Chapter 3

The rule permits the following derivation for sentences with the structure of (56):\textsuperscript{28}

\begin{align*}
\text{VP/VP} & \vdash \lambda p. \text{will} p \quad \text{copy} \\
\text{VP/NP} & \vdash \lambda p. \lambda g. \text{without}(\text{file}' g) \quad \text{and}' \\
\text{VP/NP} & \vdash \lambda p. \lambda g. \text{without}(\text{read}' g) \quad \text{reading} \\
\text{NP} & \vdash \text{these articles}' \\
\text{S/VP} & \vdash \lambda p. \text{will} p \quad \text{copy}' \\
\text{VP/NP} & \vdash \lambda p. \lambda g. \text{without}(\text{file}' g) \quad \text{and}' \\
\text{VP/NP} & \vdash \lambda p. \lambda g. \text{without}(\text{read}' g) \quad \text{reading} \\
\text{NP} & \vdash \text{these articles}' \quad \phi
\end{align*}

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

(62) *a [picture of]_NP [by]_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 \textbf{S} satisfying the following condition: \textbf{S}X\text{yc} = x(\text{yc})."

"Why is that bird so important?" asked Craig.

"You will find that out when you reach the Master Forest," replied Bravura.

Raymond Smullyan, \textit{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.\textsuperscript{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 \textit{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 (\textit{<S}_X)

\[ Y/Z \quad (X/Y)/Z \Rightarrow_S X/Z \]

where \( X = S \$ \)

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 \textit{leftward} for an argument of type \( Y \) with a \textit{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:

1. **The Principle of Adjacency:**
   Combining rules may only apply to finitely many phonologically realized and string-adjacent entities.

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

3. **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.

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:

4. **X/Y Y \not\equiv X**

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

5. **X/Y Y/Z \neq X/Z**

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

6. **X/Y \text{ CONJ} X/Y \neq 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. 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 \( Y/Z \) by \( (Y/Z)/\$, \((Y/Z)/\$, etc. It is again assumed that such schemata are limited to a bounded number of arguments:

7. **Functional composition**
   a. \( X/Y \ Y/Z \Rightarrow B X/Z \) \( (>B) \)
   b. \( X/Y \ Y/Z \Rightarrow B X/Z \) \( (>B) \)
   c. \( Y/Z \ X/Y \Rightarrow B X/Z \) \( (<B) \)
   d. \( Y/Z \ X/Y \Rightarrow B X/Z \) \( (<B) \)

8. **Functional substitution**
   a. \( (X/Y)/Z \ Y/Z \Rightarrow S X/Z \) \( (>S) \)
   b. \( (X/Y)/Z \ Y/Z \Rightarrow S X/Z \) \( (>S) \)
   c. \( Y/Z \ (X/Y)/Z \Rightarrow S X/Z \) \( (<S) \)
   d. \( Y/Z \ (X/Y)/Z \Rightarrow S X/Z \) \( (<S) \)

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, \( >B \), and \( <B \), 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
The significance of this departure is as follows.

The composition rules \( \succ \mathcal{B} \) and \( \prec \mathcal{B} \) 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 \( \succ \mathcal{B}_x, \prec \mathcal{B}_x, \succ \mathcal{S}_x, \) and \( \prec \mathcal{S}_x \) 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:\(^8\)

\(\text{(10)}\) The Order-preserving type-raising rules

\[\begin{align*}
a. \quad & \mathcal{X} : a \quad \Rightarrow^\bullet \quad \mathcal{T} / (\mathcal{T} \setminus \mathcal{X}) : \lambda f. f a \\
& \text{where } \mathcal{T} / \mathcal{X} \text{ is a parametrically licensed category for the language} \\
b. \quad & \mathcal{X} : a \quad \Rightarrow^\bullet \quad \mathcal{T} / (\mathcal{T} \setminus \mathcal{X}) : \lambda f. f a \\
& \text{where } \mathcal{T} / \mathcal{X} \text{ is a parametrically licensed category for the language}
\end{align*}\]

