising from the incompatibility of their interpretations, comparable to that in the following real-life example, which I owe to Richard Shillcock:

(15) This flour is suitable for vegetarians, freezing, pizza dough, and home bread-making machines.

The increased anomaly of the pseudosubject example (13) could then be presumed to stem from the fact that pseudosubjects just don't make very good conjuncts in the first place—cf. (12a.).

4.2 Linguistic Constraints on Unbounded Dependencies

Chapter 3 showed that the involvement of combinatory rules offers a common mechanism for canonical word order, leftward extraction constructions, and right-node-raising constructions, based on a single lexical entry for the verb, in keeping with the Principle of Head Categorial Uniqueness. The combinatory theory accordingly makes a broad prediction that *any argument that can take part in a leftward extraction will also permit the corresponding rightward movement*. Adjunct island constraints and subjacency constraints, which follow from the categories of adjuncts themselves and the type-raised status of arguments, should apply similarly to either permit or prevent both varieties.

This prediction of the theory is broadly true. However, there are a number of exceptions which are considered in detail in Steedman 1996b. Here I briefly examine just two of them—namely, asymmetries associated with subject extraction and with heavy NP shift constructions.

4.2.1 Subject Extraction Asymmetries

A number of further constraints on long-range dependencies that are asymmetrical with respect to subjects and objects, and that have been argued to stem from Chomsky's (1981) Empty Category Principle (ECP), arise in present terms because the categories reflect the different directionality of the subject and object arguments of the SVO verb. This ingredient of the theory captures the concept of "canonical government configuration" or "direction of government" (see Kayne 1983, 167–169; Pesetsky 1982; and Koster 1986, 19) directly in the lexicon and its projections under the combinatory rules, as Bach (1988, 29), among others, has pointed out. In present terms, this principle is an inevitable consequence of the Principle of Inheritance.

For example, Szabolcsi (1989), Bach (1988), and I (Steedman 1987, 1996b) discuss the way that the theory predicts the following familiar asymmetry in extractability of English subjects and objects, which has been attributed in other frameworks to various constraints on subject positions, including the ECP:

(16) a. (a man who(m)) [I think that]$_{S/S}$ [Dexter likes]$_{S/NP}$

b. *(a man who(m)) [I think that]$_{S/S}$ [likes Dexter]$_{S\backslash NP}$

According to the present theory, this asymmetry is possible in languages like English that have an SVO lexicon because the crucial composition that would potentially permit subject extraction by combining S/S and $S\backslash NP$ requires a distinct *slash-crossing* instance of composition, $>\mathbf{B}_{\times}$:

(17) $X/Y \quad Y\backslash Z \quad \Rightarrow \quad X\backslash Z$

Although such rules are permitted (and therefore predicted) by the theory, we cannot by adding such a rule specify a language that is exactly like English except for allowing general subject extraction. As I pointed out in the Steedman (1996b), if we did so, the grammar would lack another distinguishing property of English, namely, its configurationality. Word order would collapse entirely, allowing "scrambling" examples like the following:

(18) *I Dexter [think (that) likes Warren]$_{(S\backslash NP)\backslash NP}$

Thus, the theory predicts that asymmetries in extractability for categories that are arguments of the same verb depend upon asymmetries in the directionality of those arguments.[12] The fact that this particular asymmetry tends to be characteristic of configurational SVO languages and constructions therefore follows without the stipulation of any subject-specific condition or ECP.

A number of further phenomena including binding possibilities for certain negative polarity items such as *personne* in French and *nessuno* in Italian have been ascribed to the operation of the ECP at the level of LF (Kayne 1983; Jaeggli 1981; Rizzi 1982). These phenomena are shown in Steedman 1996b to also follow from the way in which directionality is projected in a combinatory grammar in Surface Structures, without the stipulation of subject-specific conditions or the equivalent of the ECP. Some related restrictions on quantifier scope alternation are discussed in section 4.4 below.

In Steedman (1996b) I also discuss some obvious exceptions to the general nonextractability of subjects, including the fact that English subjects can be extracted from bare complements:

(19) a. a man who(m) I think likes Dexter
b. a man who(m) I think Dexter likes

We cannot include such sentences by allowing a rule of crossed forward composition, no matter how restricted. Such a mechanism would immediately cause overgenerations parallel to (18). The only degree of freedom that remains within the present theory is to assume that this phenomenon arises in the lexicon. We must assume that, in addition to obvious categories like VP/S' and VP/S, verbs like *think* bear a special subject-extracting category. I will assume that this category takes the form (20).

(20) think := $(VP/NP_{+ANT,agr})/(S\backslash NP_{agr})$

In essence, this category embodies the GPSG analysis of extractable subjects proposed by Gazdar (1981) and Gazdar et al. (1985), as modified by Hepple (1990, 58) within a different categorial framework. (The advantage of the present proposal lies in the way most subject extraction is *excluded*.)[13] The *NP* argument of this category bears a feature $+ANT$ (mnemonic for the GB concept of "antecedent government"), which, like Hepple's corresponding "modality" $\triangle$, prevents this argument from being saturated by a normal lexically realized argument of any kind. The feature is in every respect like the agreement features discussed earlier. Indeed, the argument in question includes a number agreement feature *agr*, which works in the usual way via the relative clause category $(N_{agr} \backslash N_{agr})/(S/NP_{agr})$ to exclude the following examples:

(21) a. *a man who(m) I think like marmalade
b. *some men who(m) I think likes marmalade

Category (20) is clearly an exception to the Principle of Head Categorial Uniqueness, and as such counts against the theory as a stipulation. However, to the extent that it is a stipulation confined to the small number of subjects that *do* extract, rather than a negative constraint on the majority of subjects that do not, this lexicalist account may yet compare favorably with Chomsky and Lasnik (1977) **That*-Trace Filter–based and Chomsky's (1981) ECP-based accounts, especially in view of the evidence from Maling and Zaenen (1978; see also e.g. Chung 1983 and Engdahl 1985) that the general prohibition against subject extraction is not universal, and appears to correlate with canonical word order, as the present theory would predict.

On the assumptions (a) that arguments other than topicalized ones and relative pronouns are marked as $-ANT$, (b) that the restriction of X in the order-preserving type-raising rules to argument types includes this property by definition, (c) that all normal arguments of verbs are $?ANT$—that is, compatible with either + or − on this feature—and (d) that the argument of the relative-pronoun category is $S/NP_{?ANT,agr}$, we capture the following asymmetry:

(22) a man who(m)$_{(N\backslash N)/(S/NP_{?ANT,3sm})}$ [I think likes marmalade]$_{S/NP_{+ANT,3sm}}$

(23) *[I think likes marmalade]$_{S/NP_{+ANT,3sm}}$ [this very heavy man.]$_{NP^{\uparrow}_{-ANT,3sm}}$

Further details of the Fixed-Subject Condition and the bare-complement exception are explored in Steedman 1996b, where the feature $\pm ANT$ is called $\pm LEX$. Further support for the proposal that bare complement subject extraction is mediated by a special-case lexical category like the one proposed here is to be found in the careful corpus-based developmental work of Stromswold (1995), who shows that children acquire long-distance object questions before they acquire long-distance subject questions. In fact, embedded subject questions are extremely rare in young children's speech. Stromswold finds them in only 1 out of 11 children's data, and then only from age 5.0 (Stromswold 1995, sec. 8.2), suggesting that this construction is among the very last details of English grammar to be mastered, a point to which I will return in chapter 10.

4.2.2 Other Extraction Asymmetries

Although the restricted possibilities for subject extraction in English do not involve the forward crossed composition rule, the grammar of English does appear to require the other non-order-preserving composition rule permitted by the Principles of Consistency and Inheritance. In order to accommodate heavy NP shift and related coordinations like the following, Moortgat (1988a) and I (Steedman 1987) have proposed rules of backward crossed composition:

(24) *Backward crossed composition (preliminary version)* ($<\mathbf{B}_{\times}$)
$Y/Z \quad X\backslash Y \quad \Rightarrow_{\mathbf{B}} \quad X/Z$
where $X, Y = S\$$

Like the backward crossed substitution rule (1), this non-order-preserving rule must be restricted to combinations where Y is a verbal category $S\$$ such as VP or VP_{ing}.[14]

(25)
I : $S/(S\backslash NP)$ — shall : $(S\backslash NP)/VP$ — buy : VP/NP — today : $VP\backslash VP$ — and : $CONJ$ — cook : VP/NP — tomorrow : $VP\backslash VP$ — the mushrooms … : NP
buy today: VP/NP ($<\mathbf{B}_{\times}$)
cook tomorrow: VP/NP ($<\mathbf{B}_{\times}$)
buy today and cook tomorrow: VP/NP ($<\Phi>$)
buy today and cook tomorrow the mushrooms …: VP ($>$)

I will come back to the reason why the "shifted argument" must be "heavy" in chapter 5.

The same rule correctly applies to type-raised arguments:

(26)
I will : S/VP — give : $(VP/PP)/NP$ — to my sister : $VP\backslash(VP/PP)$ — an engraving by Rembrandt : NP
give to my sister: VP/NP ($<\mathbf{B}_{\times}$)
I will give to my sister: S/NP ($>\mathbf{B}$)
I will give to my sister an engraving by Rembrandt: S ($>$)

The rule also allows leftward extraction of "nonperipheral" arguments, including examples like the following, relevant to the earlier discussion of parasitic gaps:

(27) (articles) which I will file tomorrow

$$
\begin{array}{cccc}
\text{which} & \text{I will} & \text{file} & \text{tomorrow}\\
\overline{(N\backslash N)/(S/NP)} & \overline{S/VP} & \overline{VP/NP} & \overline{VP\backslash VP}\\
 & & \multicolumn{2}{c}{\underline{\hspace{6em}}{<}\mathbf{B}_{\times}}\\
 & & \multicolumn{2}{c}{VP/NP}\\
 & \multicolumn{3}{c}{\underline{\hspace{10em}}{>}\mathbf{B}}\\
 & \multicolumn{3}{c}{S/NP}\\
\multicolumn{4}{c}{\underline{\hspace{20em}}{>}}\\
\multicolumn{4}{c}{N\backslash N}
\end{array}
$$

Identical compositions are crucial in the derivation of other relativizations of nonperipheral elements including the following:

(28) a. an engraving which I will *buy today and sell tomorrow*
b. an engraving which I will *show to him and give to you*
c. a man who(m) I will *show a painting and give a flower*

However, the last of these shows that the general rule (24) must be replaced by a number of more specific instances, since examples like the following show that nominal ditransitives are an exception to the general rule that whatever can leftward extract can also rightward extract:

(29) *I will give a flower this very heavy policeman.

This asymmetry is related to the observation of Ross (1967) that Heavy NP shift, unlike relativization and right node raising, also cannot induce preposition stranding:

(30) a. the city which I will travel to today and return from tomorrow
b. I will travel to and return from the beautiful city of Dublin.
c. *I will travel to tomorrow the beautiful city of York.

As a technical device to capture these asymmetries, we can replace the backward crossed composition rule (24) by two more specific instantiations. The first allows both leftward extraction (28a,b) and heavy shift (25) and (26) of any argument not explicitly marked as forbidden to shift by a negative value of a feature *SHIFT*, with respect to which all normal arguments are assumed to be unspecified:[15]

(31) *Backward crossed composition I* ($<\mathbf{B}_{\times}$)
$Y/Z_{+SHIFT} \quad X\backslash Y \quad \Rightarrow_{\mathbf{B}} \quad X/Z_{+SHIFT}$
where $X, Y = S\$$

We may then assume that the dative NP argument of ditransitive verbs and the complement of prepositions are marked as $-SHIFT$:

(32) a. give := $(VP/NP)/NP_{-SHIFT}$
b. to := PP/NP_{-SHIFT}

The second instance of the heavy-shifting backward crossed composition rule then allows such $-SHIFT$ nonperipheral arguments to leftward-extract, but not to rightward-extract, by marking them for antecedent government only:[16]

(33) *Backward crossed composition II* ($<\mathbf{B}_{\times}$)
$Y/Z_{-SHIFT,+ANT} \quad X\backslash Y \quad \Rightarrow_{\mathbf{B}} \quad X/Z_{-SHIFT,+ANT}$
where $X, Y = S\$$

The rule will therefore allow leftward-extraction in examples like (28c), while excluding (29). I will return to this restriction in chapter 6, where it will be apparent that a related restriction applies to Dutch (but not German) main-clause order.

As Ades and I note (Ades and Steedman 1982), the crossed composition mechanism automatically excludes extraction out of the shifted-over PP in examples like (26), even when the PP is subcategorized for, to exclude sentences like the following, which violate the Clause Non-Final Incomplete Constituent Constraint proposed by Kuno (1973):

(34) *a woman who(m) I will $[\text{give}]_{(VP/PP)/NP}$ $[\text{to}]_{PP/NP}$ [an engraving by Rembrandt$]_{NP}$

The crossed composition mechanism also automatically excludes heavy shift of subjects in (35a). However, it is only the stipulative restriction of the backward crossed rule to composition into verbs that prevents heavy shift out of subjects (35b):[17]

(35) a. *$[\text{Smiled}]_{S\backslash NP}$ [the man in the grey flannel suit.$]_{S/(S\backslash NP)}$
b. *[Every friend of$]_{NP/NP}$ $[\text{smiled}]_{S\backslash NP}$ [the man in the grey flannel suit$]_{NP}$

4.3 Linguistic Constraints on Bounded Dependencies

It will be clear from the discussion in the previous sections that combinatory grammars embody an unusual view of Surface Structure, according to which strings like *Anna married* are, quite simply, surface constituents. We will see that this view directly generalizes to cover all of the "fragments" that arise under coordination and related constructions, including *a policeman a flower*, the serial verb clusters that are characteristic of Germanic "verb raising," and the strings that can be isolated as intonational phrases.

According to this view, Surface Structure is also more ambiguous than has previously been realized, for such strings must also be possible constituents of noncoordinate sentences like *Anna married Manny* and *Give a policeman a flower*, as well. It follows that such sentences must have several Surface Structures, corresponding to different sequences of composition, type-raising, and application.

For example, the derivation in (37) is allowed for the former sentence, in addition to the traditional derivation in (7) of chapter 3, and repeated here as (36).

(36)
$$\begin{array}{ccc}
\text{Anna} & \text{married} & \text{Manny}\\
\hline
NP & (S\backslash NP)/NP & NP\\
: anna' & : \lambda x.\lambda y.marry'xy & : manny'\\
\overline{\qquad\qquad}{>}\mathbf{T} & & \overline{\qquad\qquad}{<}\mathbf{T}\\
T/(T\backslash NP) & & T\backslash(T/NP)\\
: \lambda p.p\ anna' & & : \lambda q.q\ manny'\\
\hline
 & \multicolumn{2}{c}{S\backslash NP : \lambda y.marry'manny'y \quad <}\\
\hline
\multicolumn{3}{c}{S : marry'manny'anna' \quad >}
\end{array}$$

(37)
$$\begin{array}{ccc}
\text{Anna} & \text{married} & \text{Manny}\\
\hline
NP & (S\backslash NP)/NP & NP\\
: anna' & : \lambda x.\lambda y.marry'xy & : manny'\\
\overline{\qquad\qquad}{>}\mathbf{T} & & \overline{\qquad\qquad}{<}\mathbf{T}\\
T/(T\backslash NP) & & T\backslash(T/NP)\\
: \lambda p.p\ anna' & & : \lambda q.q\ manny'\\
\hline
\multicolumn{2}{c}{S/NP : \lambda x.marry'x\ anna' \quad >\mathbf{B}} & \\
\hline
\multicolumn{3}{c}{S : marry'manny'anna' \quad <}
\end{array}$$

The most important property of such families of alternative derivations is that they form semantic equivalence classes, for as the derivations show, the semantics of the combinatory rules guarantees that all such derivations will deliver an interpretation determining the same function-argument relations.[18] Indeed, there is a close relation between the canonical interpretation structures that they deliver according to the theory sketched above, and traditional notions of constituent structure.

One could in fact argue that the dominance of the traditional notion of Surface Structure is an instance of an error that Chomsky (1957, chap. 7) warned against, namely, that of depending upon ill-defined intuitions about meaning (with which most traditional tests for constituency other than susceptibility to coordination are confounded), rather than empirical data concerning syntactic form.

In CCG, the work that is done in GB by *wh*-trace, mediated or constrained by indexing, subject to structural relations such as c-command, is done without empty categories. CCG uses only the same lexical mechanism that binds arguments in situ to verbs, projected from lexical categories by the combinatory rules that mediate long-range dependencies.[19] This amounts to saying that the present theory eschews the distinction between "antecedent" government and "head" or "θ" government. There is only lexical head government, as the Principle of Head Categorial Uniqueness requires. All syntactic dependencies are projections of that, without mediation by empty categories or Ā structural positions. The consequences of this move for capturing *wh*-constructions and the constraints upon them is investigated in more detail in Steedman 1996b.

However, it follows that the entire range of grammatical phenomena that depend on structural relations like c-command, which have traditionally been dealt with in Surface Structure, must in a combinatory grammar be handled at the level of interpretation or predicate-argument structure.

4.3.1 Binding and Control

This important class of phenomena includes most of the class of constructions that were identified earlier as "bounded," importantly including the systems of (anaphor) binding and control. This proposal, including certain interactions of binding theory and *wh*-constructions, is also investigated in further detail in Steedman 1996b.

For example, the following category for "equi" verbs such as *tries* is proposed there:

(38) $(S\backslash NP_{3s})/(S_{to-inf}\backslash NP_{3s}) : \lambda p.\lambda y.try'\ (p\ (ana'\ y))\ y$

This is essentially identical to the standard GB analysis, with two slight departures. First, the responsibility for determining dependencies that have sometimes been accounted for in terms of bounded movement has been relegated to the lexicon and to the relation between syntactic category and predicate-argument structure. Second, rather than merely using a constant such as *PRO* to represent the controlled argument at Surface Structure or S-Structure, leaving to the binding theory or an autonomous module of control theory the task of establishing the antecedent, the present theory makes the lexical entry for the control verb do part of that work, by making (the interpretation of) its complement's subject a "pro-term" $ana'\ y$ resembling an anaphor bound to the (interpretation y of the) subject of the control verb.[20] On the assumption that the infinitival verb *like* has the obvious category (39a), and that the complementizer *to* has the trivial category (39b), which can compose with either infinitival, the category (38) will yield the results in (40):

(39) a. like := $(S_{inf}\backslash NP_{agr})/NP : \lambda x.\lambda y.like'xy$
  b. to := $(S_{to-inf}\backslash NP_{agr})/(S_{inf}\backslash NP_{agr}) : \lambda x.x$

(40)
$$
\begin{array}{cccc}
\text{Dexter} & \text{tries} & \text{to like} & \text{Warren} \\
\overline{NP_{3sm}} & \overline{(S\backslash NP_{3s})/(S_{to-inf}\backslash NP_{3s})} & \overline{(S_{to-inf}\backslash NP_{agr2})/NP} & \overline{NP} \\
: dexter' & : \lambda p.\lambda y.try'\ (p\ (ana'y))y & : \lambda x.\lambda y.like'xy & : warren' \\
 & & \multicolumn{2}{c}{\overline{S_{to-inf}\backslash NP_{agr2} : \lambda y.like'warren'y}^{>}} \\
 & \multicolumn{3}{c}{\overline{S\backslash NP_{3s} : \lambda y.try'\ (like'warren'(ana'y))y}^{>}} \\
\multicolumn{4}{c}{\overline{S : try'\ (like'warren'(ana'dexter'))dexter'}^{<}}
\end{array}
$$

An appropriate binding theory can then be defined in terms of a relation of command over predicate-argument structures (say, as in Steedman 1996b).

