Notes

Chapter 1

1. The HOLD-register analysis of *wh*-movement was in part anticipated in earlier work by Thorne, Bratley and Dewar (1968), who called their register *.

2. Wood (1993) provides a useful review of theories by Lambek (1958), Ades and Steedman (1982), Bach (1979), Dowty (1979), Steedman (1987), Oehrle (1988), Hepple (1990), Jacobson (1990, 1992b), Szabolcsi (1989, 1992), and Wood (1988), although my colleagues should not be assumed to endorse all the assumptions of the version that is outlined here. The present proposal is more distantly related to a number of other generalizations of the early categorial systems of Ajdukiewicz, Bach, Bar-Hillel, Dowty, Lambek, Geach, Lewis, Montague, van Benthem, Cresswell, and von Stechow, to many of which the conclusions of this book also apply. In particular, Oehrle (1987), Moortgat (1988a), and Morrill (1994) explicitly relate Lambek-style categorial grammars to prosody.

3. Marr expressed some doubt about whether natural language is in fact a modular system, apparently because he was aware of the way knowledge and inference interact with language understanding. I will argue against this conclusion in chapter 9.

Chapter 2

1. This claim should not be taken as denying that such learning can be usefully thought of in terms of supervised machine learning techniques, or as excluding the possibility that the substrate of such conceptual representations may be associative or probabilistic.

2. The "Standard Theory" presented in Chomsky 1965 did not explicitly recognize any level of Logical Form distinct from Deep Structure. However, had it done so, it would have had to derive it from Deep Structure. The fact that later "Extended," "Revised Extended," and "Principles and Parameters" or "Government-Binding (GB)" versions of Chomsky's theory derived Logical Form from a level called "S-Structure" should not be allowed to confuse the point. S-Structure is not the same as Surface Structure, as will become clear when this level is discussed in more detail below. The rather different view of Logical Form sketched in Chomsky 1971 is discussed in chapter 5.

3. There are a number of well-known exceptions to this generalization, which I will

discuss in chapter 4. The most obvious one in English arises in subject extraction.

4. Certain apparent exceptions to Ross's generalization are discussed in chapter 7.

5. The proposal is also implicit in the approach of Cresswell (1973) and von Stechow (1991) based on "structured propositions."

6. Selkirk's proposal in some respects constitutes a descendant of the proposal in Chomsky 1971 to derive aspects of Logical Form relating to focus from Surface Structure, a level corresponding to what in more modern versions of Transformational Grammar is called Phonetic Form.

7. The indicated dependencies are those between semantically related arguments and predicates, rather than surface dependencies between verbs and NP arguments that would be attributed on a VP analysis of the construction. However, in either case the Dutch dependencies cross.

8. A number of apparent exceptions to Ross's generalization have been noted in the literature, including Zapotec (which is VSO but allows SO and VSO; see Rosenbaum 1977) and German (which is SOV but allows SOV and SO; see van Oirsouw 1987). These are discussed in chapter 7, where I show how they are made possible by the fact that Zapotec (unlike Irish) allows SOV as a main-clause word order, while German and Dutch (unlike Japanese) allow VSO/SVO as main-clause orders. I restate Ross's constraint in terms of overall order properties of languages and constructions rather than any notion of "underlying" word order.

9. Languages that order object before subject are sufficiently rare as to apparently preclude a comparable data set, although any result of this kind would be of immense interest.

10. This Categorial broad church is also closely related to the Dependency Grammar tradition of Mel'čuk and Pertsov (1987), Hudson (1984), and the Prague School (Hajičová and Sgall 1987, 1988), with which it shows some sign of merging (see Pickering and Barry 1993 and Milward 1991, 1994).

11. Wood 1993 has already been mentioned as a helpful guide. Moortgat 1997 and Steedman 1993 offer more partisan reviews. The collections edited by Buszkowski, Marciszewski and van Benthem (1988) and Oehrle, Bach and Wheeler (1988) are important. The reader is warned that because this body of work addresses several distinct and sometimes non-overlapping concerns, there is a variety of notations for categorial grammars.

Chapter 3

1. This chapter and chapter 4 are for the most part a review of earlier work. They can be skimmed or skipped entirely by those already familiar with the approach. For further details see Steedman 1996b.

2. Other superficially similar but different notations are used by some of the other authors referred to here. The present notation has the advantage of maintaining a consistent order of range and domain across the page for leftward and rightward functions. Because of the present concern with semantics, and with comparisons across languages with similar semantic types but different word orders, this consistency is crucial to

readability. It will be apparent in chapter 4, where the full theory is presented, that the notation used here is an abbreviation for a notation in which directionality—that is, relative position—is a property of the argument of a function.