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 \( \mathcal{X} \) 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:

\(\text{(11)}\) a man whom I think that Dexter likes

\[\begin{align*}
& \frac{S / (S \setminus NP)}{S / (S \setminus NP)} \prec \mathcal{B} \\
& \frac{S \setminus S}{S \setminus S} \succ \mathcal{B} \\
& \frac{S / (S \setminus NP)}{S / (S \setminus NP)} \succ \mathcal{B}
\end{align*}\]

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):

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

b. [Dexter]_{S / (S \setminus NP)} \text{ and } [Warren]_{S / (S \setminus NP)} \text{ 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 ⇔ 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 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/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\NP requires a
distinct slash-crossing instance of composition, $>\mathbf{B}_x$:

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

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 combinatorial 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 non-extractability 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 $\coloneqq (VP/NP_{+ANT,agr})/(S\backslash NP_{agr})$

Explaining Constraints on Natural Grammar

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” $\Delta$, 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 topocalized 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)$\{(N,N)/(S/NP_{+ANT,3om}) \}$ I think likes marmalade$S/NP_{+ANT,3om}$

(23) *[I think likes marmalade]$S/NP_{+ANT,3om}$ (this very heavy man)$NP_{+ANT,3om}$

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) \(<B_x>\)

\[
\begin{array}{c}
Y/Z \quad X/Y \quad \Rightarrow_B \quad X/Z \\
\text{where } X, Y = SS
\end{array}
\]

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

(25) I shall buy today and cook tomorrow the mushrooms ...

\[
\begin{array}{c}
S/(S/\text{NP}) \\
(S/\text{NP})/V_P \\
V_P/\text{NP} \\
V_P/\text{CONJ} \\
V_P/\text{NP} \\
V_P/\text{VP} \\
V_P/\text{VP} \\
V_P/\text{NP} \\
V_P/\text{NP} \\
V_P/\text{NP} \quad \Rightarrow_B \\
\end{array}
\]

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 give to my sister an engraving by Rembrandt

\[
\begin{array}{c}
S/V_P \\
(V_P/(V_P/P_P)) \\
V_P/(V_P/P_P) \\
V_P/(V_P/P_P) \quad \Rightarrow_B \\
\end{array}
\]

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

(27) (articles) which \(_{}\) I will file tomorrow

\[
\begin{array}{c}
(N/N)/(S/\text{NP}) \\
S/V_P \\
V_P/\text{NP} \\
V_P/\text{NP} \quad \Rightarrow_B \\
\end{array}
\]

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:

(28) a. an engraving which I will \textit{buy today and sell tomorrow}

b. an engraving which I will \textit{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 leftist extract can also rightward extract:

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

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 \textit{SHIFT}, with respect to which all normal arguments are assumed to be unspecified:\textsuperscript{15}

(31) Backward crossed composition I \(<B_x>\):

\[
\begin{array}{c}
Y/Z_{\text{SHIFT}} \quad X/Y \quad \Rightarrow_B \quad X/Z_{\text{SHIFT}} \\
\text{where } X, Y = SS
\end{array}
\]
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 ($B\times$)

\[
Y/Z\_SHIFT,\_ANT \rightarrow_{B} X/Z\_SHIFT,\_ANT
\]

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 [give]$_{(VP/PP)/NP}$ 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. *[Smiled]$_{S/NP}$ [the man in the grey flannel suit]$_{S/(S/NP)}$

b. *[Every friend of]$_{NP/NP}$ [smiled]$_{S/NP}$ [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 Manny 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).