A similar analysis can be applied to object control. The following is the full category of the verb *persuades* $((S\backslash NP)/(S\backslash NP))/NP$, reflecting the assumption that predicate-argument structures observe the obliqueness hierarchy:

(41) persuades := $((S\backslash NP_{agr_2})/(S_{to-inf}\backslash NP_{agr_1}))/NP_{agr_1}$
  $: \lambda x.\lambda p.\lambda y.persuade'(p\ (ana'x))\ x\ y$

The category (41) again embodies a "wrap" analysis of object control verbs, akin to that proposed by Bach (1979, 1980), Dowty (1982), Szabolcsi (1989), Jacobson (1990), and Hepple (1990), again at the level of lexical predicate-argument structure or Logical Form rather than syntactic or phrasal derivation. That is, the command relation between the interpretation of the object NP and the predicate-argument is *reversed* with respect to the derivation. This represents a minor departure from the Montagovian mainstream, in which such use of Logical Form is frowned upon. However, the reason for embodying this widespread assumption in the lexicon rather than in active WRAP rules in syntax is implicit in the analysis of coordinations like (40) of chapter 3, and I will return to it frequently below.

When applied to an object like *Martha* and an infinitival like *to go*, the category (41) gives rise to derivations like the following:

(42)

$$
\begin{array}{cccc}
\text{George} & \text{persuades} & \text{Martha} & \text{to go}\\
\overline{NP_{3sm}} & \overline{((S\backslash NP_{3s})/(S_{to-inf}\backslash NP_{agr}))/NP_{agr}} & \overline{NP_{3sf}} & \overline{S_{to-inf}\backslash NP_{agr2}}\\
:george' & :\lambda x.\lambda p.\lambda y.persuade'(p\ (ana'x))xy & :martha' & :\lambda x.go'x\\
 & \multicolumn{2}{c}{\overline{\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad}>} & \\
 & \multicolumn{2}{c}{(S\backslash NP_{3s})/(S_{to-inf}\backslash NP_{3sf})} & \\
 & \multicolumn{2}{c}{:\lambda p.\lambda y.persuade'(p\ (ana'martha'))martha'y} & \\
 & \multicolumn{3}{c}{\overline{\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad}>}\\
 & \multicolumn{3}{c}{S\backslash NP_{3s}}\\
 & \multicolumn{3}{c}{:\lambda y.persuade'(go'(ana'martha'))martha'y}\\
\multicolumn{4}{c}{\overline{\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad}>}\\
\multicolumn{4}{c}{S:persuade'(go'(ana'martha'))martha'george'}
\end{array}
$$

The predicate-argument structure of the category (41) embodies a very widespread generalization about categories (see the discussion of VSO and VOS languages in Keenan 1988, according to which rightward functor categories like VSO verbs, which obey the very strong tendency of the languages of the world to reflect obliqueness ordering in SO string order, must wrap—that is, reverse in the predicate-argument structure the command relations implicit in the syntactic category itself). This generalization can also be observed in the ditransitives discussed at the end of chapter 2, to which I will return in chapter 7:

(43) showed := $((S\backslash NP)/NP)/NP : \lambda x.\lambda y.\lambda z.show'yxz$

Following Clark (1997), auxiliaries can be handled as "raising" verbs with the following kind of category:[21]

(44) might := $(S\backslash NP)/VP : \lambda p.\lambda x.might'(p\,x)$

The following kind of derivation results:

(45)

$$
\begin{array}{cccc}
\text{Anna} & \text{might} & \text{marry} & \text{Manny}\\
\overline{S/(S\backslash NP)} & \overline{(S\backslash NP)/VP} & \overline{VP/NP} & \overline{S\backslash(S/NP)}\\
:\lambda p.p\ anna' & :\lambda p.\lambda x.might'(p\ x) & :marry' & :\lambda p.p\ manny'\\
\multicolumn{2}{c}{\overline{\quad\quad\quad\quad\quad\quad\quad\quad}>\mathbf{B}} & & \\
\multicolumn{2}{c}{S/VP:\lambda p.might'(p\ anna')} & & \\
\multicolumn{3}{c}{\overline{\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad}>\mathbf{B}} & \\
\multicolumn{3}{c}{S/NP:\lambda x.might'(marry'x\ anna')} & \\
\multicolumn{4}{c}{\overline{\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad}>}\\
\multicolumn{4}{c}{S:might'(marry'manny'anna')}
\end{array}
$$

To handle binding, raising, and control lexically via a level of interpretation or predicate-argument structure is not especially controversial. A similar move has been proposed within a Montague Grammar framework by Bach and Partee (1980), within LFG by Bresnan (1982), within GB by Lasnik and Saito (1984), and within HPSG by Pollard and Sag (1992). A similar position seems to be implicit in Hale and Keyser 1993 and Brody 1995, which suggest that much of the apparatus in GB and the Minimalist Program amounts to a theory of the

CCG lexicon—a component of the present theory that has been conspicuous by its absence in the presentation so far.

However, within this broad consensus concerning the domain of binding and/or control, two camps should be distinguished. In the first are the Montague Grammarians and the proponents of virtually all varieties of Flexible Categorial Grammar since Bach (1976), including Shaumyan (1977), Dowty (1979), Jacobson (1987), Szabolcsi (1989), Moortgat (1988b), Hepple (1990), Morrill (1994), and many others. These authors deny the existence of any autonomous level of semantic representation such as predicate-argument structure intervening between syntactic derivation and the model theory. That is to say, although they may use a Logical Form for notational transparency, they eschew the exploitation of any structural property of such notations, such as the analogue of GB c-command. There is a strong affinity between these researchers and the model-theoretic tradition in Mathematical Logic.

The members of the other camp, which includes most researchers in LFG, GB, G/HPSG, and the theory proposed here, as well as virtually all computational linguists, define an autonomous structural level of predicate-argument structure or Logical Form, and define the grammar of binding in terms of structural dominance and command, making an intrinsic use of predicate-argument structure. These researchers resemble in spirit the proof-theoretic tradition in Mathematical Logic.

The analogy with the proof-theory/model-theory duality in logic suggests that this difference may turn out not to be an empirically testable one. It is likely to be the case that anything that can be captured one way can be captured the other way, and vice versa. The question will probably be resolved on the basis of the simplicity of the rival theories. Since CCG is principally concerned with the unbounded constructions, it diverges from many of its categorial relatives in adopting an explicitly predicate-argument-structural account of binding, control, and the other bounded phenomena, simply because it seems to make life easier for the linguistically oriented reader.

This should not be taken as constituting a serious disagreement with these other categorial approaches. The very fact that these phenomena are all bounded by the local domain of the verb means that the mapping from linear order to obliqueness order is essentially trivial.

4.3.2 Adjunct-Argument Asymmetry

English prepositions heading adverbial PPs that one would normally think of as adjuncts rather than arguments can be "stranded" by relativization, as in the following example:

(46) the painting that I folded the rug over

One would not normally think of *fold* as subcategorizing for such a prepositional phrase. However, reflexive binding and the impossibility of parasitic gapping makes it clear that the PP argument is more oblique than—that is, is c-commanded at LF by—the object:[22]

(47) a. I folded the rug over itself.
b. *the rug which I folded over

In categorial terms, there is really only one way to permit preposition stranding into such adjunct PPs while still accounting for the above facts. Any tactic that makes the PP a rightmost argument—such as type-raising the VP over the adjunct category, or assigning a particle-like higher-order category over VPs to the preposition—will fail to yield the scope or c-command relations that the binding phenomena require, unless it is accompanied by nontrivial manipulations of Logical Form, in violation of the Principle of Combinatory Type Transparency.

Instead we must assume that the lexical categories for the relevant class of verbs already allow for optional additional rightmost adverbial categories as arguments, and that like all rightward arguments, they wrap at Logical Form, so that their interpretations are more oblique than the obligatory arguments. (Although I will not go into the question of exactly what Logical Form is involved, I assume that it amounts to a form of control—see (42). Hence, (47b) is illicit for the same reason as **Who did you persuade to vote for?*.)

I will largely ignore such optional adjunct-arguments in what follows, but occasionally it will be necessary to recall that they and certain other adjuncts often behave like arguments, rather than true adjuncts.

4.4 Quantification in CCG

Another phenomenon that is naturally analyzed in terms of relations of command at the level of Logical Form is quantifier scope.

It is standard to assume that the ambiguity of sentences like (48) is to be accounted for by assigning two Logical Forms which differ in the scopes assigned to these quantifiers, as in (49a,b):

(48) Every boy admires some saxophonist.

(49) a. $\forall x.boy'x \rightarrow \exists y.saxophonist'y \wedge admires'yx$
b. $\exists y.saxophonist'y \wedge \forall x.boy'x \rightarrow admires'yx$

The question then arises of how the grammar can assign all and only the correct interpretations to sentences with multiple quantifiers.

This process has on occasion been explained in terms of "quantifier movement" or essentially equivalent computational operations of "quantifying in" or "storage" at the level of Logical Form. However, such accounts present a problem for monostratal and monotonic theories of grammar like CCG that try to do away with movement or the equivalent in syntax. Having eliminated nonmonotonic operations from the syntax, to have to restore them at the level of Logical Form would be dismaying, given the strong assumptions of transparency between syntax and semantics from which this and other monotonic theories begin. Given the assumptions of syntactic/semantic transparency and monotonicity that are usual in the Frege-Montague tradition, it is tempting to try to use nothing but the derivational combinatorics of surface grammar to deliver all the readings for ambiguous sentences like (48). Two ways to restore monotonicity have been proposed, namely: enriching the notion of derivation via type-changing operations; or enriching the lexicon and the semantic ontology.

It is standard in the Frege-Montague tradition to begin by translating expressions like *every boy* and *some saxophonist* into "generalized quantifiers"—in effect exchanging the roles of arguments like NPs and functors like verbs by type-raising the former (Lewis 1970; Montague 1973; Barwise and Cooper 1981; see Partee, ter Meulen and Wall 1990, 359 for a review):

In terms of the notation and assumptions of CCG, one way to incorporate generalized quantifiers into the semantics of CG determiners is to transfer type-raising to the lexicon, assigning the following categories to determiners like *every* and *some*, making them functions from nouns to type-raised NPs, where the latter are simply the syntactic types corresponding to a generalized quantifier:

(50) every := $(\mathrm{T}/(\mathrm{T}\backslash NP))/N : \lambda p.\lambda q.\forall x.px \rightarrow qx$
every := $(\mathrm{T}\backslash(\mathrm{T}/NP))/N : \lambda p.\lambda q.\forall x.px \rightarrow qx$

(51) some := $(\mathrm{T}/(\mathrm{T}\backslash NP))/N : \lambda p.\lambda q.\exists x.px \wedge qx$
some := $(\mathrm{T}\backslash(\mathrm{T}/NP))/N : \lambda p.\lambda q.\exists x.px \wedge qx$

Given the categories in (50) and (51), the alternative derivations that CCG permits will deliver the two distinct Logical Forms shown in (49), entirely monotonically and without involving structure-changing operations, as shown in (52) and (53).

(52)

$$
\begin{array}{ccccc}
\text{Every} & \text{boy} & \text{admires} & \text{some} & \text{saxophonist} \\
\hline
(\mathrm{T}/(\mathrm{T}\backslash NP))/N & N & (S\backslash NP)/NP & (\mathrm{T}'\backslash(\mathrm{T}'/NP))/N & N \\
:\lambda p.\lambda q.\forall y.py \rightarrow qy & :\lambda x.boy'x & :\lambda x.\lambda y.admires'xy & \lambda p.\lambda q.\exists x.px \wedge qx & :\lambda x.saxophonist'x \\
\hline
\multicolumn{2}{c}{\mathrm{T}/(\mathrm{T}\backslash NP) \quad >} & & \multicolumn{2}{c}{\mathrm{T}'\backslash(\mathrm{T}'/NP) \quad >} \\
\multicolumn{2}{c}{:\lambda q.\forall y.boy'y \rightarrow qy} & & \multicolumn{2}{c}{:\lambda q.\exists x.saxophonist'x \wedge qx} \\
\hline
& & \multicolumn{3}{c}{S\backslash NP \quad <} \\
& & \multicolumn{3}{c}{:\lambda y.\exists x.saxophonist'x \wedge admires'xy} \\
\hline
\multicolumn{5}{c}{S : \forall y.boy'y \rightarrow \exists x.saxophonist'x \wedge admires'xy \quad >}
\end{array}
$$

(53)

$$
\begin{array}{ccccc}
\text{Every} & \text{boy} & \text{admires} & \text{some} & \text{saxophonist} \\
\hline
(\mathrm{T}/(\mathrm{T}\backslash NP))/N & N & (S\backslash NP)/NP & (\mathrm{T}'\backslash(\mathrm{T}'/NP))/N & N \\
:\lambda p.\lambda q.\forall y.py \rightarrow qy & :\lambda x.boy'x & :\lambda x.\lambda y.admires'xy & \lambda p.\lambda q.\exists x.px \wedge qx & :\lambda x.saxophonist'x \\
\hline
\multicolumn{2}{c}{\mathrm{T}/(\mathrm{T}\backslash NP) \quad >} & & \multicolumn{2}{c}{\mathrm{T}'\backslash(\mathrm{T}'/NP) \quad >} \\
\multicolumn{2}{c}{:\lambda q.\forall y.boy'y \rightarrow qy} & & \multicolumn{2}{c}{:\lambda q.\exists x.saxophonist'x \wedge qx} \\
\hline
\multicolumn{3}{c}{S/NP \quad >\mathbf{B}} & & \\
\multicolumn{3}{c}{:\lambda x.\forall y.boy'y \rightarrow admires'xy} & & \\
\hline
\multicolumn{5}{c}{S : \exists x.saxophonist'x \wedge \forall y.boy'y \rightarrow admires'xy \quad <}
\end{array}
$$

The idea that semantic quantifier scope is limited by syntactic derivational scope in this way has some very attractive features. For example, it immediately explains why scope alternation is both unbounded as in (54a) and sensitive to island constraints, as in (54b).

(54) a. At least one witness said that the accused knew every victim.
b. Some witness who saw every plaintiff denounced him.

However, linking derivation and scope as simply and directly as this makes the obviously false prediction that in sentences where there is no ambiguity of CCG derivation there should be no scope ambiguity. In particular, object topicalization and object right node raising are derivationally unambiguous in the relevant respects, and force the displaced object to command the rest of the sentence in derivational terms. So they should only have the wide scope reading of the object quantifier. This is not the case:

(55) a. Some saxophonist, every boy admires.
b. Every boy admires, and every girl detests, some saxophonist.

Both sentences have a narrow scope reading in which every individual has some attitude toward some saxophonist, but not necessarily the same saxophonist. This observation appears to imply that even the relatively free notion of derivation provided by CCG is still too restricted to explain all ambiguities arising from multiple quantifiers.

Nevertheless, (55b) has a further property, first observed by Geach (1972), that makes it seem as though scope phenomena are strongly restricted by surface grammar. Although the sentence has one reading where all of the boys and girls have strong feelings toward the same saxophonist—say, John Coltrane—and the reading already noted, according to which their feelings are all directed at possibly different saxophonists, it does not have a reading where the saxophonist has wide scope with respect to *every boy*, but narrow scope with respect to *every girl*—that is, where the boys all admire John Coltrane, but the girls all detest possibly different saxophonists. There does not even seem to be a reading involving separate wide scope saxophonists respectively taking scope over boys and girls—for example, where the boys all admire Coltrane and the girls all detest Lester Young.

These observations are very hard to reconcile with semantic theories that invoke powerful mechanisms like abstraction or Quantifying In and its relatives, (Montague 1973; Cooper 1983; Hobbs and Shieber 1987; Pereira 1990; Keller 1988), or quantifier movement. (Carden 1973; May 1985). For example, if quantifiers are mapped from syntactic levels to canonical subject, object (etc.) position at predicate-argument structure in both conjuncts in (55b) and then migrate up the Logical Form to take either wide or narrow scope, it is not clear why *some saxophonist* should have to take the *same* scope in both conjuncts. The same applies if quantifiers are generated in situ, then lowered to their surface position. Such observations might be countered by the invocation of a "parallelism condition" on coordinate sentences, of the general kind discussed by Fox (1995). But such rules are of very expressively powerful "transderivational" kind that one would otherwise wish to avoid. (See Jacobson (1998) for discussion and arguments against transderivational parallelism.)

Related observations led Keenan and Faltz (1978, 1985), Partee and Rooth (1983), Jacobson (1992a), Hendriks (1993), Oehrle (1994), and Winter (1995, to appear), among others, to propose considerably more general use of type-changing operations than are required in CCG, engendering considerably more flexibility in derivation than seems to be required by the purely syntactic phenomena that have motivated CCG up till now.[23]

Although the tactic of including such order-preserving type-changing operations in the grammar remains a valid alternative for a monotonic treatment of scope alternation in CCG and related forms of categorial grammar, there is no doubt that it complicates the theory considerably. The type-changing operations necessarily engender infinite sets of categories for each word, requiring heuristics based on (partial) orderings on the operations concerned, and raising questions about completeness and practical parsability. Some of these questions have been addressed by Hendriks and others, but the result has been to dramatically raise the ratio of mathematical proofs to sentences analyzed.

It seems worth exploring an alternative response to these observations concerning interactions of Surface Structure and scope-taking. The present section follows Woods (1975), VanLehn (1978), Webber (1978), Fodor (1982), Fodor and Sag (1982), and Park (1995, 1996) in explaining scope ambiguities in terms of a distinction between true generalized quantifiers and other purely referential categories. For example, in order to capture the narrow scope object reading for Geach's right-node-raised sentence (55b), in whose CCG derivation the object must command everything else, the present proposal follows Park in assuming that the narrow scope reading arises from a nonquantificational interpretation of *some saxophonist*, one that gives rise to a reading indistinguishable from a narrow scope reading when it ends up in the object position at the level of Logical Form. The obvious candidate for such a nonquantificational interpretation is some kind of referring expression.

The claim that many NPs that have been assumed to have a single generalized quantifier interpretation are in fact purely referential is not new. Recent literature on the semantics of natural quantifiers has departed considerably from the earlier tendency for semanticists to reduce all semantic distinctions of nominal meaning such as *de dicto/de re*, reference/attribution, etc. to distinctions in scope of traditional quantifiers. There is widespread recognition that many such distinctions arise instead from a rich ontology of different types of (collective, distributive, intensional, group-denoting, arbitrary, etc.) individual to which nominal expressions refer. (See for example Webber 1978, Barwise and Perry 1980, Fodor and Sag 1982, Fodor 1982, Hobbs 1983, 1985, Fine 1985, and papers in the recent collection edited by Szabolcsi 1997.)