3. This principle and the following Principles of Head Categorial Uniqueness and Categorial Type Transparency replace the Principle of Categorial Government in Steedman 1996b.

4. Interestingly, pure Categorial Grammar is only "weakly" equivalent to Context-Free Phrase Structure Grammar (CFPSG). That is, although we can write a pure categorial grammar to recognize any context-free language, it will not necessarily be possible to write one that will deliver the same trees as a given CFPSG for the same language (see Bar-Hillel, Gaifman and Shamir 1964). For real natural languages, because of the close dependence of syntactic categories on the same semantics as the corresponding CFPSGs, this seems in practice never to be a problem.

5. Some caution must be exercised in invoking feature-matching mechanisms of this kind. In particular, if features are allowed to take unbounded structures as values, then their use can lead to very unconstrained theories indeed (Carpenter 1991). All features used here are simple, in the sense that they do not take unbounded structure as values.

6. However, the theory expressly does not exclude the possibility of a natural language in which the interpretation for the corresponding transitive verb is $\lambda y.\lambda x.marry' xy$, a term in which the λ-binders for the same predicate-argument structure are not eliminable in this way.

7. I do not exclude the possibility of accounting for binding phenomena without such an "intrinsic" use of Logical Form, by using nonconcatenative rules, such as the WRAP rules of Bach (1979, 1980), Dowty (1979), Jacobson (1990), and others. However, such rules appear to complicate the analysis of coordination considerably.

8. The notion "like category" is of course problematic for any theory of syntax, as well-known examples like the following reveal:

(i) Pat is a Republican and proud of it.

I will ignore such problems here, assuming that some finer-grained specification of categories can be applied to the present theory. See Sag et al. 1985, Morrill 1994, and Bayer and Johnson 1995 for discussion of some alternatives.

9. The natural expedient for a categorial approach might therefore seem to be to eschew such syncategorematic rules and drive coordination off the lexical category of the conjunct, following Lambek (1958, 1961), who associated the categorial type in (i) with sentential conjunctions like *and*, where the lexical category itself includes a variable ranging over categories of the same type, written X:

(i) and:= $(X\backslash X)/X$

However, unless further refined, such a category will cause the grammar to overgeneralize, because the constituents such as $[and\ he\ talks]_{s\backslash s}$ that this analysis induces do not behave like other categories with respect to the rules in the rest of the grammar. In particular, the rule of "backward composition" that is central to the analysis of example (40), here given as (iia), interacts with this analysis to wrongly permit (iib).

(ii) a. I gave the teacher an apple and a policeman a flower.
b. *a man who walks and he talks

Since as we will see the rule involved is a theorem of the Lambek calculus, this observation applies to the entire family of "associative" generalized categorial grammars.

10. The annotation Φ^n on the reduction arrow in the rule is inspired by the combinatory notation of Curry and Feys (1958). Although as usual the semantic details can safely be ignored, the dots schematize over the following family of functionals combining the interpretation of the conjunction *b* with the two argument interpretations, following Partee and Rooth 1983.

(i) $\Phi^0 bxy \equiv bxy$
$\Phi^1 bfg \equiv \lambda x.b(fx)(gx)$
$\Phi^2 bfg \equiv \lambda x.\lambda y.b(fxy)(gxy)$
$\Phi^3 bfg \equiv \lambda x.\lambda y.\lambda z.b(fxyz)(gxyz)$
$\Phi^4 bfg \equiv \lambda x.\lambda y.\lambda z.\lambda w.b(fxyzw)(gxyzw)$

They correspond to the different instances of coordination discussed here, and (for English at least) we can assume a bound of $n = 4$, which (as noted earlier) seems to be the highest valency in the English lexicon. The ...convention is therefore merely an expository convenience. Every theory must schematize its constituent coordination rule semantically in the same way.

11. Curry and Feys (1958, 184 fn.) note that Curry called the operation **B** because that letter occurs prominently in the word *substitution*, and because the names **S** and **C** were already in use.

12. The equivalence sign in these definitions is supposed to indicate that the combinators are *primitives*, not that they are defined or interpreted in terms of the abstraction operator λ.

13. We could at some cost to readability write the semantics entirely in combinatory terms, without any use of variables, as Shaumyan (1977) does.