\[
\begin{array}{ccc}
NP & \rightarrow_{T} & (S/NP)/NP \\
\quad : anna' & \rightarrow_{T} & \lambda x.\lambda y.\text{marry}'xy \\
\quad : manny' \\
\quad : T/(\lambda p.p anna') \\
\quad : \lambda q.q\text{manny}'y \\
\quad : S/NP \\
\quad : \lambda y.\text{marry}'\text{manny}'y \\
\quad : S : \text{marry}'\text{manny}'anna'
\end{array}
\]

\[
\begin{array}{ccc}
NP & \rightarrow_{T} & (S/NP)/NP \\
\quad : anna' & \rightarrow_{T} & \lambda x.\lambda y.\text{marry}'xy \\
\quad : manny' \\
\quad : T/(\lambda p.p anna') \\
\quad : \lambda q.q\text{manny}'y \\
\quad : S/NP \\
\quad : \lambda x.\text{marry}'x\text{anna}' \\
\quad : S : \text{marry}'\text{manny}'anna'
\end{array}
\]

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 Sur-
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” and y resembling an anaphor bound to the (interpretation y of the) subject of the control verb. 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\NP\agr)/NP : λx.λy.λx'λy'
y
   b. to := (S\inf\NP\agr)/(S\inf\NP\agr) : λx.x

(40) Dexter tries to like Warren

\[ NP_{3sm} \rightarrow (S\NP)_(S\inf\NP\agr)/(S\inf\NP\agr) : λx.λy.λx'λy' y \rightarrow NP \\
\rightarrow (S\NP)_(S\inf\NP\agr) : λx.λy.λx'λy' y \rightarrow NP \\
\rightarrow (S\NP) : λx.λy.λx'λy' y \rightarrow NP \\
\rightarrow S : trial' (like'warren'(and'dexter'))dexter' \]

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

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\NP)\NP\agr : λp.λy.\try'(p(\and'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” and y resembling an anaphor bound to the (interpretation y of the) subject of the control verb. 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\NP\agr)/NP : λx.λy.λx'λy'
y
   b. to := (S\inf\NP\agr)/(S\inf\NP\agr) : λx.x

(40) Dexter tries to like Warren

\[ NP_{3sm} \rightarrow (S\NP)_(S\inf\NP\agr)/(S\inf\NP\agr) : λx.λy.λx'λy' y \rightarrow NP \\
\rightarrow (S\NP)_(S\inf\NP\agr) : λx.λy.λx'λy' y \rightarrow NP \\
\rightarrow (S\NP) : λx.λy.λx'λy' y \rightarrow NP \\
\rightarrow S : try'(like'warren'(and'dexter'))dexter' \]

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\NP)/(S\NP))/NP, reflecting the assumption that predicate-argument structures observe the obliqueness hierarchy:

(41) persuades := ((S\NP\agr)/(S\inf\NP\agr))/NP\agr : λx.λp.λy.persuaded(p(\and'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:
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.

(42) showed := ((S\NP)/NP)/NP : \ax.\ay.\ax.\ay.\ax.\ay.show/\ax/\ay

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

(43) might := (S\NP)/VP : \lp.\ax.\ay.might/\lp/\ax

The following kind of derivation results:

(44) Anna

(45) marry

To handle binding, raising, and control lexically via a level of interpretation of 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 Shaumyana (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 primarily concerned with the unbounded constructions, it diverges from many of its categorical 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 categorical 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 standing 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. \( \forall x. \text{boy} \, x \rightarrow \exists y. \text{saxophonist} \, y \land \text{admires} \, y \, x \)
   b. \( \exists y. \text{saxophonist} \, y \land \forall x. \text{boy} \, x \rightarrow \text{admires} \, y \, x \)

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 function 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 := (T/(T/NP))/N : \lambda p. \lambda q. \forall x. p \, x \rightarrow q \, x
    every := (T/(\text{T/NP}))/N : \lambda p. \lambda q. \forall x. p \, x \rightarrow q \, x

(51) some := (T/(\text{T/NP}))/N : \lambda p. \lambda q. \exists x. p \, x \land q \, x
    some := (T/(\text{T/NP}))/N : \lambda p. \lambda q. \exists x. p \, x \land q \, x

Given the categories in (50) and (51), the alternative derivations that CCG permits will deliver the two distinct Logical Forms shown in (49), entirely
(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 at 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/theme. The corresponding structural

![Figure 4.1](attachment:image.png)

Architecture of a Combinatory Grammar, 1

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 monostral, in the sense that it builds only a single level of structure, namely, Logical Form.
18. The same generalization is implicit in the coordination rule (20), whose semantics
was defined by Curry’s schematic combinator $\Phi^n$. We will see that both families of
combinatory rules must be limited to bounded $n$.