One example of such nontraditional entity types (if an idea that apparently originates with Aristotle can be called nontraditional) is the notion of "arbitrary objects" (Fine 1985). An arbitrary object is an object with which properties can be associated but whose extensional identity in terms of actual objects is unspecified. In this respect, arbitrary objects resemble the Skolem terms that are generated by inference rules like Existential Elimination in proof theories of first-order predicate calculus.

I will argue that arbitrary objects so interpreted are a necessary element of the ontology for natural language semantics, and that their involvement in CCG explains not only scope alternation (including occasions on which scope alternation is *not* available), but also certain cases of anomalous scopal binding that are unexplained under any of the alternatives discussed so far.

4.4.1 Donkeys as Skolem Terms

One example of an indefinite that is probably better analyzed as an arbitrary object than as a quantified NP occurs in the following famous sentence, first brought to modern attention by Geach (1962):

(56) Every farmer who owns a donkey$_i$ beats it$_i$.

The pronoun looks as though it might be a variable bound by an existential quantifier associated with *a donkey*. However, no purely combinatoric analysis in terms of the generalized quantifier categories offered earlier allows this, since the existential cannot both remain within the scope of the universal, and come to *c*-command the pronoun, as is required for true bound pronominal anaphora, as in (57):

(57) Every farmer$_i$ in the room thinks that she$_i$ deserves a subsidy

One reaction to this observation has been to treat the existential as a universal in this case, as in Discourse Representation Theory (DRT, Kamp and Reyle 1993), or to generalize the notion of scope, as in Dynamic Predicate Logic (DPL, Groenendijk and Stokhof 1990). However, Webber (1978), Cooper (1979), Evans (1980), Hobbs (1983), and others have pointed out that donkey pronouns in many respects look more like *non*-bound-variable or discourse-bound pronouns, in examples like the following:

(58) Everybody who knows Dexter$_i$ likes him$_i$.

I will assume for the sake of argument that *a donkey* translates at predicate-argument structure as something we might write as $arb'donkey'$. I will assume that the function arb' yields a Skolem term—that is, a term applying a unique functor to all variables bound by universal quantifiers in whose extent $arb'donkey'$ falls. Call it $sk_{donkey}x$ in this case, where sk_{donkey} maps individual instantiations of x—that is, the variable bound by the generalized quantifier *every farmer*—onto objects with the property $donkey'$ in the database.

The mechanism by which arb' "knows" what scopes it is in is presumably the same mechanism whereby a bound variable pronoun "knows" about its binder. This mechanism might be formalized in terms of such "environment passing" devices as "storage" (Cooper 1983) or the related device proposed by Jacobson (1999). However, in the present theory, unlike those of Cooper, Keller (1988), Hobbs and Shieber (1987), Pereira (1990), and Shieber, Pereira and Dalrymple (1996), the mechanism in question offers no autonomous degrees of freedom to determine available readings. An arbitrary object is deterministically bound to *all* scoping universals at the time it is evaluated, and the available readings are thereby entirely determined by the combinatorics of syntactic derivation.

An ordinary discourse-bound pronoun may be bound to this arbitrary object, but unless the pronoun is in the scope of the quantifiers that bind any variables in the Skolem term, it will include a variable that is outside the scope of its binder, and fail to refer.

This analysis is similar to but distinct from the analyses of Cooper (1979) and Heim (1990), who assume that *a donkey* translates as a quantified expression and that the entire subject *every farmer who owns a donkey* establishes a contextually salient function mapping farmers to donkeys, with the pronoun specifically of the type of such functions. However, by making the pronoun refer instead to a Skolem term or arbitrary object, we free our hands to make the inferences we draw on the basis of such sentences sensitive to world knowledge. For example, if we hear the standard donkey sentence and know that farmers may own more than one donkey, we will probably infer on the basis of knowledge about what makes people beat an arbitrary donkey that the farmer beats all of them. On the other hand, we will not make a parallel inference on the basis of the following sentence (attributed to Jeff Pelletier), and the knowledge that some people have more than one dime in their pocket:

(59) Everyone who had a dime in their pocket put it in the parking meter.

The reason is that we know that the reason for putting a dime into a parking meter, unlike the reason for beating a donkey, is voided by the act itself.

The proposal to translate indefinites as Skolem termlike discourse entities is anticipated in much early work in Artificial Intelligence and Computational Linguistics, including Kay (1973), Woods (1975 76-77), VanLehn (1978), and Webber (1983, 353, cf. Webber 1978, 2.52), and also by Chierchia (1995), and Schlenker (to appear). Skolem functors are closely related to, but distinct from, "Choice Functions" (see Reinhart 1997, Winter 1997, Sauerland 1998, and Schlenker to appear for discussion. Webber's (1978) analysis is essentially a choice functional analysis, as is Fine's.)

4.4.2 Scope Alternation and Skolem Entities

If indefinites can be assumed to have a referential translation as an arbitrary object, rather than a meaning related to a traditional existential generalized quantifier, then other supposed quantifiers, such as *some/a few/two saxophonists*, may also be better analyzed as referential categories.

We will begin by assuming that *some* is not a quantifier, but rather a determiner of a (singular) arbitrary object. It therefore has the following pair of subject and complement categories, in place of those in (51):

(60) a. some := $(\mathrm{T}/(\mathrm{T}\backslash NP))/N : \lambda p.\lambda q.q(arb'p)$
b. some := $(\mathrm{T}\backslash(\mathrm{T}/NP))/N : \lambda p.\lambda q.q(arb'p)$

In this pair of categories, the constant arb' is the function identified earlier from properties p to entities of type e with that property, such that those entities are functionally related to any universally quantified NPs that have scope over them at the level of Logical Form. If $arb'p$ is not in the extent of any universal quantifier, then it yields a unique arbitrary constant individual.

We will assume that *every* has at least the generalized quantifier determiner category given in (50), repeated here:

(61) a. every := $(\mathrm{T}/(\mathrm{T}\backslash NP))/N : \lambda p.\lambda q.\forall x.px \rightarrow qx$
b. every := $(\mathrm{T}\backslash(\mathrm{T}/NP))/N : \lambda p.\lambda q.\forall x.px \rightarrow qx$

These assumptions, as in Park's related account, provide everything we need to account for all and only the readings that are actually available for Geach's sentence (55b), *Every boy admires and every girl detests some saxophonist.* Thus the narrow scope saxophonist reading of this sentence results from the (backward) referential category (60b) applying to the translation of *Every boy admires and every girl detests* of type S/NP (whose derivation is taken as read), as in (62).[24]

(62)

Every boy admires and every girl detests	some saxophonist
S/NP	$S\backslash(S/NP)$
$: \lambda x.and'(\forall y.boy'y \rightarrow admires'xy)(\forall z.girl'z \rightarrow detests'xz)$	$: \lambda q.q(arb'sax')$

$$\underline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{<}$$
$$S : and'(\forall y.boy'y \rightarrow admires'(arb'sax')y)(\forall z.girl'z \rightarrow detests'(arb'sax')z)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$S : and'(\forall y.boy'y \rightarrow admires'(sk'_{sax_1}y)y)(\forall z.girl'z \rightarrow detests'(sk'_{sax_2}z)z)$$

Crucially, if we evaluate the latter Logical Form with respect to a database after this reduction, as indicated by the dotted underline, for each boy and girl that we examine and test for the property of admiring/detesting an arbitrary saxophonist, we will find (or in the sense of Lewis (1979) "accommodate" or add to our database) a potentially different individual, dependent via the Skolem functors sk'_{sax_1} and sk'_{sax_2} upon that boy or girl. Each conjunct thereby gives the appearance of including a variable bound by an existential within the scope of the universal.

The wide scope saxophonist reading arises from the same categories as follows. If Skolemization can act *after* reduction of the object, when the arbitrary object is within the scope of the universal, then it can also act *before*, when it is not in scope, to yield a Skolem constant, as in (63).

(63)

Every boy admires and every girl detests	some saxophonist
S/NP	$S\backslash(S/NP)$
$: \lambda x.and'(\forall y.boy'y \rightarrow admires'xy)(\forall z.girl'z \rightarrow detests'xz)$	$: \lambda q.q(arb'sax')$
	$\dots\dots\dots\dots\dots$
	$: \lambda q.q(sk'_{sax})$

$$S : and'(\forall y.boy'y \rightarrow admires'sk'_{sax}y)(\forall z.girl'z \rightarrow detests'sk'_{sax}z) \quad <$$

Since the resultant Logical Form is in all important respects model-theoretically equivalent to the one that would arise from a wide scope existential quantification, we can entirely eliminate the quantifier reading (51) for *some*, and regard it as bearing only the arbitrary object reading (60).[25]

Consistent with Geach's observation, these categories do not yield a reading in which the boys admire the same wide scope saxophonist but the girls detest possibly different ones. Nor do they yield one in which the girls also all detest the same saxophonist, but not necessarily the one the boys all admire. Both facts are necessary consequences of the monotonic nature of CCG as a theory of grammar, without any further assumptions of parallelism conditions.

In the case of the following scope-inverting relative of the Geach example, the outcome is subtly different.

(64) Some woman likes and some man detests every saxophonist.

The scope-inverting reading arises from the evaluation of the arbitrary woman and man *after* combination with *every saxophonist*, within the scope of the universal:

(65) $\forall x.saxophonist'x \rightarrow and'(like'x(sk'_{woman}x))(detest'x(sk'_{man}x))$

The reading where *some woman* and *some man* appear to have wider scope than *every saxophonist* arises from evaluation of (the interpretation of) the residue of right node raising, *some woman likes and some man detests*, before combination with the generalized quantifier *every saxophonist*. This results in two Skolem constants—here, sk'_{woman} and sk'_{man}—liking every saxophonist, again without the involvement of a true existential quantifier:

(66) $\forall x.saxophonist'x \rightarrow and'(like'x\, sk'_{woman})(detest'x\, sk'_{man})$

These readings are obviously correct. However, since Skolemization of the arbitrary man and woman has so far been assumed to be free to occur any time, it seems to be predicted that one arbitrary object might become a Skolem constant in advance of coordination and reduction with the object, while the other might do so after. This would give rise to further readings in which only one of *some man* or *some woman* takes wide scope—for example:

(67) $\forall x.saxophonist'x \rightarrow and'(like'x\ sk'_{woman})(detest'x(sk'_{man}x))$

The apparent nonavailability of such readings might again seem to call for transderivational parallelism constraints. Quite apart from the theoretically problematic nature of such constraints, they must be rather carefully formulated if they are not to exclude apparently legal conjunction of narrow scope existentials with explicitly referential NPs, as in the following:

(68) Some woman likes, and Fred detests, every saxophonist.

We will assume instead that the non-parallel mixed-scope reading (67) is in fact available, but is pragmatically disfavoured, on the following argument.

On the analysis of intonation contour and its interaction with coordinate structures given in Steedman 1991a and chapter 5 below, the coordinate fragments that result from right node raising like *Some woman likes and some man detests* must coincide with information structural units of the utterance, such as the "theme." Such information structural units carry presuppositions about contextually available alternatives that must hold or be accommodated for the utterance to be felicitous, and are evaluated as a whole in the course of derivation. In the present framework, readings like (67) can therefore be eliminated without parallelism constraints, by the further assumption that Skolem binding of arbitrary objects can only be done over complete information structural units—that is, entire themes, rhemes, or utterances. When any such unit is resolved in this way, *all* arbitrary objects concerned are obligatorily bound.[26]

Although the present account of indefinites might appear to mix derivation and evaluation in a dangerous way, this is in fact what we would expect from a monotonic semantics that supports the use of incremental semantic interpretation to guide parsing, as humans appear to (see Crain and Steedman 1985 and below).

Further support for a nonquantificational analysis of indefinites can be obtained from the observation that certain nominals that have been talked of as quantifiers entirely fail to exhibit scope alternations of the kind just discussed. One important class is the nonspecific or non-group-denoting counting quantifiers, including the upward-monotone, downward-monotone, and nonmonotone quantifiers (Barwise and Cooper 1981) such as *at least three*, *few*, *exactly five*, and *at most two* in examples like the following, which are of a kind discussed by Liu (1990), Stabler (1997), and Beghelli and Stowell (1997):

(69) a. Some linguist can program in at most two programming languages.
b. Most linguists speak at least three /few/exactly five languages.

In contrast to true quantifiers like *most* and *every*, these quantified NP objects appear not to be able to invert or take wide scope over their subjects. That is, unlike *Some linguist can program in every programming language*, which has a scope-inverting reading meaning that every programming language is known by some linguist, (69a) has no reading meaning that there are at most two programming languages that are known to any linguist, and (69b) cannot mean that there are at least three/few/exactly five languages for which there are distinct majority sets of linguist speakers, one set for each language.[27]

Beghelli and Stowell (1997) account for this behavior in terms of different landing sites (or in GB terms, functional projections) at the level of LF for the different types of quantifier. However, another alternative is to believe that in syntactic terms these NPs have the same category as any other but in semantic terms they are (plural) arbitrary objects rather than quantifiers, like *some, a few, six* and the like. This in turn means that they cannot engender dependency in the Skolem term subject arising from *some linguist* in (69a). As a result the sentence has a single meaning, to the effect that there is an arbitrary linguist who can program in at most two programming languages.

4.4.3 Binding Theory and Distributional Scope

The proposal that the nonspecific and counting so-called quantifiers aren't quantifiers at all does not explain how they induce the appearance of taking wide scope when they are subjects, in sentences like the following:

(70) a. Few/at most two/three boys ate a pizza.
b. Few/at most two/three farmers who own a donkey beat it.

There is every reason to doubt that the distributive reading of this sentence, according to which the boys ate different pizzas, arises from a quantified subject. Unlike the behavior of true quantifiers, distributivity is strictly subject to the same *c*-command condition as the binding conditions as defined in CCG terms in Steedman 1996b, p19. That is, only those terms that *c*-command another term at Logical Form can take bind or take scope over it. Thus the unavailability of scope inversion in (71a) is paralleled by the unacceptability of the reflexive in (71b):

(71) a. Some linguist knows at most two languages.
b. *Himself shaves Harry

Even more strikingly (since it is independent of derivational command, according to the treatment of binding in Steedman 1996b) the asymmetry in anaphor binding illustrated in (72) also shows up in scope alternation in (73). (That is, an indirect object can bind or take scope over a direct object a, but not vice versa b.)

(72) a. I showed the dogs themselves/each other.
b. *I showed themselves/each other the dogs.

(73) a. I showed three dogs some rabbit. *(ambiguous)*
b. I showed some dog three rabbits. *(unambiguous)*

It is therefore natural to follow Link (1983), van der Does (1992), and van den Berg (1996) in explaining the distributive behavior of plurals as arising from the Logical Form of verbs, in rather the same way as the behavior of reflexives. We will assume that, as well as having the normal translation (74a), and the "reflexivized" translation (74b) (adapted from Steedman (1996b), 21), all transitive verbs with plural agreement like *eat'* have a "distributivizing" category like (74c).

(74) a. eat a pizza := $S\backslash NP_{agr} : \lambda y.eat'(arb'pizza')y$
b. eat himself := $S\backslash NP_{3sm} : \lambda y.self'(eat'(ana'y))y$
c. eat a pizza := $S\backslash NP_{pl} : \lambda y.dist'(eat'(arb'pizza'))z$

Thus, subjects in examples like (70) can optionally distribute over the function that applies to them at the level of Logical Form, such as $eat'(arb'pizza)$, to yield not only standard forms like a, below, but also b:

(75) a. $eat'(arb'pizza')(arb'3boys')$
b. $dist'(eat'(arb'pizza'))(arb'3boys')$

If $dist'$ translates as (76), then an argument x can distribute over more oblique arguments.

(76) $\lambda f.\lambda x.\forall y \in x.fy$

For example:

(77) Three boys ate a pizza

$$NP^{\uparrow} : \lambda p.p(arb' 3boys') \quad (S\backslash NP)/NP : \lambda x.\lambda y.dist'(eat'x)z \quad NP^{\uparrow} : \lambda p.p(arb' pizza')$$

$$NP^{\uparrow} : \lambda p.p(sk'_{3boys})$$

$$S\backslash NP : \lambda y : \forall z \in y.eat'(arb' pizza')z \quad <$$

$$S : \forall z \in sk'_{3boys}.eat'(arb' pizza')z \quad <$$

$$S : \forall z \in sk'_{3boys}.eat'(sk'_{pizza}z)z$$

In invoking a "subordinated" use of universal quantification, this proposal resembles the treatment of distributive nonquantifiers in Kamp and Reyle 1993, 326-8. However, by tying it to the verb and its predicate-argument structure we explain the *c*-command condition on distribution, as evidenced by the fact that pizzas cannot distribute over boys in the following:

(78) A boy ate three pizzas.

4.4.4 Interaction of Word order and Quantifier Scope

We may assume (at least for English) that even the nonstandard constituents created by function composition in CCG cannot increase the number of quantifiable arguments for an operator beyond the limit of three or so imposed by the lexicon. It follows that the observation of Park (1995, 1996) that only quantified arguments of a single (possibly composed) function can freely alternate scope places an upper bound on the number of readings. The Logical Form of an n-quantifier sentence is a term with an operator of valency 1, 2 or 3, whose argument(s) must either be quantified expressions or terms with an operator of valency 1, 2 or 3, and so on. The number of readings for an n quantifier sentence is therefore bounded by the number of nodes in a single spanning tree with a branching factor b of up to three and n leaves. This number is given by a polynomial whose dominating term is $b^{log_b n}$—that is, it is linear in n, albeit with a rather large constant (since nodes correspond up to $3! = 6$ readings). For the relatively small n that we in practice need to cope with, this is still a lot of readings in the worst case.

However, the actual number of readings for real sentences will be very much lower, since it depends on how many true quantifiers are involved, and in exactly what configuration they occur. For example, the following three-quantifier sentence is predicted to have not $3! = 6$ but only 4 distinct readings, since the nonquantifiers *some dog* and *three rabbits* cannot alternate scope with each other independently of the truly quantificational dependency-inducing *Every boy*.[28]

(79) Every boy showed some dog three rabbits.

This is an important saving for the parser, as redundant syntactic analyses can be eliminated on the basis of identity of Logical Forms, a standard method of eliminating such "spurious ambiguities" (Karttunen 1989; Komagata 1999).

Similarly, as well as the restrictions that we have seen introduced by coordination, the SVO grammar of English means (for reasons discussed in Steedman 1996b) that embedded subjects in English are correctly predicted neither to extract nor take scope over their matrix subject in examples like the following:

(80) a. *a boy who(m) I know that admires John Coltrane
b. Somebody knows that every boy admires some saxophonist.

Cooper (1983) and Williams (1986, (100)) argue that sentences like the latter have no readings where *every boy* takes scope over *somebody*. This three-quantifier sentence therefore has not $3! = 6$, not $2! * 2! = 4$, but only $2! * 1 = 2$ readings. Since such embeddings are crucial to obtaining proliferating readings, it is likely that in practice the number of available readings is usually quite small.