14. There is a lot of semantic detail here that can be skipped. It is fairly obvious that if we have an interpretation for the function *might* (say, as a set of ordered pairs from complement VPs and tensed predicates) and an interpretation for the function *marry* (in the same terms, a set of ordered pairs from NPs like *Manny* to VPs like *marry Manny*), then we know everything necessary to construct a new function from NPs to predicates (e.g. as an ordered set of pairs like *Manny* and *might marry Manny*).

15. The predicate-argument structures as they are given here are simplified and leave out certain details relevant to the implementation of the binding theory and control in the present framework. These further details are made explicit in Steedman 1996b.

16. Though slightly clumsy, the example is parallel to ones that Abbott (1976) shows to be grammatical in English. The related *wh*-extractions like *the policeman to whom I offered and may give a flower*, which (as we will see) hinge on the same generalization of composition, are impeccable.

17. In some early papers the corresponding schema was defined slightly more generally, leading to some overgeneration, which I am grateful to Glyn Morrill for calling to my attention.

18. The same generalization is implicit in the coordination rule (20), whose semantics was defined by Curry's schematic combinator Φ^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 morphological case and grammatical relations like subject is invariably straightforward. Icelandic (Rögnvaldsson 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 include. 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 back to Oehrle 1974, Rodman 1976, and Cooper 1982.

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 subcategorized 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 $[\text{fold}]_{VP/NP}$ $[\text{over}]_{(VP\backslash VP)/NP}$

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 $[\text{I will fold the cloth}]_{S/PP}$ $[\text{over}]_{PP/NP}$

27. If the rule could apply with X equal to N, native speakers would accept the following with the meaning *a good dog with a dog*:

(i) $*[\text{a}]_{NP/N}$ $[[\text{good}]_{N/N}$ $[\text{with a}]_{(N\backslash N)/N}]_{N/N}$ $[\text{dog}]_N$

28. Infinitival and gerundival predicate categories are abbreviated as *VP* and VP_{ing}, and the raised NP object as $NP^{\uparrow}$, 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 allowed. 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 $<\mathbf{B}$ causes an overgeneration in a fragment of English that includes nonrestrictive relatives, allowing (ib) as well as (ia) with the categories shown.

(i) a. [John,]$_{NP}$ [who speaks Russian,]$_{NP\backslash NP}$ [reluctantly]$_{(S\backslash NP)/(S\backslash NP)}$ [came]$_{S\backslash NP}$
b. *[John]$_{NP}$ [reluctantly,]$_{(S\backslash NP)/(S\backslash NP)}$ [who speaks Russian,]$_{NP\backslash NP}$ [came]$_{S\backslash NP}$

Although the example certainly shows that composition may induce new word orders, the category $NP\backslash NP$ for nonrestrictive relatives is not necessarily correct. 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 $NP\backslash NP$ 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. $X \Rightarrow_{\mathbf{T}} \mathrm{T}/(\mathrm{T}/X)$ $(>\mathbf{T}_{\times})$
b. $X \Rightarrow_{\mathbf{T}} \mathrm{T}\backslash(\mathrm{T}\backslash X)$ $(<\mathbf{T}_{\times})$

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 $S_{TOPIC}/(S/X)$, 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 case.

10. Examples like *Dexter, and I think Warren, are geniuses* are better, but 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 NP/NP and S/S can in the Lambek calculus come to have the same category $(S/(S\backslash NP))/NP$, 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 VP/S categories is a first cousin to the Slash Termination Metarule 2 of Gazdar et al. (1985) (see Hukari and Levine 1987; Hepple 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 $[\text{the}]_{NP/N}$ $[\text{good}]_{N/N}$ $[\text{with a toy}]_{N\backslash N}$ $[\text{boy}]_{N}$ from meaning *the good boy with a toy.*

15. The feature-value $+SHIFT$ 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 are offered as no more than technical solutions.

17. However, if type-raising were lexical, so that *every* were $NP^{\uparrow}/N$, 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 λ-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 $sk'_{sax_1}y$ 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 Bierner 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

theme in terms of the related but technically distinct (and as Jackendoff points out less appropriate) notion of "open proposition," or proposition including an unbound free variable.

4. Oehrle (1987), Moortgat (1988b), Morrill (1994), and Hendriks (1994, 1998) explicitly relate other Lambek-style varieties of categorial grammars to prosody.

5. Dwight Bolinger and Julia Hirschberg in personal communications have at least half convinced me that there are circumstances under which one or the other is allowed. However, the only claim I make is that if such circumstances exist, they are such as to make *both* more felicitous.