19. T is also the remaining prominent consonant besides B and S in the word substitution
(see note 11).

20. This remark should not be taken as assuming that the relation between morphologi-
cal case and grammatical relations like subject is invariably straightforward. Icelandic
(Rognvaldsson 1982; Bresnan and Thráinsson 1990, 361-362) is a frequently-cited
example that reveals some complexities that are passed over here.

21. Dowty (1996) has since pointed out that this analysis is incompatible with his
(1979) and Bach’s (1979) analyses of binding and control, and in fact he now disowns
it. However, other analyses of the latter phenomena are discussed below.

22. I omit certain details here, including agreement, which both categories must in-
clude. See Steedman 1996b (which unhelpfully uses the $-$ convention to schematize
over the last three categories) for details.

23. Since multiple extractions involve non-string-peripheral arguments, they require a
generalization to the composition rules that will be discussed in chapter 4.

24. The suggestion that island phenomena are related to semantic interpretation goes

25. This property is shared by the GPSG analysis of parasitic gaps in Gazdar et al.
1985.

26. The analysis requires us to assume that even PPs that are apparently not subcate-
gorized for, like the one in I folded the rug over the painting, must in some sense be
arguments PP rather than adjuncts VP $\backslash$ VP. Otherwise, the following example would
be accepted with an analysis parallel to (55), to mean I folded the rug over itself:

(i) * a rug which I will fold $\backslash$ VP $\backslash$ NP [over] $\backslash$ VP $\backslash$ VP $\backslash$ NP

The unacceptability of this example is discussed further in chapter 4, but the idea that
the PP is in some sense an argument is borne out by the fact that such prepositions can
strand and therefore must be “composable into,” as in the following example:

(ii) the picture which I will fold the cloth $\backslash$ PP [over] $\backslash$ PP $\backslash$ NP

27. If the rule could apply with $X$ equal to $N$, native speakers would accept the follow-
ing with the meaning a good dog with a dog:

(i) *a $\backslash$ NP $\backslash$ N [with a] $\backslash$ N $\backslash$ N [dog] $\backslash$ N

28. Infinitival and gerundival predicate categories are abbreviated as VP and $\backslash$ VP $\backslash$ NP$^*$
and the raised NP object as NP$^1$, for ease of reading. Again details of the predicate-
argument structure are simplified (see Steedman 1996b for a fuller account).

Chapter 4

1. That is not to say that further explanations might not be found—say, in arguments
from learnability. But the onus of providing them remains upon us.

2. The Principle of Adjacency defined here is distinct from the similarly named principle
used in Dependency Grammar (see Robinson 1970).
3. In terms of standard transformationalist theory, this principle is equivalent to a ban on the use of variables in transformations.

4. This position is closely related to the notion of "directionality of government" in Kayne 1983.

5. Identification of the individual instances that the generalization gives rise to for \( n \) up to about 3 is suggested as an exercise. It is easy to see that each combinator \( X^n \) makes \( 2^n \) syntactic instances available to Universal Grammar. However, the parametric specification of the lexicon of any given language will mean that many of these rules cannot apply in that language. The fact that there appear to be strong learnability-related constraints on possible lexicons for human languages, such as tendencies toward consistent head finality or head initiality, will further restrict the range of rules in any real language.