To the extent that the availability of wide scope readings for the true quantifiers depends upon syntactic derivability, we may expect to find interactions of phenomena like scope inversion with word order variation across languages of the kind discussed by Bayer (1996) and Kayne (1998). In particular, the failure of English complement subjects to take scope over their matrix generalizes to a wider class of embedded arguments in verb-final complements in languages like German and Dutch.

I will return to this point in chapters 6 and 7 in part II.

4.4.5 Disambiguation and Underspecification

It is interesting to speculate finally on the relation of the above account of the available scope readings with proposals to minimize search during processing by building "underspecified" Logical Forms by Kempson and Cormack (1981), Alshawi and Crouch (1992), Reyle (1992), Poesio (1995), Asher and Fernando (1997), Joshi and Vijay-Shanker (1999) and Willis and Manandhar (1999). There is a sense in which arbitrary individuals are themselves underspecified quantifiers, which are disambiguated by Skolemization. However, under the present proposal, they are disambiguated during the derivation itself.

The alternative of building a single underspecified Logical Form can under some circumstances dramatically reduce the search space and increase efficiency of parsing. However, few studies of this kind seem to have looked at the problems posed by the restrictions on available readings exhibited by sentences like (55b). If they are to be disambiguated efficiently, then the disambiguated representations must embody or include those restrictions. However, the restriction that Geach noted seems intrinsically disjunctive, and hence appears to threaten efficiency in both parsing and disambiguating underspecified representations.

The fact that relatively few readings are available and that they are so tightly related to Surface Structure and derivation means that the technique of incremental semantic or probabilistic disambiguation of fully specified partial Logical Forms mentioned earlier may be a more efficient technique for computing the contextually relevant readings. For example, in processing (81) (adapted from Hobbs and Shieber 1987), which Park 1995 claims to have only four readings, rather than the five predicted by their account, such a system can build both readings for the S/NP *every representative of three companies saw* and decide which is more likely, before building both compatible readings of the whole sentence and similarly resolving with respect to statistical or contextual support:

(81) Every representative of three companies saw some sample.

This is only possible because of the strictly monotone relation between Logical Form and syntactic derivation.

The above observations imply that among the so-called quantifier determiners in English, only those that can engender dependency-inducing scope-inversion have interpretations corresponding to genuine quantifiers. The others are not quantificational at all, but are various types of arbitrary individuals translated as Skolem terms. These give the appearance of taking narrow scope when they are bound to truly quantified variables, and of taking wide scope when they are unbound and therefore "take scope everywhere." In addition the plural arbitrary individuals can distribute over or bind other arbitrary objects that they *c*-command at the level of Logical Form. Available readings can be computed monotonically from the combinatorics of syntactic derivation alone. The notion of syntactic derivation embodied in CCG is the most powerful limitation on the number of available readings, and allows all logical-form level constraints on scope orderings to be dispensed with, a result related to, but more powerful than, that of Pereira (1990), as extended in Shieber, Pereira and Dalrymple (1996).

It is interesting to note in this connection that Baker (1995), Bittner (1995), Bittner and Hale (1995), Faltz (1995), Jelinek (1995), and Damaso Vieira

(1995), together with other papers in Bach et al. (1995), show that the universal involvement of generalized quantifiers in explaining scope phenomena assumed by Barwise and Cooper (1981) is not easy to reconcile with the properties of various nonconfigurational, pronominal-argument, or agglutinative languages including Mohawk, Warlpiri, Navajo, Lakhota, Straits Salish, and Asurini. It seems quite likely that further examination will reveal quantification in such languages to be mediated by explicitly nonquantificational devices similar to those proposed here for English.

4.5 Summary: Surface Structure and Interpretation

According to the theory of grammar proposed here (as 36 and 37 and the preceding discussion reveal), surface derivation is less closely tied to predicate-argument structure than we are used to assuming. There are in general several alternative surface derivations for any given reading of a sentence, in some of which the object may structurally command the subject as in (82b) (or may even command a subject in a higher clause)

(82) a. Anna married Manny b. Anna married Manny

At the same time, at the level of the interpretation all these derivations yield the same Logical Form, in which function-argument relations of dominance and command over subjects and other elements hold in pretty much their traditional form.

The proliferation of surface derivations in CCGs creates problems for the processor (to which I will return in part III), because it compounds the already grave problems of local and global ambiguity in parsing by introducing numerous semantically equivalent potential derivations for each reading. This has been referred to as the problem of "spurious" ambiguity by Wittenburg (1986). Although it clearly does not matter which member of any equivalence class the parser finds, the parser must find *some* member of *every* semantically distinct class of analyses. The danger is that the entire forest of possible analyses will have to be examined in order to ensure that all such analyses have been found. Since the number of distinct derivations (in the sense of distinct sequences of rule applications leading to a derivation) can grow as the Catalan function of the length of the sentence, the problem is potentially serious.

Nevertheless, this problem has been overstated. Standard PS grammars give rise to a similar proliferation of derivations, in the sense that for any syntactic structure there are many alternative orders in which rules can be applied in a derivation to yield the same tree. This fact tends to be forgotten, because of the isomorphism of trees representing syntactic structures and those representing derivations, but efficient parsing algorithms have to deal with this problem. They do so either by using a "normal form" algorithm (e.g. Earley's) that is guaranteed to find only one derivation per tree, or by making sure (as in the CKY parser discussed in chapter 9) that only one copy of each subtree is kept. These techniques will be discussed in chapter 9, where we will see that a variant of the latter technique originally proposed by Karttunen (1989) can be directly applied to CCG derivations.

In fact, the term "spurious ambiguity" is distinctly misleading. Far from being spurious this nondeterminism is simply a property of all natural languages. Any theory that captures the range of phenomena discussed here, in particular in respect of coordination, will necessarily encounter the same nondeterminism. Given the degree of ambiguity that is tolerated elsewhere in the language, it is not even particularly surprising to find that there is a bit more of it from this source.

To say this is not to deny that nondeterminism (of all kinds) remains a problem for processors. It is simply to deny that this particular nondeterminism indicates anything wrong in the combinatory categorial competence theory.

The architecture of this theory can be represented to a first approximation as in figure 4.1, which is a version of the transformationalists Y-diagram seen in figure 2.1 of 2, using the two derivations of one of the two interpretations of the sentence *Every boy admires some girl* as an example.

In the generative direction, according to this theory, derivations can be regarded as projected by the combinatory rules from the lexicon, which also pairs types or categories with interpretations or Logical Forms.[29]

In the analytic direction, combinatory derivations map Phonetic Form directly onto constituents, with a category consisting of a syntactic type and a Logical Form. The Logical Form associated with the S category that is monotonically specified as a result of the derivation is a quantified predicate-argument structure, which can be thought of as an unordered tree representing traditional dominance/command relations. We will see in chapter 5 that the interpretations of the immediate derivational antecedents of the root S node can be regarded as the elements of Information Structure and as identifying the content of topic/theme and comment/rheme. The corresponding structural

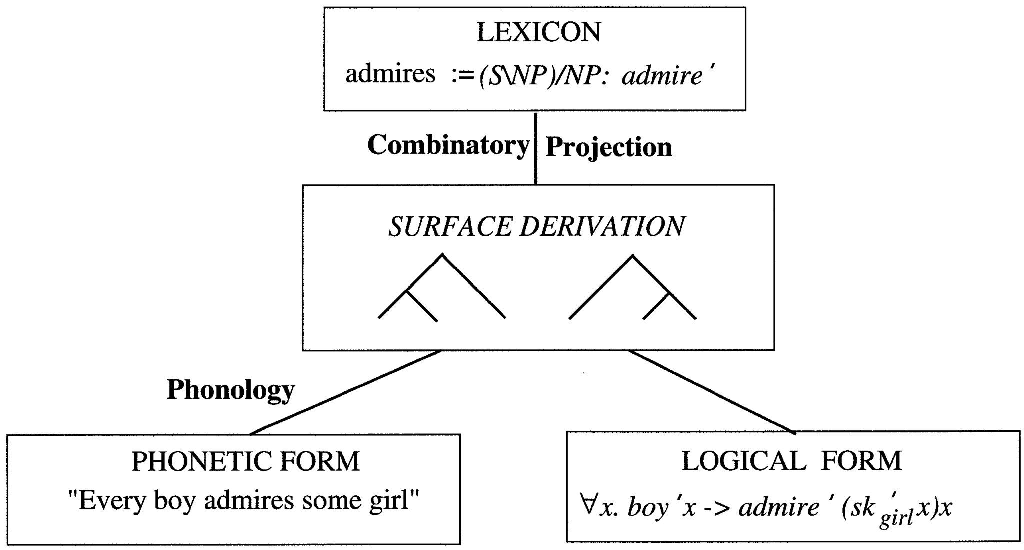


Figure 4.1
Architecture of a Combinatory Grammar, I

units directly coincide with phrasal intonational boundaries, where these are present, justifying the earlier identification of Surface Structure with Phonological Form. However, Surface Structure is strictly a record of the process of deriving such Logical Forms via combinatory operations that are type-driven, rather than structure-dependent. Surface Structure is therefore not a grammatical "level of representation" at all. To that extent, the theory is not only monotonic in the sense of never revising the structures it builds. It is also monostratal, in the sense that it builds only a single level of structure, namely, Logical Form.

BLANK PAGE

Chapter 5

Structure and Intonation

Take care of the sense, and the sounds will take care of themselves.
Lewis Carroll, *Alice in Wonderland*

In chapter 2, (13), we considered the following exchange, in which intonation imposes a perceptual grouping of words in the spoken utterance into fragments that are inconsistent with traditional linguistic notions of syntactic constituency.[1]

(1) a. I know that Alice married Alan. But who did ANNA marry?
b. (ANNA married) (MANNY).

The prosody informally indicated in (1b) by capitals (for stress) and parentheses (for intonational phrase boundaries) is one possibility (among others that we will come to later) for an answer to the question (1a). It consists in not only marking the focused information in the answer by the use of high pitch on the stressed first syllable of the word *MANNY*, but also in stressing the first syllable of *Anna*, using a high pitch-accent, and placing a final rise at the end of *married*, with lower pitch interpolated in between. This utterance might give rise to a pitch contour something like the sketch in figure 5.1, which is an idealized version of the actual pitch track shown in 5.2. This contour conveys the contrast between the previous topic concerning Alice's marriage and the new one concerning Anna's, and it imposes the perceptual grouping indicated by the brackets. Such a grouping cuts across the traditional syntactic analysis of the sentence as a subject and a predicate VP.

Many authorities, such as Chomsky (1971), Jackendoff (1972), Cooper and Paccia-Cooper (1980), Gee and Grosjean (1983), Kaisse (1985), and Croft (1995) have continued to argue, nevertheless, that intonation can be driven directly from Surface Structure. There is an immediate intuitive appeal to the idea, as noted in chapter 2, for it is hard to see why intonation should depart from the Constituent Condition on Rules in any language. However, the apparent complexities engendered by examples like (1) have led many others, such

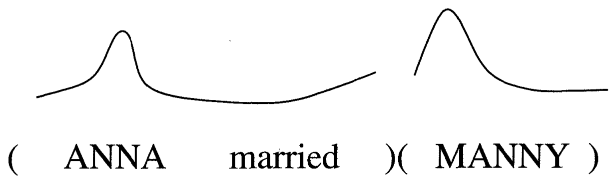


Figure 5.1
Idealized pitch contour for (1b)

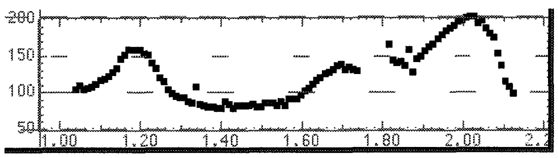


Figure 5.2
Actual pitch contour for (1b)

as Liberman (1975), Goldsmith (1976), Pierrehumbert (1980), Selkirk (1984), and Nespor and Vogel (1986), to postulate an autonomous level of "Intonation Structure" independent of Surface Structure and related only indirectly to Logical Form or function/argument structure, via Information Structure.

However compelling the logic of this argument may appear, we noted in chapter 2 that the involvement of two apparently autonomous levels of structure, related to two autonomous levels of meaning representation, complicates the theory considerably. The picture becomes even bleaker when it is realized that the two levels of structure must communicate, because of the presence of certain focusing constructions and operators, such as the English topicalization construction or the focusing particle *only*, the latter exemplified in the following sentence:

(2) John introduced only BILL to Sue.

Such constructions and particles, which have been discussed by Rooth (1985), von Stechow (1991), Hoeksema and Zwarts (1991), and others, have effects in both domains. These observations have seemed to demand the quite complex theoretical architecture shown in figure 2.3. Such a theoretical architecture offers a view of sentence structure as having an "autosegmental" topology which Halle influentially likened to that of a spiral-bound notebook (cf. Halle and Vergnaud 1987 p78-79). This notebook has phonetic segments arranged along the spine, and different autonomous levels of structure—prosodic, syntactic, and others—written on different leaves, each of which may refer to descriptions on other pages. As Zwicky and Pullum (1987, 4) have pointed out, such theories are potentially very unconstrained, in the absence of a principled statement about which of the pages may cross-refer, and why. The simplest possible constraint upon such a theory would be a demonstration that certain communicating levels involve isomorphic structural descriptions, for those levels at least could be combined upon a single page of the notebook.

However, a strong hint that a simplification might be possible seems to be provided by the observation that Intonation Structure is, despite its apparent partial independence from syntax, nonetheless constrained by meaning, and in particular by distinctions of focus, information, and propositional attitude toward concepts and entities in the discourse model. The intonation contour in the response in (1) seems to divide the utterance into a topic or theme to do with *Anna marrying*, and a comment or rheme *Manny*.[2] These terms will be defined formally below, but informally the theme can be thought of as denoting what the speaker assumes to be *the question under discussion* and the rheme can be thought of as what the speaker believes to be *the part of the utterance that advances the discussion*. Even in advance of a more formal definition, it will be convenient to refer to such partitions of the information in the proposition as the "Information Structure" of an utterance.

A theme in the present sense of the term can be more concretely exemplified as *that which is introduced into the discourse context by a* wh-*question*. By now we are familiar with the idea that such an entity can be expressed as a functional abstraction, as Jackendoff (1972) and Sag (1976) point out, equivalent in this case to the following λ-term:[3]

(3) $\lambda x.marry'x\ anna'$

Establishing a theme with content (3) in the context via a *wh*-question such as *Who did Anna marry?* is one way to make the intonation contour in (1) felicitous. (Of course, it is not claimed that an explicit mention, via a *wh-question* or otherwise, is necessary for interpreting the response. Nor does this *wh*-question uniquely determine this response—for example, for reasons that we will come to later, it is possible to answer the same question with the "fall-rise" contour confined to *Anna* and the boundary before the VP *married Manny*. There is also no claim that intonation contours in general determine Information Structure unambiguously. We return to these points in section 5.5.2 below.)

The close relation in English of Intonation Structure to Information Structure, first proposed by Halliday (1967a), has recently been endorsed by Selkirk

(1984, 284) as "The Sense Unit Condition" on intonational constituency, which says in essence that intonational constituents must have coherent translations at Information Structure.

However, we have seen in previous chapters that natural languages include a number of other constructions whose semantics is also reminiscent of functional abstraction. The most obvious and theoretically tractable class are Wh-constructions, in which many of the same fragments that can be delineated by a single intonation contour appear as the residue of the subordinate clause. Another and much more diverse class are the fragments resulting from the coordinate constructions discussed in previous chapters. The latter constructions are doubly interesting, because they and certain other sentence-fragmenting constructions, such as parentheticals, interact very strongly with intonation, on occasion making intonation breaks obligatory, rather than optional, as Downing (1970) and Bing (1979), among others, have noted. For example, the intonation indicated on the following ambiguous sentence forces one syntactic analysis with an absurd reading, and leaves the sensible analysis quite inaccessible (the example is from Pierrehumbert (1980), who attributes it to Mark Liberman):

(4) *(Harry likes the NUTS) (and bolts APPROACH).

It is therefore tempting to think that the nonstandard concept of Surface Structure and constituency that has been developed in earlier chapters in order to explain coordination and unbounded dependency might directly provide the notion of structure that is required to account for intonational prosody. If this conclusion is correct, then both the camps identified earlier are in a sense correct. Intonation can indeed be specified directly from surface syntactic structure, without the mediation of an autonomous Intonation Structure. However, the syntactic structure in question corresponds to Information Structure rather than traditional Surface Structure, and hence directly subsumes the Intonation Structure of English.

5.1 Surface Structure and Intonation Structure

According to the combinatory theory, conjoinable strings like *Anna married* and even *a policeman a flower* correspond to constituents in their own right, without syntactic operations of deletion or movement. It follows that they must also be possible constituents of simple noncoordinate sentences like *Give a policeman a flower* and (1b), *Anna married Manny*, as well. Such sentences have several Surface Structures, corresponding to different sequences of composition and application. As we have seen, the nonstandard derivation (5) is allowed for the latter sentence, as well as the traditional derivation (6):

(5) Anna married Manny

$$NP : anna' \quad (S\backslash NP)/NP : marry' \quad NP^{\uparrow} : manny'$$

$$\underline{\qquad\qquad}{>}\mathbf{T}$$

$$S/(S\backslash NP) : \lambda f.f\ anna'$$

$$\underline{\qquad\qquad\qquad\qquad}{>}\mathbf{B}$$

$$S/NP : \lambda x.marry'x\ anna'$$

$$\underline{\qquad\qquad\qquad\qquad\qquad\qquad}{<}$$

$$S : marry'manny'anna'$$

(6) Anna married Manny

$$NP : anna' \quad (S\backslash NP)/NP : marry' \quad NP^{\uparrow} : manny'$$

$$\underline{\qquad\qquad}{>}\mathbf{T}$$

$$S/(S\backslash NP) : \lambda f.f\ anna'$$

$$\underline{\qquad\qquad\qquad\qquad}{<}$$

$$S\backslash NP : \lambda y.marry'manny'y$$

$$\underline{\qquad\qquad\qquad\qquad\qquad\qquad}{>}$$

$$S : marry'manny'anna'$$

Such families of derivations form equivalence classes, for the semantics of the combinatory rules guarantees that all such derivations will deliver an interpretation determining the same function-argument relations—in this case, *marry'manny'anna'*. Moreover, the interpretation of the nonstandard constituent *Anna married* of type S/NP bears an interpretation equivalent to the abstraction (3).

It is therefore tempting to believe that these semantically equivalent derivations convey distinctions of discourse information and that they are on occasion distinguished by intonational markers in spoken language.[4]

For example, the following bracketings correspond to alternative CCG Surface Structures, arising out of different sequences of compositions and applications, each of which corresponds directly to a possible intonation contour:

(7) a. (I)(want to begin to try to write a play).
b. (I want)(to begin to try to write a play).
c. (I want to begin)(to try to write a play).
d. (I want to begin to try)(to write a play).
e. (I want to begin to try to write)(a play).