6. The conspiracy between prosodic phenomena and the notion of constituency that emerges from related grammars including associative operations is noted in Steedman (1985, 540), and by Moortgat (1987, 1988a), and Oehrle (1988). Related points concerning "bracketing paradoxes" in morphology are made by Hoeksema (1984), Hoeksema and Janda (1988), Moortgat (1988a,b), and Morrill (1994). See also other categorial analyses of Wheeler (1981, 1988) and Schmerling (1981).

7. Steedman (2000) generalizes the system to the remaining pitch accents and boundary tones, and explores further the discourse functions of the tunes considered here.

8. These were originally written L- and H-, a usage that continues in the ToBI notation (Silverman et al. 1992). (ToBI stands for "Tones and Break Indices".) Here I will follow the notation in Pierrehumbert and Beckman (1988).

9. Neither Pierrehumbert's theory nor its combinatory expression below should be taken as implying that the null tone corresponds to an absence of pitch. Nor does either version imply that an element bearing the null tone is always realized with the same intonation contour. They merely imply that the intonation is independently specified. It follows that the null tone may carry information about what pitch accents and other tones are downstream of it. As in the case of downstep, it also follows that a processor might make use of this information.

10. For the moment, the distinction between the intonational phrase proper and what Pierrehumbert and her colleagues call the "intermediate" phrase is ignored. These categories will be seen to differ with respect to boundary tone sequences.

11. The reason for notating the latter boundary as LL% rather than L again has to do with the distinction between intonational and intermediate phrases, to which I will return.

12. See Veltman 1984, Heim 1983, Harel 1984, Landman 1986, Groenendijk and Stokhof 1990, Gabbay, Hodkinson and Reynolds 1994, and Steedman 1997 for discussion of various "dynamic" logics that can be used to formalize the notion of updates and side effects.

13. Discourse models of the kind just sketched can be more or less directly realized using logic programming languages such as Prolog. A model of this kind has been investigated in some detail and implemented computationally by Prevost (1995).

14. It will be recalled from chapter 3 that *ana'x* is the equivalent of a PRO controlled subject.

15. An alternative prosody, in which the contrastive tune is confined to *Anna*, seems

15. An alternative prosody, in which the contrastive tune is confined to *Anna*, seems equally coherent in this context. In Steedman 1991a, I argue that this alternative is informationally distinct and arises from an ambiguity about whether the theme established by the question is *Anna* or *who Anna married*. It too is accepted by the rules below.

16. It is important to know that the term "focus" is used in the literature in several conflicting ways. The present use is common among phonologists, who use the term simply to denote the material marked by the pitch accent(s). Grosz and Sidner (1986), use it to denote something like topic, or theme in present terms. Other authors such as Chomsky (1971), Jackendoff (1972), Gussenhoven (1983), Hajičová and Sgall (1987, 1988), Vallduví (1990), Lambrecht (1994), Erteschik-Shir (1998), and Zubizarreta (1998), in different ways confine its use to the rheme. Still other authors, notably Selkirk (1984), Rooth (1985, 1992), Jacobs (1991), Krifka (1991), É. Kiss (1998), and Rochemont (1998), invoke "two levels" of focus, using the term to cover both comment/rheme *and* phonological focus.

17. The difference between pre- and postnuclear material is underlined by the fact that it is also possible to have an H* pitch accent on *woman* in this example, but not on *musical*.

18. This algorithm is related to one proposed by Dale and Haddock (1991), and forms a central module of the generation system described in Prevost and Steedman (1994).

19. There is a question about how to notate the tune of this example that we will return to below.

20. In this I follow Bird (1991) and Prevost (1995). As noted earlier, there are several other pitch accents in Pierrehumbert's system that are not explicitly covered here. See Steedman (to appear) for discussion.

21. This could be made explicit in the notation by writing the category for the verb with null tone as $(S_\eta \backslash NP_\eta)/NP_\eta : ate'$, with η a variable ranging over θ and ρ.

22. The categories given here constitute a modification to previous versions of the present theory that brings it more closely into line with the proposals in Pierrehumbert and Hirschberg 1990. The idea that boundaries are in categorial terms functors was proposed in Prevost and Steedman 1994; Prevost 1995 and independently by Kirkeby-Garstad and Polgárdi (1994).

23. The three remaining boundary-tones of Pierrehumbert's system, H, HH% and HL%, are omitted because they do not arise in the limited varieties of discourse considered here. They are discussed in Steedman to appear.

24. The present treatment of boundary tones, as exhibited in this derivation, shows how it is possible for a pitch accent and a boundary tone to affect the same word in the string. This detail was left unspecified in earlier work, as Kirkeby-Garstad and Polgárdi (1994) point out in a critique of the earlier papers. The symbol *the'* is no more than a placeholder for a proper semantics of the determiner.