6. The categories given in (43) of chapter 3 for the relative pronoun is an example of a higher-order functor that induces a word order that would not otherwise be allowable. Dowty (1988) and Steedman (1987) stated the order-preserving property too generally, as Milward (1991, 1994) points out. Milward offers an example that purports to show that rule \(<B\) causes an overgeneration in a fragment of English that includes nonrestrictive relatives, allowing (ib) as well as (ii) with the categories shown.

(i) a. [John\_NPNP [who speaks Russian\_NPNP \( x \) NP\_NPNP \( x \) \( \text{reluctantly} \) \( x \) \( S\_SNP \)_NP\_NPNP \( x \) \( \text{came} \) \( x \) \( S\_SNP \)_NP\_NPNP]

    b. *(John\_NPNP [reluctantly] \( x \) \( S\_SNP \)_NP\_NPNP \( x \) \( \text{who speaks Russian} \) \( x \) \( \text{NP\_NPNP \( x \) \( \text{came} \) \( x \) \( S\_SNP \)_NP\_NPNP])

Although the example certainly shows that composition may induce new word orders, the category \( \text{NP\_NPNP} \) for nonrestrictive relatives is not necessarily incorrect. In fact, the general immunity of nonrestrictives to constraints on movement makes it likely that they should not be regarded as "in construction" at all, but rather as a variety of parenthetical. The ill-formedness of (ib) can therefore be regarded as proof that the \( \text{NP\_NPNP} \) category is incorrect, rather than as raising a problem for CCG.

7. See chapter 8 and Vijay-Shanker and Weir (1990, 1994) for discussions of power and complexity of CCG, including a polynomial time worst-case complexity parsing result under the assumptions used here.

8. Under at least some formalizations of the Principles of Consistency and Inheritance, including the one offered in chapter 8, two further non-order-preserving instances of type-raising are potentially allowed:

(i) The Non-order-preserving type-raising rules

   a. \( X \Rightarrow T \) \( T(T(X)) \)

   b. \( X \Rightarrow T \) \( T(X(T)) \)

If a construction is to be classified as configurational at all, it must entirely exclude such non-order-preserving instances of argument type-raising. We will see in chapter 8 that there is a natural formulation of the Principle of Consistency under which languages must either entirely exclude syntactic non-order-preserving type-raising or must sacrifice configurationality. There is no middle ground. Categories such as relative pronouns, which as we saw in (43) of chapter 3 are related to but distinct from non-order-preserving type-raised categories, are lexically unique words, strikingly prone to case marking. Because of their unique position and intonational markedness, it is assumed that English topicalized arguments are also in effect lexically special items, such as \( \text{TOPIC}/(S/X) \), distinguished by intonation or the related comma in written sentences like "This law, the Supreme Court has ruled unconstitutional. We might plausibly assume that this special category is assigned by a unary rule that only applies to the leftmost item in the sentence, as proposed in Steedman 1987.

9. Whether the specific overgeneralizations identified by Houtman can all be dealt with in this way this way depends on a more formal definition of the constraint than has been given here, and upon some open questions that we will not go into here about his examples, which involve numeric determiners that are known to be special cases.

10. Examples like "Dexter, and I think Warren, are geniuses but better, only to the extent that I think can be read as parenthetical (as agreement reveals).

11. This problem in CCG is more restricted than "Dekker’s paradox" concerning the Lambek calculus. Dekker noticed that two legitimate English categories \( \text{NP\_NPNP} \) and \( S/S \) in the Lambek calculus come to have the same category \( (S/(S\_SNP))/NP \), giving rise to anomalous conjunctions like the following (Dekker 1988, cited in Houtman 1994, 85–89).

   (i) *The brother of, and John believes that, Pete slept.

Since CCG lacks the Geach “Division” rule, it does not fall prey to Dekker’s overgeneration.

12. This question, including the alternative proposal of Perlmutter (1971), which relates exemption from the Fixed-Subject Condition to the parameter of pro-drop, is discussed at further length in Steedman 1996b.