The leftmost element is in every case a fragment that can be coordinated. For example:

(8) *I wanted, and you expected,* to write a play.

Conversely, the following are at least as strange (and pragmatically demanding) as coordinations as they are as intonational phrases:

(9) a. ?(I want to BEGIN to), (try to write a PLAY).
b. ?I wanted to, and you actually expected to, try to write a play.

(Examples like (7) and (9a) are used by Selkirk (1984, 294) to motivate a definition of the Sense Unit Condition in terms of a relation over the heads of constituents.) A stronger example emerges from comparing the following examples, in which the string *three mathematicians* is as hard to make an intonational phrase as it is to coordinate. (The unacceptability of (10a) is also used by Selkirk as evidence for the Sense Unit Condition.)[5]

(10) a. ?(Three MATHEMATICIANS) (in ten prefer MARGARINE).
b. ?Three mathematicians, in ten prefer margarine, and in a hundred can cook a passable soufflé.

It is irrelevant to the present purpose to ask *how* sentences like (10b) might be excluded, or even to ask whether what is wrong with them is a matter of syntax, semantics, or pragmatics. The important point for present purposes is that the *same* constraint applies in syntactic and prosodic domains. That is, the Sense Unit Condition on prosodic constituents simply boils down to the Constituent Condition on Rules of grammar. This result is a very reasonable one, for what else but a constituent could we expect to be subject to the requirement of being a semantic unit?

It follows that we predict the strongest possible conspiracy between prosodic constituency and coordinate structure. Noncoordinate sentences typically have many equivalent combinatory derivations, because composition is optional and associative. These analyses can give rise to many different intonation contours. On the other hand, coordinate sentences, like relative clauses, have fewer equivalent analyses, because only analyses that make the conjuncts into constituents are allowed. Two predictions follow. First, we must expect that any substring that can constitute a prosodic constituent will also be able to coordinate. Second, of all the intonational tunes that distinguish alternative prosodic constituencies in noncoordinate sentences, we predict that only the ones that are consistent with the constituents demanded by the coordination rule will be allowed in coordinate sentences. Intonation contours that are appropriate to the alternative constituencies are syntactically ruled out. So for example, there are many prosodic constituencies for the example (7), *I want to begin to try to write a play*, realized by a variety of intonational contours. However, there are many fewer possible intonation contours for the following coordinate sentence, and they seem intuitively to be closely related to the ones that impose the corresponding bracketing (7e) in the simpler sentence:

(11) I want to begin to try to write, and you hope to produce, a musical based on the life of Sir Stafford Cripps.

Observations like the above make it seem likely intonation often determines which of the many possible bracketings permitted by the combinatory syntax of English is intended, and that the interpretations of the constituents are related to distinctions of information-structural significance among the concepts that the speaker has in mind. Thus, whatever problems for parsing written text arise from the profusion of equivalent alternative Surface Structures engendered by this theory, these "spurious" ambiguities seem to be to some extent resolved by prosody in spoken language. The theory therefore offers the possibility that prosody and syntax are one system, and that speech processing and parsing can be unified in a single process.[6]

This section and the next show that the combinatory rules of syntax that have been proposed in order to explain coordination and unbounded dependency in English do indeed induce Surface Structures that subsume the structures that have been proposed by Selkirk (1984) and others in order to explain the possible intonation contours for all sentences of English. The proof of this claim depends upon two results. First, it must be shown that the rules of combinatory grammar can be made sensitive to intonation contour, so as to limit the permissible derivations for spoken sentences like (1b). Second, it must be shown that the interpretations of the principal constituents of these derivations correspond to the Information Structure established by the context to which they are appropriate, such as (1a).

5.2 Two Intonation Contours and Their Functions

I will use a notation for intonation contours that is based on the theory of Pierrehumbert (1980), itself a development of proposals by Bruce (1977), Liberman (1975), and Goldsmith (1976). The version used here is roughly as presented in Selkirk 1984, Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988, and Pierrehumbert and Hirschberg 1990, although it will become clear below that I have extended Pierrehumbert's notation in a couple of minor respects. I have tried as far as possible to take my examples and the associated intonational annotations from those authors.

The advantage of this theory is that it specifies intonation contour independently of the string, in terms of just two kinds of fixed points or "tones." The contour between tones can be determined by interpolation. The first group of tones are the "pitch-accents," which are the substantial local pitch maxima and minima that coincide with perceived contrastive emphasis. The other group of tones are the "boundaries," that mark the right-hand edge of a prosodic phrase. I follow Pierrehumbert in assuming that intonation contours can be described in terms of two abstract or relative pitch levels, H and L, denoting high or low abstract pitch.

Of Pierrehumbert's six pitch accent tones, I will consider only two here, H* and L+H*.[7] The phonetic or acoustic realization of pitch accents is a complex matter. Roughly speaking, the L+H* pitch accent that is extensively discussed below in the context of the L+H* LH% melody generally appears as a maximum that is preceded by a distinctive low level and peaks later than the corresponding H* pitch accent when the same sequence is spoken with the H* LL% melody. (See Silverman 1988 for discussion. Nothing in the combinatory theory hinges on the precise identities of the pitch accent types. All that matters is that the two complete melodies are distinct, a matter on which all theories agree.)

The intonational constituents of interest here are made up of one or more pitch accents (possibly preceded by other material), followed by a boundary. In recent versions of the theory, Pierrehumbert and her colleagues distinguish two distinct levels of prosodic phrase: the intonational phrase proper and the "intermediate phrase." The intermediate phrase boundary is a bare phrasal tone, either L or H.[8] Intonational phrase boundaries are L or H phrasal tones plus an boundary tone written H% or L%. We will principally be concerned here with the intonational phrase boundaries that are written LH%, and the boundary L or LL%.

The intermediate phrase is distinguished in Pierrehumbert's theory as defining the domain of a phenomenon known as "downstep." If more than one pitch accent occurs without an intervening boundary—that is, within an intermediate phrase—then the entire pitch range of each successive pitch accent is shifted downward from its predecessor. At the intermediate phrase boundary (and therefore at any higher-level boundary including the intonational phrase boundary), the pitch levels are reset to the normal level. Although this aspect of the pitch contour is completely rule-governed, so that in Pierrehumbert's own notation, downstepped pitch accents are not distinguished, it is sometimes useful to include such a notation. On such occasions will use "!" as a prefix to the pitch accent type, a notation that originates with Ladd (1980) and has been included in the ToBI conventions, writing such sequences as follows:

(12) blueberries, bayberries, raspberries, mulberries, and brambleberries
H* !H* !H* !H* !H* LL%

I have followed Beckman and Pierrehumbert in regarding boundaries of all kinds as confined to the right-hand end of the prosodic phrase. However, the position and nature of the phrasal tone is one of the more controversial details of the theory (Pierrehumbert and Beckman 1988, 236–237). The influence of, say, an L or LL% boundary on a preceding H* pitch accent is apparent immediately after the maximum, no matter how distant the right-hand boundary is. Pierrehumbert and Beckman point out that this influence may be apparent by the end of the word bearing the pitch accent preceding the boundary. Indeed, in the framework of the British School, the event corresponding to the phrasal tone component L or H of the boundary is considered to be part of the pitch accent, rather than part of the boundary event.

For all other regions of the prosodic phrase, notably the region before the (first) pitch accent, the regions between pitch accents, and the region between pitch accent and boundary, the pitch contour is merely interpolated. In Pierrehumbert's notation, such substrings therefore bear no indication of abstract tone whatsoever. It is sometimes convenient to regard such elements as bearing a "null" tone.[9]

Thus, according to this theory, the shape of a given pitch accent in a prosodic phrase, and of its phrase accent and the associated right-hand boundary, is essentially invariant. If the constituent is very short—say, a monosyllabic NP—then the whole intonation contour may be packed into that one syllable. If the constituent is longer, then the pitch accent will appear further to the left of the phrasal tone and boundary tone at the right-hand edge. The intervening pitch contour will merely be interpolated, as will any part of the contour preceding the pitch accent(s). In this way, the same tune can be spread over longer or shorter strings, in order to mark the corresponding constituents for the particular distinction of information and propositional attitude that the melody denotes.

Consider the prosody of the sentence *Anna married Manny* in the following pair of discourse settings, which are adapted from Jackendoff (1972, 260) and Steedman (1991a). To aid the exposition, words bearing pitch accents are printed in capitals, and prosodic phrase boundaries are explicitly marked in the sentences, using parentheses. (These devices are not part of Pierrehumbert's notation.)

(13) Q: Well, what about MANNY? Who married HIM?
A: (ANNA) (married MANNY).
H* L L+H* LH%

(14) Q: Well, what about ANNA? Who did SHE marry?
A: (ANNA married) (MANNY).
L+H* LH% H* LL%

In these contexts the stressed syllables on both *Anna* and *Manny* receive a pitch accent, but a different one. In answer (13A) there is a prosodic phrase on *Anna* made up of the sharply rising pitch accent that Pierrehumbert calls H*, immediately followed by an L boundary, perceived as a rapid fall to low pitch. There is another prosodic phrase having the somewhat later- and lower-rising pitch accent called L+H* on *Manny*, preceded by null tone (and therefore interpolated low pitch) on the word *married*, and immediately followed by a rising "continuation" boundary, written LH%. (See Pierrehumbert and Hirschberg 1990, (33), for discussion of a similar example.)[10] In answer (14A) the two tunes are reversed (see figures 5.1 and 5.2): this time the tune with pitch accent L+H* and boundary LH% is spread across a prosodic phrase *Anna married*, and the other tune with pitch accent H* and boundary LL% is carried by the prosodic phrase *Manny*, again starting with an interpolated or null tone.[11]

The intuition that there is some systematic distinction in meaning between these tunes seems to be very compelling, though it has in the past proved hard to formalize. The tunes have been associated with such factors as social attitude (O'Connor and Arnold 1961; Merin 1983; Bartels 1997), illocutionary acts (Liberman and Sag 1974; Sag and Liberman 1975; Liberman 1975), propositional attitudes (Ward and Hirschberg 1985), maintenance of mutual belief (Pierrehumbert and Hirschberg 1990), and Information Structure (Halliday 1967a; Jackendoff 1972; Schmerling 1976; Ladd 1980; Gussenhoven 1983; Selkirk 1984; Terken 1984, Cormack 1992, Terken and Hirschberg 1994; Morel 1995; Rochemont 1986, 1998; Rochemont and Culicover 1990; Steedman 1991a,b; Zubizarreta 1998).

The present chapter concentrates on certain aspects of intonation that primarily concern Information Structure, in the sense of that term proposed by Vallduví (1990), and Steedman (1991a), although these proposals differ in detail (see Vallduví and Engdahl (1996) for a survey). These theorists follow Halliday (1967b, 1970) in assuming that there are two independent dimensions to

Information Structure that are relevant to intonation. The first corresponds to the distinction, informally introduced at the start of the chapter, between theme and rheme. In English we will see that this dimension of Information Structure determines the overall shape of the intonational tune or tunes imposed upon an utterance. The second dimension is one of salience or contrast. In English this dimension is reflected in the position of pitch accents on particular words. The presence of a pitch accent of any shape is generally agreed to assign salience or contrast independently of the particular shape or contour of the pitch accent or overall phrasal melody (see Pierrehumbert and Hirschberg 1990, 288-289). The next sections consider these two dimensions in turn.

5.3 Theme and Rheme

The λ-abstraction operator is closely related to the existential quantifier $\exists$. It is therefore natural to associate the notion of theme with the set of propositions among all those supported by the conversational context that could possibly satisfy the corresponding existential proposition. In the case of the exchange in (14) it is the following, in which $\Diamond$ indicates possibility:

(15) $\exists x.\Diamond marry' x\ anna'$

This might be a set like the following:

(16) $$\left\{\begin{array}{c} \Diamond marry'\,alan'\,anna' \\ \Diamond marry'\,fred'\,anna' \\ \Diamond marry'\,manny'\,anna' \\ \ldots \end{array}\right\}$$

This extensional interpretation of the notion theme resembles the Alternative Semantics approach to presupposition and focus of Karttunen and Peters (1979), Rooth (1985, 1992), and Kratzer (1991), and the related analysis of German intonational meaning of Büring (1995, 1997). Specifically, the alternative set in question is the one that Rooth and Büring call C, the "contextual" alternative set. Since all alternative sets are contextual, I will refer to it here as the "rheme" alternative set.

Alternative sets are of course in many cases not exhaustively known to hearers, and in practice one would want to compute with something more like the quantified expression (15) or the λ-term itself, as in the structured-meanings approach of Cresswell (1985) and others. However, alternative sets are easy to grasp and are used here for reasons of exposition.

In semantic terms the theme and rheme can therefore be characterized as follows:

(17) a. The Theme *presupposes* the rheme alternative set.
b. The Rheme *restricts* the rheme alternative set.

The sense in which a theme "presupposes" the rheme alternative set is much the same as that in which a definite expression presupposes the existence of its referent. That is to say, there is a pragmatic presupposition that the relevant alternative set is available in the contextual "mental model" (Johnson-Laird 1983) or database. The presupposition may be "accommodated" in the sense of Lewis (1979)—that is, be added by the hearer after the fact of utterance to a contextual model that is consistent with it.

5.3.1 Update Semantics for Theme and Rheme

One way of making such referents available is to think of the theme of an utterance as *updating* or having side-effects on the context or discourse model.[12] Following Jacobs (1991) and Krifka (1991), it can be characterized as in general causing one or more existing referents or "facts" such as ($\theta' \lambda x.marry' x\ anna'$), where θ marks the λ-term as a theme, to be *retracted* or removed from the context model, and causing a new thematic referent or fact to be *asserted* or added. If the theme is unmarked by any accent, then it will simply be the corresponding thematic referent that is retracted and asserted. Unless a fact of the appropriate form is already present in (or is at least consistent with) the context, the first of these effects will cause the discourse to fail. Otherwise, the thematic referent will be reasserted.

The rheme should also be thought of as updating the context with a similar type of referent, which may become the theme of a subsequent utterance. However, the rheme does not require a preexisting referent or cause any existing thematic referents to be retracted (although we will see that it may have other effects on the database, via the entailments and implicatures discussed above).

The exact form of the retracted and/or asserted informational referents in all of the above examples depends upon the location of focus and pitch accents in the utterance and is determined in a manner discussed in section 5.4.

Noncompositional, procedural notions like assertion and retraction must eventually be declarativized, if we are to be able to prove anything about the expressive power of this theory. However, procedural descriptions can be very transparent, probably because they remain very close to what is actually going on in our heads, and for the moment it will be helpful to think of the problem in these terms.[13]

The claim that the L+H* LH% tune when present marks the theme in one or the other of these closely related senses is implicit in the accounts of Jackendoff (1972), Ladd (1980), and others, but it remains controversial. Pierrehumbert and Hirschberg (1990, 294–297) propose a compositional semantics for intonational tunes that is based on scalar values on dimensions of propositional attitude such as certainty concerning relevance and degree of commitment to belief revision. According to their account, the L+H* pitch accent is used "to convey that the accented item—and not some alternative related item—should be mutually believed" (p. 296).

As an example, Pierrehumbert and Hirschberg discuss the following dialogue (adapted from Jackendoff (1972, 258–265), and also discussed by Ladd (1980, 157–159), and Steedman (1991a)), which is isomorphic to (13):

(18) Q: What about the beans? Who ate THEM?
A: Fred ate the BEANS.
H* L L+H*LH%

In support of their claim that the L+H* pitch accent evokes a set of alternatives besides the accented item, they correctly observe that the utterance implicates the possibility that other people may have eaten other things. However, this particular alternative set has already been introduced into the context by the question, and in the absence of such a question (or some other utterance establishing a context that supports or is at least consistent with this theme), the intonation contour is inappropriate. The example therefore does not exclude the possibility that the L+H* LH% tune evokes this set of alternatives by marking a part of the theme.

The following minimal pair of dialogues will be helpful in deciding between these claims, because it appears at first glance to raise problems for both.

(19) Q: Does Mary like corduroy?
A: Mary likes BOMBAZINE.
H* LL%

(20) Q: Does Mary like corduroy?
A: Mary likes BOMBAZINE.
L+H* LH%

In (19), the entire response is marked with the H*LL% tune that we have identified as marking the rheme, constituting what the speaker believes the hearer needs. Depending on the context, the speaker may thereby be committed by the usual Gricean principles to a number of conversational implicatures. For example, if liking bombazine entails hating corduroy, then this response implicates denial. If on the other hand liking bombazine entails liking corduroy, then the response implicates affirmation. Either way, the speaker's intonation commits them only to the claim that *bombazine* is the rheme—that is, that it restricts the set of alternatives to just one—rather than to a particular change in belief.

More specifically in both cases, the rheme of the yes-no question adds a theme *theme*′(*like*′*corduroy*′*mary*′) to the facts making up the respondent's context. (The alternative set here is confined to the proposition and its negation.) The respondent then constructs the corresponding query and evaluates it with respect to the context. If the query immediately succeeds, or fails altogether, then it is appropriate to respond with a direct yes or no. If the query succeeds but a step of inference involving the respondent's rule that *Everyone who likes bombazine likes corduroy* and the respondent's knowledge that *Mary likes bombazine* is needed to establish the answer, then one of the following cases may apply. If the respondent's discourse model implies that the questioner knows neither the rule nor the truth of the premise, then the respondent should state them both, as in the extended example (23). On the other hand, if the discourse model implies that the respondent knows the rule, but not the premise, then the response should be either as in (19) or as in (20). If there is reason to believe that this is the only relevant difference between the questioner's knowledge and the respondent's own, then stating the premise as a rheme, as in (19), is appropriate, since the respondent can sincerely claim that it is everything the questioner needs. But if the respondent has reason to suspect that there may be other differences and therefore cannot sincerely claim that the questioner can make this inference, then the respondent should mark the premise as a theme, as in (20), and leave the questioner to derive the rheme or not, as the case may be.

As is often the case, the respondent may for reasons of politeness or other pragmatic footwork use an utterance of an isolated theme to conversationally implicate lack of willingness to commit to the adequacy of their information, simultaneously being perfectly certain of the outcome. Nearly all speech acts like the response in (20) have the smell of indirection about them, and we should not expect to capture them in terms of literal meaning alone.

Example (20), which is of a kind discussed by Jackendoff (1972, 264–265), Liberman and Sag (1974), and Ladd (1980, 146–148), appears at first glance to be almost equivalent. In particular, the possibilities for conversational implicature of either affirmation or denial seem identical. Any difference seems to lie in the degree of conviction that the utterance constitutes an answer to the question.

Since in other respects the two utterances seem similar, there is a temptation to believe that the L+H* LH% tune in this case might mark a rheme, rather than a theme, differing from the standard H*LL% rheme tune in terms of the degree of commitment to whether it does in fact restrict the set of rheme alternatives sufficiently.

However, it is also possible to believe that this utterance is in fact a theme and that what the respondent has actually done is to offer a new set of alternatives, without stating a rheme at all, leaving the other party to supply the rheme. This would certainly explain the lack of commitment to whether the utterance restricts the rheme alternative set, since that is exactly what a theme does not do. It is also likely that the effect of not taking responsibility for a rheme in this utterance will be that of *conversationally implicating* a lack of confidence in either the relevance of the theme or the certainty of the inference that might be drawn. But that would not be a matter of literal or conventional meaning of the utterance or the intonation contour itself.