13. More precisely, the relation that such categories bear to the basic \( \text{VP/S} \) categories is a first cousin to the Slash Termination Metarule 2 of Gazdar et al. (1985) (see Hukari and Levine 1987; Hople 1990, 59). The analysis differs from that presented in Steedman 1987; see Steedman 1996b for further discussion of its ancestry.

14. This restriction (which is discussed in more detail in Steedman 1996b) prevents \( \text{the} \) \( \text{NP\_NPNP} \) \( \text{good} \) \( \text{N\_NPNP} \) \( \text{boy} \) from meaning the good boy with a toy.

15. The feature–value +\( \text{SHIFT} \) seems to be related to the notion “most oblique.” The analysis given here for the restrictions on heavy NP shift differs in minor details from the one given in Steedman 1996b, 68–69.

16. Such rules raise obvious problems for the theory of language acquisition, which are briefly discussed in chapter 10. They appear as no more than technical solutions.

17. However, if type-raising were lexical, so that every were \( \text{NP\_NPNP} \), then the restriction would follow without stipulation, since the predicate is not a raisable category.

18. It is assumed here that the level of interpretation in question is neutral with respect to non-argument-structure-dependent aspects of meaning such as quantifier scope, which is discussed in section 4.4.

19. Here, this mechanism is represented by \( \lambda \)-binding of variables. But this mechanism should not be confused with the combinatory rules that establish the long-distance dependency itself. These variables just represent the normal mechanism for binding all arguments to predicates.
20. A more complete account of anaphor binding in CCG is offered in Steedman 1996b. The idea that control verbs involve implicit anaphors goes back at least to Helke (1971) and has more recently been proposed by Manzini (1993).

21. Clark’s analysis replaces the eccentric proposal to treat auxiliaries as control verbs in Steedman 1996b.

22. The latter is an example of Taraldsen’s (1979) anti-c-command condition on parasitic gaps, discussed in Steedman 1996b.

23. For example, in order to obtain the narrow scope object reading for (55b), Hendriks (1993), subjects the category of the transitive verb to “argument lifting” to make it a function over a type-raised object type, and the coordination rule must be correspondingly semantically generalized.

24. Technically, this analysis raises questions which we will pass over here about the status of expressions like $\delta_{\text{max},y}$ with respect to the binding theory as it is defined in Steedman 1996b, since they can be bound but do not count as pro-terms for the purposes of c- or f-command.

25. Similar considerations give rise to apparent wide and narrow scope versions of the existential donkey in (56).

26. I am grateful to Gann Biemer for discussions on this problem.

27. The exact facts are hard to pin down in this area, and some judges claim to get the dependent readings. In the terms of the present theory this may mean that they have true quantifier meanings for upward monotone quantifiers. One case where almost everyone seems to get a dependent reading is the following, pointed out to me by Yoad Winter:

(i) A flag was hanging in front of at most two/exactly three/at least four windows.

However, the indefinites that give the appearance of dependency in this way seem to be quite restricted, and may be confined to entities that are inherent duplicates, like flags and books, even when they are definite. So we get

(ii) a. The American flag was hanging in front of at most two windows.
    b. A copy of The Little Red Songbook was seized from at least three bathrooms.
    c. A woman was waving from at most two/exactly three/at least four windows.

28. The corresponding example for this point in Steedman (1999) is in error.

29. The architecture is therefore close to that implicit in Chomsky (1993, 1995), although the idea is implicit in much earlier systems, including Chomsky’s own (1957), Montague Grammar, Woods’s (1973) ATN grammars, and earlier versions of the present theory.

Chapter 5

1. This chapter is a completely revised and reworked version of Steedman (1991a).

2. The terms “theme” and “rheme” are taken from Halliday (1967b, 1970), although I follow Lyons (1977) and Bolinger (1989) in rejecting Halliday’s requirement that the theme be sentence-initial. I also leave open the possibility that an utterance may involve multiple or “discontinuous” themes and/or rhemes.

3. Wilson and Sperber (1979) and Prince (1986), define the notion here identified as