This is essentially the analysis proposed by Ladd (1980, 153-156), who relates "fall-rise" contours to the function of evoking a set of alternatives established by the preceding context—a notion I have identified with the notion of theme and have interpreted above in terms of the Alternative Semantics of Karttunen and Peters (1979) and Rooth (1985).

In the case of answer (20A) the new theme is simply the following:

(21) $like'bombazine'mary'$

Since this is a fully saturated proposition, with no λ-bound variables, the corresponding rheme alternative set is a singleton:

(22) $\{like'bombazine'mary'\}$

Since it contains only one member, it also entails an answer to the question via exactly the same chain of inference from shared beliefs as (19).

Further support for the claim that the L+H* LH% tune marks theme in (20), establishing a new set of alternatives, and that any effect of lack of commitment arises by conversational implicature, can be found in the fact that this intonation remains appropriate when the step of inference that generates the rheme itself is explicitly spelled out, as in the following deliberately exaggerated example, in which *liking bombazine* is necessarily distinct from the rheme:

(23) Q: Does Mary like corduroy?

A: Well, she likes BOMBAZINE,
L+H* LH%
And people who like BOMBAZINE like CORDUROY.
L+H* LH% H*LL%
So I am sure that Mary likes CORDUROY.
H*LL%

(Note that *likes bombazine* in the first conjunct could equally well be uttered with an H*L% rheme accent, but in the second it really must be marked as a theme. Under most circumstances the first and third conjuncts could be omitted entirely, as being implicated by the second.)

Still more evidence for the claim that L+H* LH% invariably marks a theme can be found in the fact that when a similarly implicative reply states a law from which the conclusion necessarily follows, so that there is no plurality of alternatives, then only the rheme tune is felicitous, as in the following minimal pair:

(24) Q: Will Mary come to the meeting on time?

A: Mary is ALWAYS on time.
H* LL%

(25) Q: Will Mary come to the meeting on time?

A: #Mary is ALWAYS on time.
L+H* LH%

Conversely, if the content of the response necessarily implicates a plurality of alternatives, then the position is reversed: only the theme tune is felicitous:

(26) Q: Will Mary come to the meeting on time?

A: Mary is USUALLY on time.
L+H* LH%

(27) Q: Will Mary come to the meeting on time?

A: #Mary is USUALLY on time.
H* LL%

It is clear that the L+H*LH% tune in the latter example indicates no lack of commitment to the probabilistic claim about Mary's punctuality; rather, it indicates a lack of commitment to the adequacy of this information to fully answer the question.

5.3.2 Unmarked Themes

Extravagantly informative intonation contours like those in (13) and (14) are the exception. It is only appropriate to mark the theme with an L+H* pitch accent when it stands in contrast to a preceding different theme. If the rheme alternative set that a theme presupposes is unambiguously established in the context, it is common to find that the theme is deaccented throughout—in Pierrehumbert's terms, without any pitch accent or obvious boundary, as in the following exchange:

(28) Q: Who did Anna marry?
A: (Anna married) (MANNY).
H* LL%

We would be missing an important semantic generalization if we failed to note that examples (14) and (28) are identical in Information Structure as far as the theme-rheme division goes. We will therefore need to distinguish the "marked" theme in the former from the "unmarked" theme in the latter. Unmarked intonation, unlike the marked variety, is always ambiguous with respect to Information Structure. In the following context the same contour will have the Information Structure of (13):

(29) Q: What do you know about Anna?
A: (Anna) (married MANNY).
H* LL%

In these terms it is clear that the context-establishing questions in (13) and (14) can also be analyzed in terms of a theme and a rheme. In both cases, as Prevost (1995) points out, the *wh*-item constitutes the theme. Usually such themes are prosodically unmarked, but they may also bear the marked L+H*LH% theme tune. In either case the phrase *who*, *what*, or *which man* translates as an abstraction such as the following:

(30) $\lambda p.\lambda x.man'x \wedge px$

Such a theme defines a very unspecific set of alternatives, namely, the set of all contextually supported predicates applying to persons, things, men, or whatever. However, the remainder of the *wh*-question, which must bear the H*LL% rheme tune, restricts this set to one particular predicate. It is this predicate that typically becomes the theme of the answer.

5.3.3 Multiple Themes and Rhemes

It is quite possible for an utterance to mark more than one disjoint segment as either theme or rheme. An example of a multiple or discontinuous theme can be found in the following exchange:

(31) Q: I know which team Mary EXPECTS to LOSE. But which one does she WANT to WIN?
A: (Mary WANTS) (IPSWICH) (to WIN.)
L+H* LH% H* L L+H* LH%

The theme established by the question is *Which one Mary wants to win.* That is:[14]

(32) $\exists x.\Diamond * want'(*win'(ana'x))x\ mary'$

We may assume that the rheme alternative set includes propositions like the following:

(33) $\left\{\begin{array}{l}\Diamond want'(win'(ana'watford'))watford'mary'\\ \Diamond want'(win'(ana'ipswich'))ipswich'mary'\\ \Diamond want'(win'(ana'sunderland'))sunderland'mary'\end{array}\right\}$

In the answer the words *wants* and *win* get L+H* pitch accents, because the theme alternative set includes the previous theme, *Which one Mary expects to lose*, or $\exists x.\Diamond want'(lose'(ana'x))x\ mary'$. Since elements of the theme are separated by the rheme *Ipswich* (which of course has its own H* pitch accent and boundary), there are two L+H* LH% theme tunes. These fragments work independently to have the effect of a "discontinuous theme." The first presupposes that the rheme alternative set consists entirely of propositions of the form $want'x\ y\ mary'$; the second presupposes that it consists of propositions of the form $p(win'(ana'y))y\ z$. Both presuppositions are compatible with the same rheme alternative set, so together they require that it consists of propositions of the form $want'(win'(ana'x))x\ mary'$, just as if they constituted a single discontinuous theme.

5.4 Focus and Background

The possibility of such unmarked themes, lacking any pitch accent, draws attention to a second independent dimension to discourse Information Structure that affects intonational tune. In (14) the L+H* LH% tune is spread across the entire substring of the sentence corresponding to the theme in the above sense—that is, over the substring *Anna married*.[15] In (13) the same tune L+H* LH% is confined to the object of the theme *married Manny*, because the intonation of the original question indicates that marrying Manny *as opposed to someone else* is the new topic or theme. In (28) and (29) there is no L+H* LH% tune at all.

The position of the pitch accent in the phrase has to do with a further dimension of Information Structure *within both theme and rheme*, corresponding to a distinction between *the interesting part(s)* of either information unit, and the rest. Halliday (1967b), who was probably the first to identify the orthogonal nature of these two dimensions, called it "new" information, in contrast to "given" information (cf. Brown 1983). The term "new" is not entirely helpful, since (as Halliday was aware) the relevant part of the theme need not be novel to the discourse, as in the examples to hand. Here I will follow the phonological literature and Prevost (1995) in calling the information marked by the pitch accent the "focus," distinguishing theme focus and rheme focus where necessary, and use the term "background" for the part unmarked by pitch accent or boundary. Again there are a number of other taxonomies, most of which are fairly straightforwardly compatible with the present proposal.[16]

The following example serves to illustrate the full range of possibilities for the distribution of focus and background within the theme and the rheme.

(34) Q: I know that Mary envies the man who wrote the musical. But who does she ADMIRE?

A: (Mary ADMIRES) (the woman who DIRECTED the musical)

Mary	ADMIRES	the woman who	DIRECTED	the musical
	L+H* LH%		H*	LL%
Background	*Focus*	*Background*	*Focus*	*Background*
Theme		*Rheme*		

Here the theme is something that I will call *Mary admires*, as an informal shorthand referring to the translation of that part of the utterance as the abstract proposition $\lambda x.admires'x\ mary'$. Only the word *admires* is emphasized, because the previous theme was also about Mary. The presence of pitch accents in themes like that in (34) is marked by distinguishing the corresponding constant in the translation *admires'* with an asterisk:

(35) $\exists x.\Diamond * admires'x\ mary'$

Unless a compatible prior theme—that is, one that matches (35) when $*admires'$ is replaced by some other constant, as in (36)—can be retrieved or accommodated, the utterance is simply infelicitous, and the analysis will fail at this point:

(36) $\exists x.\Diamond like' x\ mary'$

The set of alternative themes in this case is the following:

(37) $\left\{\begin{array}{l}\exists x.\Diamond admires' x\ mary' \\ \exists x.\Diamond like' x\ mary'\end{array}\right\}$

The set of alternative themes is closely related to Büring's (1995) Q, or "question" alternative set. Here I will call it the "theme" alternative set, in contrast to the earlier rheme alternative set.

The rheme alternative set presupposed by the theme is therefore a set of propositions about Mary admiring various people. The rheme is *the woman who directed the musical*, where only the word *directed* is contrasted.

It is important to note that it is all and only the material marked by the pitch accent(s) that is contrasted. This applies when there is more than one pitch accent, as the reader can verify by observing the effect of adding a further pitch accent on the word *musical*. Anything not so marked, including the material between the pitch accent(s) and the boundary, is background. Examples like this suggest that the focusing property of pitch accents applies at the level of words and their interpretations, not at the level of larger constituents, unlike the theme/rheme distinction.

However, there is an asymmetry between the "prenuclear" background material *the woman who ...* that precedes the pitch accent on *directed*, and the background material that succeeds it (*the musical*). The fact that there is no pitch accent on the latter seems to demand that all individuals in the context have the property of having something to do with this particular musical. It would actually be wrong in this context to have a pitch accent. However, the lack of accent on the former does not seem similarly to demand that all the individuals that we are restricting over are women, and in fact in the example they are not. The implication is that in this context the property of directing the musical is sufficient to distinguish the individual uniquely—the fact that this individual is *also* unique by virtue of being a woman need not be marked.[17]

Why does this asymmetry hold? Could it work the other way around? Prevost (1995) has proposed that pitch accents are assigned to words corresponding to properties that successively limit the extension of an NP to the desired referent. If we assume that the order in which these predicates are successively evaluated is "bottom up" with respect to some kind of interpretation or predicate-argument structure, then we predict that if a modifier like the relative clause *who directed the musical* completely specifies the extension, then words corresponding to predicates higher up the predicate-argument structure, such as the head noun, need not be stressed.[18] Thus, the pragmatic difference between prenuclear and postnuclear background material arises from the relation of word order to predicate-argument structure in English.

5.5 Grammar and Information Structure

What is the relation of such Intonation Structures to syntax and semantics, and how is Information Structure computed? Many of the intonational constituents that we have just been examining—such as the string *Anna married*—do not correspond to traditional syntactic constituents. Jacobs (1991) and Krifka (1991, sec. 4.8) have pointed out the problems that ensue for the semantics of focus.

Such "nonconstituent" intonational units are very widespread and can cooccur with other intonational tunes, including the H*+L rheme tune considered here. Consider the following utterance (adapted from Ladd 1980—see below), uttered in the context of a question like *I know that Harry keeps up with the newspapers, but has he read* War and Peace*?*:

(38) Harry doesn't READ BOOKS!
H*L L+H* LH%

Here the theme seems to be *books*, marked because the concept stands in contrast to *newspapers*. The rheme seems to be *Harry not reading something*, with the usual final H*L tune on *read*. The theme can also be unmarked, even in this context, as in Ladd's original example:[19]

(39) Harry doesn't READ books.
H* LL%

The interest of such derivations for present purposes will be obvious. The claim is simply that the nonstandard Surface Structures that are induced by the combinatory grammar to explain coordination in English subsume the Intonation Structures that are postulated by Pierrehumbert and her colleagues to explain the possible intonation contours for sentences of English. The claim is that in spoken utterances, intonation helps to determine which of the many possible bracketings permitted by the combinatory syntax of English is intended, and that the interpretations of the constituents that arise from these derivations, far from being "spurious," are related to distinctions of Information Structure and discourse focus among the topics that the speaker has in mind and the comments that the speaker has to contribute.

The proof of this claim lies in showing that the rules of combinatory grammar can be made sensitive to intonation contour, which limits their application in spoken discourse. We must also show that the major constituents of intonated utterances like (14), under the analyses that are permitted by any given intonation, correspond to the Information Structure required by the context to which the intonation is appropriate. A preliminary proposal, to which the next section proposes a revision and an extension, was sketched in Steedman 1991a,b, and Prevost and Steedman 1994.

5.5.1 Combinatory Prosody

The papers just cited followed Pierrehumbert and colleagues in taking an "autosegmental" approach to the tones, in the sense of the term introduced by Goldsmith (1976), assigning a distinct prosodic CG category to all elements of the sentence, as well as a grammatical one. Like grammatical categories, prosodic categories could be either functions or arguments and could apply to or compose with one another. Syntactic combination was made subject to intonation contour by the assumption of a "Prosodic Constituent Condition", which only permitted combination of syntactic categories via a syntactic combinatory rule if their result was also a prosodic constituent.

The present version of the theory takes a different approach, integrating prosodic information with the standard grammatical categories to more directly capture Intonation Structure, together with its interpretation as Information Structure, in CCG.

We have already noted that the focus-marking property of pitch accents seems to belong at the level of the word, whereas the theme/rheme-marking property seems to belong at the level of phrasal constituents. We therefore begin by assuming that pitch accents both mark (some element of) the interpretation of the words they occur on for focus or contrast, and mark the syntactic category in a way that "projects" theme-rheme status to elements with which the word combines. Although eventually we will certainly want to do this by morphological rule, for present purposes we will regard this compiled out into distinct lexical entries like the following categories for the verb *ate* bearing the two pitch accents under discussion here. before syntax gets to work on them:[20]

(40) ate := $(S_\theta\backslash NP_\theta)/NP_\theta : *ate'$
L+H*

(41) ate := $(S_\rho\backslash NP_\rho)/NP_\rho : *ate'$
H*

The subscript symbols θ and ρ are mnemonic for theme and rheme respectively, and are a shorthand for a value on a feature of the whole category that I will call *INFORMATION*. A category like (40) ensures that any argument that combines with it must be phonologically compatible with being part of a theme.

The "null tone," which I will follow Pierrehumbert in leaving without any annotation in strings, does not affect the interpretation of a word that carries it, and leaves the syntactic category unspecified as to the value of the feature *INFORMATION*. It can therefore conveniently be written without any annotation, as before:

(42) ate := $(S\backslash NP)/NP : \mathit{ate}'$

Since the value of *INFORMATION* is unspecified, this category can combine with either θ, ρ, or unmarked categories. However, it is important to remember that the unspecified values on arguments and result are the *same* unspecified value of the same attribute *INFORMATION*. In the first two cases, this *INFORMATION* value becomes specified for all arguments and the result, by the usual unification mechanism, and subsequent combinations must be compatible with that value.[21]

Prosodically annotated categories of this kind allow the influence of the pitch accent to spread over arbitrarily large constituents. For example, in analyzing the first two words of the sentence *Fred ate the beans*, uttered in response to a question like *I know what Harry ate. But what did FRED eat?*, the following partial derivation can occur:

(43)
$$\begin{array}{cccccc}
\text{FRED} & \text{ate} & & \text{the BEANS} & \\
\text{L+H*} & & \text{LH\%} & \text{H*} & \text{LL\%} \\
\overline{S_\theta/(S_\theta\backslash NP_\theta) : \lambda p.p\ {*}\mathit{fred}'} & \overline{(S\backslash NP)/NP : \lambda x.\lambda y.\mathit{ate}'xy} & & & \\
\multicolumn{2}{c}{\overline{\hspace{8cm}}{>}\mathbf{B}} & & & \\
\multicolumn{2}{c}{S_\theta/NP_\theta : \lambda x.\mathit{ate}'x\ {*}\mathit{fred}'} & & &
\end{array}$$

The L+H* pitch accent on *FRED* marks all elements of the raised subject category as θ on the *INFORMATION* feature. The verb bears the null tone, but when the subject composes, all occurrences of the verb's own *INFORMATION* feature come to bear the value θ by the unification mechanism. Hence the object in the category that results from composition also bears the *INFORMATION* value θ.

In contrast to the version in Steedman 1991a, the present theory assumes that boundary tones, unlike pitch accents, are not associated with words, but are elements of the string in their own right, much like the punctuation marks that, on occasion, represent them in the orthography. Like the pitch accents, the boundary tones affect both the syntactic and the semantic components of categories. The grammatical category of a boundary is that of a function from categories marked as θ, ρ, or unspecified, into phonological phrasal categories, distinguished by a value ϕ.[22]

The boundary tones must also mark the informational units at the level of the interpretation, so that the combination of a constituent bearing a boundary tone with another including a pitch accent semantically defines the major informational elements such as the theme and the rheme.

For present purposes, the full categories for the three boundary tones under discussion here are written in full using the following notation.[23]

(44) L, LL%, LH% := $S\$_{\phi}\backslash S\$_{\eta} : \lambda f.\eta' f$

The variable $S\$$ ranges as usual over a set $\{S\$\}$ of categories including S and all functions into members of $\{S\$\}$—that is, it includes S, S/NP, and all verbs and type-raised arguments of verbs, but not nouns and the like. The subscript η, which can be thought of as a variable ranging over the two *INFORMATION* values θ and ρ, further specifies it as ranging over correspondingly marked categories S_{θ}, S_{θ}/NP_{θ}, $(S_{\rho}\backslash NP_{\rho)})/NP_{\rho)}$, etc. When it combines with such a function, it has the effect of replacing its θ or ρ marking with a distinct marker ϕ (for "phrasal"), which can only unify with itself. Such a category can only combine with other ϕ-marked prosodically phrasal categories.

Semantically, the boundary categories apply a corresponding thematic or rhematic function θ' or ρ' to the interpretation of the category with which they combine, via a corresponding variable category η'. In terms of the Logical Form, θ' or ρ' are identity functions that effectively vanish from the predicate-argument structure when they apply. However, they are assumed to cause the appropriate alternative set to be evoked from the database, and to be accompanied by the updates discussed in section 5.3.1 Until they do apply, they block any further reduction of the interpretation to the canonical predicate-argument structure.

This specification of boundaries allows them to combine with either pitch accent, consistent with Pierrehumbert's own system, in which all pitch accents and boundaries can combine freely. For example, low boundaries can combine with L+H* pitch accents, as in the following sentence, from Ladd 1996, 96-7:

(45) THAT's the whole POINT of the exercise!
H* L L+H* LL%

The implicit claim that the boundaries also project θ or ρ marking from the pitch accent, so that the phrase *the whole point of the exercise* is a theme in the above example (as it would be if it bore an LH% boundary), is more controversial, and will not be discussed here. Nevertheless, we will see later that at least one variety of theme, the unmarked variety, does bear low boundaries.

To say that the boundary projects the category of the pitch accent is not to exclude a more active role in the semantics for the boundary, analogous to a specifier such as a determiner, contributing distinctively to information structural content along the lines suggested by Pierrehumbert and Hirschberg (1990).

Since they bear exactly the same category as L boundaries, LL% boundaries are free in the present system to occur utterance-medially, in contrast to earlier versions of the theory which were criticized on this point by Beckman (1996, 63-64). Utterance medial LL% boundaries do not figure much in the examples discussed in the present chapter (although see (75) below and the discussion of example (60) in chapter 7). In particular it does not appear to be possible to substitute them freely for L intermediate phrase boundaries in examples like (13) (a fact upon which Beckman does not comment). By the same token, the present system allows L boundaries to occur utterance-finally, which is impossible. We will assume for present purposes that these details are to do with finer distinctions between the boundaries, and in particular with the distinction between intermediate and intonational phrases. The question of whether LL% boundaries are or are not categorically distinct from intermediate phrase L boundaries is a matter of some dispute among phonologists, and we will continue to pass over it here.

The following example, which completes the derivation of the theme of the earlier sentence *FRED ate the BEANS*, demonstrates the effect of the boundary tone:

(46)

$$
\begin{array}{cccccc}
\text{FRED} & \text{ate} & & \text{the BEANS} & \\
\text{L+H*} & & \text{LH\%} & \text{H*} & \text{LL\%} \\
\hline
S_\theta/(S_\theta\backslash NP_\theta) & (S\backslash NP)/NP & S\$_\phi\backslash S\$_\eta & & \\
:\lambda p.p *fred' & :\lambda x.\lambda y.ate'xy & :\lambda f.\theta' f & & \\
\hline
\multicolumn{2}{c}{S_\theta/NP_\theta : \lambda x.ate'x *fred' \quad {>}\mathbf{B}} & & & \\
\hline
\multicolumn{3}{c}{S_\phi/NP_\phi : \theta'(\lambda x.ate'x *fred') \quad {<}} & &
\end{array}
$$

The second prosodic phrase in (46) bears the H*LL% rheme tune, parallel to (14) above. The complete derivation is as follows:

(47)

$$
\begin{array}{cccccc}
\text{FRED} & \text{ate} & & \text{the} & \text{BEANS} & \\
\text{L+H*} & & \text{LH\%} & & \text{H*} & \text{LL\%} \\
\hline
S_\theta/(S_\theta\backslash NP_\theta) & (S\backslash NP)/NP & S\$_\phi\backslash S\$_\eta & NP^\uparrow/N & N_\rho & S\$_\phi\backslash S\$_\eta \\
:\lambda p.p\ *fred' & :\lambda x.\lambda y.ate'xy & :\lambda f.\theta' f & :\lambda x.\lambda p.p(the'x) & :*beans & :\lambda f.\rho' f \\
\hline
\multicolumn{2}{c}{S_\theta/NP_\theta : \lambda x.ate'x\ *fred' \quad (>\mathbf{B})} & & \multicolumn{2}{c}{NP^\uparrow_\rho : \lambda p.p(the' *beans') \quad (>)} & \\
\hline
\multicolumn{3}{c}{S_\phi/NP_\phi : \theta'(\lambda x.ate'x\ *fred') \quad (<)} & \multicolumn{3}{c}{NP^\uparrow_\phi : \rho'(\lambda p.p(the' *beans')) \quad (<)} \\
\hline
\multicolumn{6}{c}{S_\phi : \rho'\lambda p.p(the' *beans')(\theta'\lambda x.ate'x\ *fred') \quad (<)} \\
\hline
\multicolumn{6}{c}{S_\phi : ate'(the' *beans') *fred'}
\end{array}
$$

No *other* division into a theme and a rheme is possible for this intonation contour.[24]

It is only once the functions θ' and ρ' have applied that the final semantic reduction or normalization can take place, to yield the canonical predicate-argument structure. As far as the interpretation goes, these are just the identity function, but we have assumed that they are accompanied by side effects of assertion or retraction on the database. The reduction of θ' and ρ' can occur at any point in a derivation.

The division of the utterance into a property constituting the theme and an argument constituting the rheme is appropriate to a context parallel to that established in (14)—say, by the *wh*-question *What did FRED eat?* uttered in the context of a discussion of what somebody else (say, Harry) ate and a prior theme such as the following:

(48) $\theta'\lambda x.ate'x\ harry'$

The theme in derivation (47), $\theta'\lambda x.ate'x\ *fred'$, is a member of the theme alternative set in the earlier sense.[25] The new theme presupposes a rheme alternative set of propositions about Fred eating things, which the rheme $\rho'\lambda p.p(the' *beans')$ reduces to a single proposition, $ate'(the'beans')fred'$.

Since categories bearing the null tone can compose either with others bearing null tone, or with those bearing pitch accents, intonational phrasal tunes like L+H* LH% can spread unboundedly across any sequence that forms a grammatical constituent according to the combinatory grammar. For example, if the answer to *What did FRED eat?* is *MARY says he ate BEANS*, then the tune will typically be spread over *Mary says he ate ...* as in the (incomplete) derivation (49), in which the semantics has been suppressed in the interests of brevity:

(49)

$$
\begin{array}{lllll}
\text{MARY} & \text{says} & \text{he} & \text{ate} & \ldots \\
\text{L+H*} & & & & \text{LH\%} \\
\hline
S_\theta/(S_\theta\backslash NP_\theta) & (S\backslash NP)/S & S/(S\backslash NP) & (S\backslash NP)/NP & S\$_\phi\backslash S\$_\eta \\
\end{array}
$$

$$
\begin{array}{l}
\underline{S_\theta/(S_\theta\backslash NP_\theta)\quad (S\backslash NP)/S}{>}\mathbf{B} \\
S_\theta/S_\theta \\
\underline{S_\theta/S_\theta \quad S/(S\backslash NP)}{>}\mathbf{B} \\
S_\theta/(S_\theta\backslash NP_\theta) \\
\underline{S_\theta/(S_\theta\backslash NP_\theta) \quad (S\backslash NP)/NP}{>}\mathbf{B} \\
S_\theta/NP_\theta \\
\underline{S_\theta/NP_\theta \quad S\$_\phi\backslash S\$_\eta}{<} \\
S_\phi/NP_\phi
\end{array}
$$

A number of derivations are permitted for more complex rheme NPs like *the green beans* with a final pitch accent, represented by the following bracketings:

(50) a. (the green) (BEANS)
H* LL%

b. (the) (green BEANS)
H* LL%

The same derivational ambiguity is characteristic of a further intonation contour, in which the H* pitch accent is applied to the word *green*, making that word alone contrasted:

(51) Q: Did Fred eat the green beans or the yellow beans?

A: (FRED ate) (the GREEN beans.)
L+H* LH% H* LL%

It also applies when both words are marked for rheme accent with H* pitch accents, in a context in which both elements of the noun group *green beans* are contrasted or informative:

(52) Q: Did Fred eat the green beans or the yellow squash?

A. (FRED ate) (the GREEN BEANS.)
L+H* LH% H* !H* LL%

Example (52) needs further comment. First, the rheme includes more than one H* pitch accent. These pitch accents belonging to the same intonational/intermediate phrase would be seen in a pitch track to be downstepped—that is, to have successively lower peaks and baselines for approximately the same pitch range. The ToBI conventions (Silverman et al. 1992) again offer a convenient notation for downstep, extending Pierrehumbert's system with the prefix "!". However, as in Pierrehumbert's original theory, words bearing downstepped pitch accents bear identical categories to non-downstepped accents. The !-notation is redundant, and the effect on pitch-contour can be determined automatically.

Downstepped pitch accents can therefore compose, to allow derivations like the following:

(53) FRED ate the GREEN BEANS
L+H* LH% H* !H* LL%

$$\frac{}{S_\phi/NP_\phi : \theta'(\lambda x.ate'x *fred')} \quad \frac{}{NP/N : the'} \quad \frac{}{N_\rho/N_\rho : *green} \quad \frac{}{N_\rho : *beans'} \quad \frac{}{S\$_\phi \backslash S\$_\eta : \lambda f.\rho' f}$$

$$\frac{NP/N : the' \quad N_\rho/N_\rho : *green}{NP_\rho/N_\rho : \lambda x.the'(*green'x)} {>}\mathbf{B}$$

$$\frac{NP_\rho/N_\rho \quad N_\rho}{NP_\rho : (the'(*green' * beans'))} <$$

$$\frac{NP_\rho : (the'(*green' * beans'))}{S_\rho \backslash (S_\rho/NP_\rho) : \lambda f.f(the'(*green' * beans'))} {<}\mathbf{T}$$

$$\frac{S_\rho \backslash (S_\rho/NP_\rho) \quad S\$_\phi \backslash S\$_\eta}{S_\phi \backslash (S_\phi/NP_\phi) : \rho'(\lambda f.f(the'(*green' * beans')))} <$$

$$\frac{S_\phi/NP_\phi \quad S_\phi \backslash (S_\phi/NP_\phi)}{S_\phi : \rho'\lambda f.f(the'(*green' * beans'))(\theta'\lambda x.ate'x *fred')} <$$

$$\frac{}{S_\phi : ate'(the'(*green' * beans'))*fred'}$$

This time, there is another derivation for the rheme in (53). In fact, both derivations illustrated in (50) are permitted for all three stress patterns. The sentences differ only in the elements of the interpretation that are marked for contrast. These observations reinforce the earlier suggestion that the effect of the pitch accents applies at the level of words words and their interpretations, rather than at higher levels of derivation, unlike the effect of boundary tones.

Many impossible intonation contours correctly remain excluded by the fact that CCG conflates prosodic structure and syntactic structure, including examples of the kind that motivated Selkirk's Sense Unit Condition. For example, the following are disallowed because their Intonation Structure is not compatible with any syntactic analysis, owing to island constraints:

(54) a. *(FRED ate the green) (BEANS.)
L+H* LH% H* LL%

b. *(My OLDER) (sister ate the green BEANS.)
L+H* LH% H* LL%

5.5.2 Unmarked Themes

So far I have only considered sentences that include a theme and rheme that both include words marked for contrast by pitch accents. Such sentences are relatively unambiguous with regard to their Information Structure. However, sentences like the following, which in Pierrehumbert's terms consist of a single intonational phrase, are much more common:

(55) (Mary wrote a book about BATS.)
H* LL%

Such sentences are notoriously ambiguous with respect to the theme they presuppose (cf. Selkirk 1995). For example, (55) seems equally appropriate as a response to any of the following questions:

(56) a. What did Mary write a book about?
b. What did Mary write?
c. What did Mary do?
d. What's new?

Such questions could in more contrastive contexts give rise to themes marked by the L+H* LH% tune, bracketing the sentence as follows:

(57) a. (Mary wrote a book about)$_{Theme}$(BATS.)$_{Rheme}$
b. (Mary wrote)$_{Theme}$(a book about BATS.)$_{Rheme}$
c. (Mary)$_{Theme}$(wrote a book about BATS.)$_{Rheme}$
d. (Mary wrote a book about BATS.)$_{Rheme}$

It is therefore a virtue in the grammar as it stands that it already allows all of the implicit derivations for the sentence in which the theme is unmarked, while the various possible rhemes are marked as such in the derivation.

Such unmarked themes can be made explicit in the theory as follows. The boundary categories (44) are already defined so as to allow them to combine with unmarked categories, on the assumption that an unspecified *INFORMATION* value can unify with or match the variable η in the boundary category.

If we further assume that an L boundary is phonetically indistinguishable from the null tone, then such a boundary tone may be postulated anywhere there is null tone (and low pitch). Such a tactic nondeterministically allows all of the derivations indicated in (57). For example:

(58)

$$
\begin{array}{ccc}
\text{Mary wrote a book about} & \text{L} & \text{BATS} \quad \text{H}* \quad \text{LL\%} \\
\hline
S/NP & S\$_{\phi}\backslash S\$_{\eta} & S_{\phi}\backslash(S_{\phi}/NP_{\phi}) \\
:\lambda x.write'(a'(book'(about'x)))mary' & :\lambda f.\eta' f & :\rho'(\lambda p.p * bats') \\
\hline
\multicolumn{2}{c}{S_{\phi}/NP_{\phi} \quad <} & \\
\multicolumn{2}{c}{:\eta'(\lambda x.write'(a'(book'(about'x)))mary')} & \\
\hline
\multicolumn{3}{c}{S_{\phi} : \rho'(\lambda p.p * bats')(\eta'(\lambda x.write'(a'(book'(about'x)))mary')) \quad <} \\
\hline
\multicolumn{3}{c}{S_{\phi} : write'(a'(book'(about' * bats')))mary'}
\end{array}
$$

On the reasonable assumption that an unspecified η' has the same effect as a value of θ' in the interpretation, apart from necessarily applying to an argument lacking any * marker of contrast, then the representation of theme and rheme in the interpretation is exactly as in the earlier examples.

As Steedman 1991a points out, this nondeterminism can be eliminated for processing purposes by taking advantage of the fact that the unmarked theme is exclusively used when the hearer is assumed to already know the theme. Thus, the appropriateness of applying the rule to a given category can be directly decided by referring to the discourse model to see whether it supports the presupposition that the corresponding referent is theme, background, or whatever. (See Straub 1997 for experimental evidence for the systematic omission of explicit prosodic boundaries by speakers when alternative sources of disambiguating information, including contextual, are present.)

The ambiguity of such contours with respect to Information Structure appears to be correctly constrained by the assumption that Information Structure and syntactic structure must coincide. That is, the following do not appear to be possible Information Structures, because, like the related examples in (54), they are not possible syntactic structures.[26]

(59) a. *(Fred ate the green)$_{Theme}$(BEANS)$_{Rheme}$
 b. *(My older)$_{Theme}$(sister ate the green BEANS)$_{Rheme}$

The trick of nondeterministically assuming an invisible boundary to null themes may seem unnecessarily clumsy, and to compromise the theory by introducing phonological entities that have no phonetic realization, as Croft (1995) and Ladd (1996) have suggested. I will return to this point below. However, the device captures an important generalization concerning the relation of these unmarked themes to the corresponding marked ones, and to another variety of unmarked theme which *does* have an explicit boundary.

In English (as opposed to many other languages—see Ladd 1996; Zubizarreta 1998) unmarked themes can occur utterance-finally, and when they do, they end with an LL% boundary, as in the following example:[27]

(60) Q: Who ate the beans?
 A: (FRED) (ate the beans).
 H* L LL%

If the rheme *FRED* is to be a well-formed intonational phrase distinct from the unmarked theme, it must end in an L intermediate phrase boundary. Again postulating such a boundary introduces a nondeterminism—but again this nondeterminism arises only in contexts where the theme in question is entirely given, or background, and hence is recoverable by the hearer. (When evaluated, such themes must by definition yield a set of alternative propositions identical to the background set. I will return to this point.)

However, an intonational-phrase-final LL% tone cannot always be analyzed in this way. It may just be background information in the rheme. To take an old example from Schmerling (1976), one may announce the death of Nixon in the absence of any prior discourse by saying the following:

(61) NIXON died.
     H*   LL%

The second word is then part of the rheme, which of course is allowed by the grammar, and the utterance is felicitous just in case dying is a background possibility for the individual in question. (If not, as Schmerling points out, one has to say something like *Nixon DIED*.)

The other, apparently phonetically indistinguishable analysis for this sentence, with an unmarked theme, is of course still available and is appropriate to a situation where the question is *Who died?*, as Ladd (1980, 53) points out. It is this analysis that is at work in Ladd's own example (39), repeated here in the revised notation, uttered in the context of a question *Has Harry read* War and Peace*?*:

(62) Harry doesn't READ books.
                  H* L   LL%

Although the analysis proposed here is quite different from Ladd's, he supports a view of the Information Structure according to which the utterance is "about a book" (Ladd 1980, 52)—in present terms, where *books* is the theme.

5.6 Intonation and the Simplex Clause

Jackendoff (1972) exhaustively examines the effect of all possible assignments of the two tunes considered here to a simple transitive clause, *Fred ate the beans*, and it is instructive to do the same within the present framework.

In contrast to the intonation in the derivation (47), the intonation contour on (18) prevents the composition of subject and verb, because under the forward prosodic composition rule the subject is not allowed to combine with the verb. It follows that a derivation parallel to the earlier one (and the formation of the corresponding theme) is not allowed. On the other hand, the following derivation *is* allowed for (18):

(63)

$$
\begin{array}{cccccc}
\text{Fred} & & \text{ate} & \text{the} & \text{beans} & \\
\text{H*} & \text{L} & & & \text{L+H*} & \text{LH\%} \\
\hline
S_\rho/(S_\rho\backslash NP_\rho) & S\$_\phi\backslash S\$_\eta & (S\backslash NP)/NP & NP^{\uparrow}/N & N_\theta & S\$_\phi\backslash S\$_\eta \\
:\lambda p.p * fred' & :\lambda f:\rho' f & :\lambda x.\lambda y.ate'xy & :\lambda x.(the'x) & :*beans' & :\lambda f:\theta' f \\
\hline
\multicolumn{2}{c}{S_\phi/(S_\phi\backslash NP_\phi):\rho'\lambda p.p * fred' \quad <} & & \multicolumn{2}{c}{NP^{\uparrow}_\theta : the' * beans' \quad >} & \\
\hline
\multicolumn{6}{c}{S_\phi:\rho'\lambda p.p * fred'(\theta'\lambda y: ate'(the' * beans')y) \quad >} \\
\hline
\multicolumn{6}{c}{S_\phi: ate'(the' * beans') * fred}
\end{array}
$$

Again no other analysis is allowed, and again the division into rheme and theme, and the associated interpretations, are consistent with the context given in (18).

The effect of the above derivation is to annotate the entire predicate as theme, just as if the tune L+H* LH% had been spread across the whole constituent.

Other cases considered by Jackendoff are accepted under the same assumptions and in every case yield unique and contextually appropriate interpretations, as follows. (The derivations themselves are suggested as an exercise.) The first two yield derivations parallel to (63), in that the fundamental division of the sentence is into a traditional subject and predicate (again these are the only analyses that the rules permit):

(64) Q: What about FRED? What did HE do to the beans?

A: (FRED) (ATE the beans.)
L+H* LH% H* LL%
Theme *Rheme*

(65) Q: I know who COOKED the beans. But then, who ATE them?

A: (FRED) (ATE the beans.)
H* L L+H* LH%
Rheme *Theme*

The other two cases considered by Jackendoff yield derivations parallel to (47), in which the fundamental division of the sentence is orthogonal to the traditional subject-predicate structure:

(66) Q: I know what Fred COOKED. But then, what did he EAT?

A: (Fred ATE) (the BEANS.)
L+H* LH% H* LL%
Theme *Rheme*

(67) Q: Well, what about the BEANS? What did Fred do with THEM?

A: (Fred ATE) (the BEANS.)
H* L L+H* LH%
Rheme *Theme*

In the case of (66) at least, it seems obvious that the theme established by the context is indeed the one corresponding to the bracketing. In the case of (67) it is less obvious. However, the treatment of relative clauses below will show that this analysis must at least be available.

The following further derivation for (67) is also allowed, as is a parallel derivation for (66), in which *Fred* is a background component of a discontinuous theme, rather than a background part of the rheme:

(68) Q: Well, what about the BEANS? What did Fred do with THEM?

A: (Fred) (ATE) (the BEANS.)
L H* L L+H* LH%
Theme Rheme *Theme*

There seems to be little to distinguish the alternatives on pragmatic or phonetic grounds. It is the context that determines which Information Structure is felicitous.

Two further cases, which are parallel to (63) and (47) except that the theme and rheme tunes are exchanged, are also accepted, again yielding unique, contextually appropriate analyses. The first is the following:

(69) Q: I know that ALICE read a BOOK. But what about FRED? What did HE do?

A: (FRED) (ate the BEANS.)
L+H* LH% H* LL%
Theme *Rheme*

The contour on the response here is also a coherent response in the context used in (43). This possibility (which may be the one intended in Jackendoff's 1972 discussion of the example) appears to arise from an ambiguity in the context itself. However, the converse does not apply: the intonation on the response in (64) is not felicitous in the above context, as the following example shows:[28]

(70) Q: I know that ALICE read a BOOK. But what about FRED? What did HE do?

A: #(FRED ate) (the BEANS.)
L+H* LH% H* LL%
Theme *Rheme*

There is one final possibility, which Jackendoff does not distinguish from (63). It is intuitively less obvious than the others, because its discourse meaning is better expressed (at least in the written language) by a left dislocation, *As for the BEANS, FRED ate them*, or even a passive *The BEANS were eaten by FRED*, uttered with the same assignment of pitch accents to *the beans* and *Fred*. Again the use of a second pitch accent on the verb *ate* in the rheme would also improve the example. Its place in the scheme of things will become clearer in section 5.7.2.

(71) Q: Well, what about the BEANS? What happened to THEM?

A: (FRED ate) (the BEANS.)
H* L L+H* LH%
Rheme *Theme*

5.7 Intonation in Complex Constructions

The number of possible intonation contours for complex sentences is naturally even larger than those that have just been demonstrated for simple transitive sentences, and the contextual conditions that are required to make them felicitous are even more abstruse. The following sections are necessarily restricted to showing that the theory makes correct predictions concerning the complex constructions in which forward composition is necessarily implicated in syntax (in particular, reduced coordinate sentences and relative clauses), rather than merely allowing alternative derivations.

5.7.1 Coordinate Sentences

Since the coordinate sentence (72a) below necessarily involves composition of the (type-raised) subject with the verb, whereas (72b) necessarily does not, it is predicted that the intonation contours that they permit will be more restricted than for the non coordinate sentence (72c):

(72) a. Bill cooked, and Fred ate, the beans.
b. Fred ate the beans, and drank the wine.
c. Fred ate the beans.

For example, among other alternatives, we would expect the intonation contour (73) to be possible for (72a). (The example assumes the mechanism for multiple pitch accents of section 3.3. It is a possible answer to the question *What did Bill and Fred do with the beans?*)

(73) (Bill COOKED and Fred ATE) (the BEANS.)
H* !H*L L+H* LH%

By contrast, intonational tunes that assign categories that are not consistent with the crucial syntactic compositions block derivation:

(74) a. *(Bill cooked and FRED) (ate the BEANS.)
H*L L+H* LH%
b. *(Bill cooked and FRED) (ate the BEANS.)
L+H*LH% H* LL%

This effect is sufficiently strong for garden paths to be forced under the same principle, as we saw earlier:

(75) *(Harry likes the NUTS) (and bolts APPROACH.)
H*LL% H* LL%

5.7.2 Relative Clauses

Since relative clauses, like the coordinate structures of section 4.1, also force the involvement of functional composition, a similar conspiracy with intonation is predicted for them as well. And indeed, all the possible intonational tunes that appeared in Jackendoff's (1972) examples on the fragment *Fred ate*—that is, all those that allow syntactic composition under the Prosodic Constituent Condition—can also appear on the same fragment when it occurs as the residue in a relative clause:

(76) the beans that Fred ate
a. L+H* LH%
b. L+H*LH%
c. H* LL%
d. H*LL%

(The null tone is of course also allowed on the relative clause.) Each alternative conveys different presuppositions concerning the context. Since the cleft construction is often used with the *wh*-clause marked with the theme tune, L+H* LH%, (77) and (78) show one way of making the first two alternatives—(76a) and (76b) respectively—felicitous:

(77) Q: FRED didn't eat the POTATOES. HARRY ate THEM.
A: (It was the BEANS) (that FRED ate.)
H* L L+H* LH%

(78) Q: Fred didn't eat the POTATOES. He threw THEM AWAY.
A: (It was the BEANS) (that Fred ATE.)
H* L L+H*LH%

The H* LL%tune, which marks the rheme, is frequently used on restrictive relatives, so (79) and (80) may serve to make the remaining two cases—(76c) and (76d) respectively—felicitous. (I have assumed an analysis with an unmarked theme, but this detail is not crucial.)

(79) Q: It wasn't the beans that HARRY ate that looked so delicious.
A: (It was) (the beans that FRED ate.)
H* LL%

(80) Q: It wasn't the beans that Fred COOKED that looked so delicious.
A: (It was) (the beans that Fred ATE.)
H*LL%

The converse also holds. As in the case of coordination, tone sequences that are consistent with no CCG derivation are forbidden from appearing on the relative clause. Thus, we predict that (81a,b) are intonationally disallowed, for the same reason that (81c) is ruled out:

(81) a. *(The beans that FRED) (ate were DELICIOUS.)
H*L L+H* LH%
b. *(The beans that FRED) (ate were DELICIOUS.)
L+H*LH% H* LL%
c. *The beans that I, and squash that you, cooked were delicious

Thus a condition akin to Selkirk's Sense Unit Condition emerges as a theorem of the assumptions inherent in CCG, without independent stipulation.

5.8 Conclusion

At this point it should be clear that we can simply subsume both Intonation Structure and Surface Structure under a single notion of Information Structure. Such a view of Intonation Structure involves a richer structural representation than the one invoked under that name by Pierrehumbert and others, since Information Structure includes many constituent boundaries that do not have any phonetic realization. This should not seem a strange conclusion to reach. It simply means that information-structural boundaries are no more (and no less) completely specified by tones than syntactic boundaries are specified by words.

This observation means that Ladd's (1996, 224) criticism of earlier version of this account as being "compromised by its dependence on entities whose presence cannot be independently [phonetically] verified" and Croft's (1995, 856) related suggestion that it "leaves unexplained the mismatch between prosody and information structure" both rather miss the point. One might as well criticize standard syntactic theories for using lexically unattested brackets. There *is* no "mismatch" between interpretable structure and its surface markers such as function words and tones. These markers are simply rather ambiguous and highly elliptical, just like everything else in natural language.

If Information Structure boundaries and surface syntactic boundaries coincide in this way, then there are a number of other prosodic effects that should depend upon the Surface Structures afforded by CCG in as direct a manner as English intonation contour. Some obvious candidates are such vowel-harmonic effects as French *liaison* (Selkirk 1972), American English flapping (Nespor and Vogel 1986), the Rhythm Rule (Gussenhoven 1983; Selkirk 1984), Bengali /r/ assimilation (Hayes and Lahiri 1991), and Italian *raddoppiamento syntattico* (Napoli and Nespor 1979; Kaisse 1985; Nespor and Vogel 1986; Vogel 1994). The last authors in particular show information-structural effects of focus that seem likely to be capturable in this way. These phenomena would then be brought under the category of "superficial" constraints of syntax on phonology called for by Pullum and Zwicky (1988), since Surface Structure now completely specifies the input to phonology and relation to the metrical grid (Liberman and Prince 1977; Prince 1983; Hayes 1995). In fact, in the sense in which Pullum and Zwicky intend the term, the present theory of syntax is "phonology-free," although in another sense CCG syntax actually subsumes prosodic structure.

More speculatively, it seems likely that many of the "end-based" effects of syntax upon phonology argued for by Selkirk (1986, 1990), Selkirk and Shen (1990), Hirst (1993), and Truckenbrodt (1995, 1999), according to which intonation-structural boundaries coincide with either left or right edges of syntactic constituents, but not necessarily both, are artifacts of the syntactic theories within which they are being framed. That is to say, English appears to be a left-edge language because a traditional right-branching account of its Surface Structure just doesn't offer phonologists enough right brackets to work with. The present theory simply claims that those right brackets are there in the syntax, in left-branching structures like (5). Under this interpretation of Surface Structure it is unnecessary to postulate an additional independent prosodic structure, as does Selkirk, and as do Nespor and Vogel (cf. Vogel and Kenesei

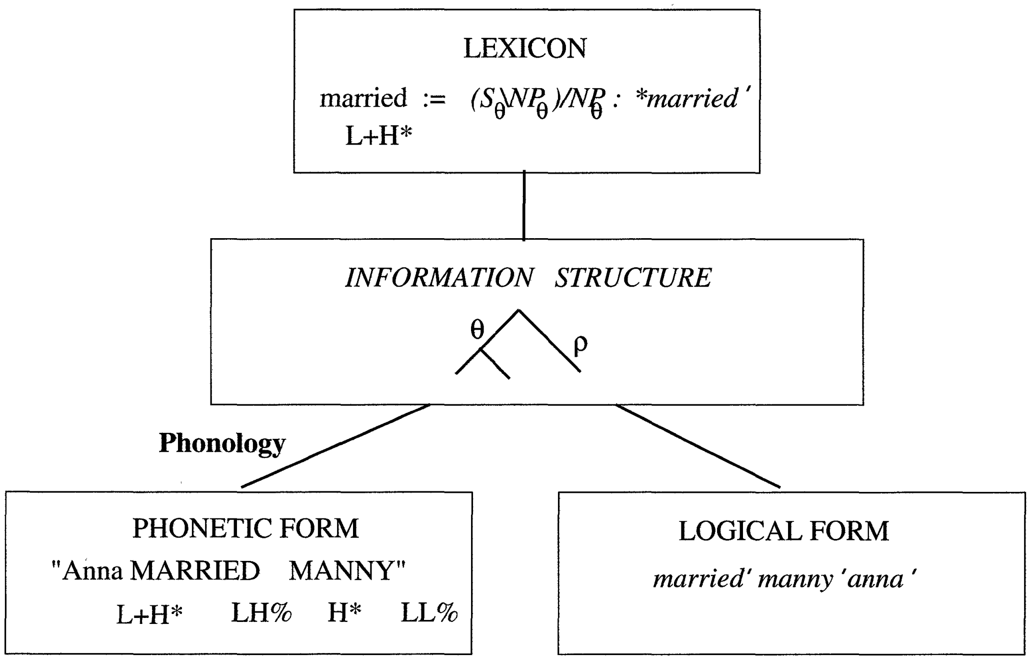


Figure 5.3
Architecture of a combinatory grammar, II

1990; Zec and Inkelas 1990). We should instead adopt a more liberal notion of syntactic structure, one that is directly compatible with the boundaries that the phonologists observe.[29]

However, the status of Surface Structure in such a theory is very different from the status of the related concepts in GB and earlier theories such as the "annotated surface structures" of Chomsky (1971) and Jackendoff (1972). To understand this point, it will again be helpful to consider the architecture of the present theory of grammar in terms of the traditional "T" or "Y" diagram, as in figure 5.3, which includes an example of an object of the kind represented in each structural module for the following sentence:

(82) Anna MARRIED MANNY.
L+H* LH% H* LL%

According to the present theory, the lexicon associates category-interpretation pairs with (the phonological specification of) each word of the language. Derived objects or constituents also pair (the phonological specification of) strings with category-interpretation pairs, which are projected in parallel from (ordered multisets of) lexical entries, via derivations using combinatory rules. In the case of both lexical items and derived objects, the category is, strictly speaking, redundant, since under the Principle of Categorial Type Transparency, it is presumed to be entirely predictable from (a) the type of the interpretation, (b) X-bar theory, and (c) a parametric description of the language specifying position of heads relative to complements. In effect, the CG category "compiles out" this information, in the sense that it represents explicitly information that could be derived. Therefore, the category-interpretation pairs really count as a single level of representation.[30]

Surface Structure does not figure at all as a level of representation in the theory. Although I have described the combinatory derivations that map phonological strings onto such category-interpretation pairs (and vice versa) in terms of structures, I have never predicated any rule or relation over such structures. They are merely a history or record of how an algorithm might get from the string to the interpretation (or vice versa). Although it is convenient to write such structures down and refer to them as Surface Structures, precisely in order to make the point that no rules of domination, deletion, or movement apply to those structures, they do not constitute a grammatical level of representation at all. No rule ever needs to know how a category that it applies to was derived.

It is the combinatory derivations that correspond to Intonation Structure in the extended sense of the term defined above, as well as capturing coordinate structure and the effects of relativization. Surface Structure or derivation in the present sense therefore subsumes some functions of S-Structure, and all those of Intonation Structure, together with some of the role of PF as it is understood in GB. Phonetic Form in present terms really is no more than an abstract specification of speech segments.

The interpretation the derivation associates with a constituent of category S (or any other derived constituent) directly reflects such information-structural distinctions as those between theme and rheme and between focus and ground.[31] Such information-structural elements are evaluated with respect to alternative sets in the contextual database, and they may be discontinuous.

The present realization of Surface Structure as Information Structure conspicuously fails to represent traditional notions of dominance and command, including c-command. However, relations of dominance and command *are* represented in the canonical predicate-argument structure that results from the trivial procedure of normalizing or "β-reducing" the alternative Information Structures yielded by the alternative derivations of a given proposition, as discussed in connection with examples (5) and (6), and as implicitly assumed in derivations throughout the chapter. It follows that all grammatical relations that depend upon c-command, notably including binding and control and such related phenomena as crossover, must be treated as properties of predicate-argument structure, not Surface Structure, a suggestion consistent with the observations of Lasnik and Saito (1992).

By incorporating the finer distinction between focus and background within both theme and rheme, the present grammar opens up further possibilities of addressing a range of questions in semantics that have been explained in terms of various notions of focus (see Chomsky 1971; Jackendoff 1972; Rooth 1985; Rochemont 1986; von Stechow 1991; Jacobs 1991; Hoeksema 1991; Kratzer 1991; Krifka 1991). In particular, one may expect some light to fall on certain phenomena that have been identified in semantic accounts of particles like *only*, which are claimed to "associate with focus" and which, as Jacobs (1991) and Krifka (1991, sec. 4.8) have noted, interact with intonation in puzzling ways. They are exemplified in sentences like the following:

(83) a. Freeman even introduced HARDY_{H*} to $\text{Willis}_{LL\%}$.
b. Freeman only introduced HARDY_{H*} to $\text{WILLIS}_{!H*\ LL\%}$.

One might have expected that availability of quantifier scope alternations in scope-ambiguous sentences like *Some boy admires every saxophonist* might be affected by intonation, since we saw in chapter 4 that scope alternations are limited by syntactic derivation, and we have seen in the present chapter that intonation may limit combinatory derivation. Such an expectation would be in line with the claims of the Prague School that Information Structure determines scope—see Hajičová, Partee and Sgall (1998) for discussion. However, the lexical mechanism for quantifier scope advanced in chapter 4, motivated by examples like (55), makes scope entirely independent of which combinatory derivation is involved, just as long as there is one.

This means that according to the theory of scope ambiguities sketched in section 4.4, the effects of intonation on availability of readings are essentially limited to changes in the relative preference or salience of the readings that the competence grammar makes syntactically and semantically available, a conclusion that appears to be consistent with the observations on both sides of the debate in Hajičová, Partee and Sgall (1998), although it remains somewhat unclear what the facts in this area actually are.

Much further work remains to be done to complete this picture of the interface between grammar and speech. Nothing has been said here about the way metrically related phenomena of rhythm, timing, and lengthening are to be accommodated. (It should be obvious nevertheless that the theory offered here is consistent with more or less any of the available theories.)

Serious difficulties also attend the automatic identification of prosodic boundaries in speech. The phonetic realizations of elements such as pitch accents and boundary tones are subject to coarticulation effects, like all phonological segments, and are hard to recognize. In fact, it is highly likely that the process of identifying them cannot be separated from that of recognizing the words that carry them. This observation might seem daunting, since current techniques for word recognition, although improving dramatically, are nonetheless not very good. However, it is likely that the task of recognizing words and intonation together will turn out to be easier than doing either task in isolation, as Pierrehumbert (1993) points out.

The problem of so-called spurious ambiguity engendered by combinatory grammars now appears in a different light. Although the semantic properties of the rules (notably the associativity of functional composition) indeed allow alternative analyses that are equivalent in terms of the function-argument structure to which their interpretations reduce, the corresponding distinctions in surface constituency are nonetheless meaning-bearing. To call them "spurious" is very misleading, for they are genuine ambiguities at the level of Information Structure. Any theory that actually addresses the range of prosodic phenomena and coordinate constructions considered here must implicate exactly the same nondeterminism.

However, the question remains, how does the parser cope with structural ambiguity in general, and with this kind in particular? Sometimes of course intonation uniquely specifies structure. But very often it does not. PP attachment ambiguities, of the kind exhibited in the following sentence, are not usually disambiguated by intonation:

(84) Put the block in the box on the table.

Moreover, in the discussion in section 5.5.2 of the null tone on unmarked themes, we saw that Information Structure boundaries need not be disambiguated by intonation either.

The pragmatic nature of sentences with unmarked themes actually provides a strong suggestion about the nature of a mechanism for resolving not only the nondeterminism inherent in the null tone, but also other structural ambiguities such as PP attachment.

The null tone is found on the theme precisely when the corresponding theme is entirely ground information—that is, when it is already established in the context and known to the listener, and when nothing else in the context stands in contrast to it. That is to say, this particular ambiguity is only permitted when the theme is already in the listener's model of the discourse. In the case of (55) this means that at successive positions in a left-to-right analysis of the string *Mary wrote a book about BATS*, the property corresponding to *Mary*, *Mary wrote*, and *Mary wrote a book about* can be derived and can be compared with the one(s) present in the model, so that choices between syntactic alternatives such as composing or not composing can be made accordingly. What is more, since the combinatory grammar allows more or less any leftmost substring to be treated as a constituent, complete with an interpretation, there exist very simple parsing algorithms that will permit incremental analysis of this kind, consistent with the strict competence hypothesis.

This may be the most significant practical benefit of the combinatory theory. In the past, syntax and semantics on the one hand, and phonology and discourse information on the other, have appeared to demand conflicting structural analyses and to require processing more or less independently. The present theory shows them to be in harmony. Processors may more easily be devised that use all these sources of information at once, potentially simplifying both problems. The fact that the combinatory notion of syntactic structure and interpretation stands in the closest possible relation both to the prosodic structure of the signal itself and to the concepts, referents, and themes represented in the discourse context should make it easier to use all of these higher-level sources of information to filter out the ambiguities that will inevitably continue to arise from processing at lower levels.

I will return to this architecture and to the question of how to process these grammars in part III (to which the more psycholinguistically or computationally inclined reader might well turn directly). But first it is important to look more deeply into the linguistic justification for the grammars proposed here. Part II presents two related case studies, which examine in detail the extent to which the theory generalizes to more complex constructions, including further varieties of coordination in English and other languages, and to their interaction with quantifier scope and intonation.

PART II

Coordination and Word Order