25. In effect, the update action first abstracts over the * constant in the new theme. Higher-order unification provides one mechanism for this task, via the algorithm proposed by Huet (1975), in a manner discussed by Shieber, Pereira and Dalrymple (1996).

Prevost (1995) discusses a related device.

26. Relativization and coordination reveal that the sentences violate island constraints.

27. The prosodic annotation of this example represents a minor departure from Pierrehumbert, whose theory does not permit boundaries without corresponding pitch accents, and who would regard the whole tune as a single H*LL% intonational phrase. The present analysis is quite close to one proposed in a different notation by Bing (1979).

28. A multiple H* pitch accent on the VP is an even more appropriate response and also serves to distinguish these contexts, since it is not appropriate to (69).

29. This freer notion of Surface Structure may also explain some of the examples that Bolinger (1985) has used to argue for an entirely autonomous, lexically oriented account of accent assignment, and that Gussenhoven (1985) has used to argue for a similarly autonomous focus-based account. It may also allow us to eliminate some of the nonsyntactic string-based rules and "performance structures" that Cooper and Paccia-Cooper (1980), Gee and Grosjean (1983), and Croft (1995) have proposed to add to the syntax-driven model.

30. However, it would be a great mistake from a processing point of view—and from the point of view of intelligibility of the present presentation—not to compile out this information, technically redundant though it may be, for it has to be used on a great many occasions.

31. As noted earlier, there is a close resemblance here between Information Structures and the structured-meanings of Cresswell (1973), von Stechow (1991), and Chierchia (1989).

Chapter 6

1. Sections 6.1–6.6 of the present chapter are a complete revision and reworking of part II of Steedman 1985, and together with sections 6.7–6.9 extend the coverage to a considerably larger fragment. The earlier paper used an early version of CCG in which slashes were nondirectional, forward and backward application were called "forward and backward combination," and composition was, following still earlier work with Tony Ades (Ades and Steedman 1982), called "partial combination." This chapter imports the earlier analysis into the more modern framework of this book. In the intervening years the basic analysis of the earlier paper has been criticized and improved upon in a number of different frameworks. The earliest and most influential of these developments was Johnson 1988, chaps. 3 and 5, which led to a number of related proposals within LFG, HPSG, and related unification-based grammar frameworks (see Netter 1988; Hepple 1990; Paritong 1992; Bouma and van Noord 1994; Reape 1994, 1996; Hinrichs and Nakazawa 1994). Many of these proposals have been influenced by Moortgat's similar proposal (discussed in Moortgat 1988a, and developed independently at approximately the same time as the present one), which used functional composition as a lexical, rather than syntactic, operation. Along the way, many of these authors and others such as de Roeck (1984), Janeaway (1991), Houtman (1994), Versmissen (1996), and van Noord and Bouma (1997) have pointed to shortcomings in the original account. I have tried to acknowledge and act upon these

criticisms in the course of the chapter. I am grateful to Gann Bierner, Mimo Caenepeel, Angeliek van Hout, Susanne Kronenberg, Jan van Kuppevelt, Marc Moens, and Annie Zaenen for help with revisions to the grammar and/or informant judgments.

2. Evers (1975, 51, 55) following Bech (1955) states that all but the two most deeply embedded verbs in sentences including tensed auxiliaries and multiple infinitives may occur in the Dutch tensed-first order, requiring forward composition within the verb group, as in (i), potentially inducing crossed dependencies in German:

(i) ...daß man ihn hier wird können lassen liegen bleiben.
...that one him here will be-able leave lie stay
'...that one will be able to let him stay lying here.'

Certain dialects of German appear to allow *sah füttern helfen* for (6), and some Swiss dialects discussed by Shieber (1985) and Cooper (1988) even appear to allow the full Dutch order, that is, the equivalent of *sah helfen füttern*.

3. There is an increasing processing load that makes such multiple embeddings increasingly unacceptable. By well-known arguments (see Chomsky and Miller 1963), such considerations are irrelevant to questions of grammaticality.

4. See Seuren 1985 for a discussion of certain problems that Dutch clitic pronouns posed for the earlier version of this theory.

5. S_{+SUB} is an abbreviation for $S_{+SUB}^{-CP,+IP}$, a tensed subordinate bare complement clause or IP. *VP* is an abbreviation for the infinitival predicate category $S_{?SUB}^{-CP,-IP}\backslash NP$, unmarked on the feature *SUB*. The categories given here for infinitival complement verbs differ significantly from those in Steedman 1985. The present analysis of bare-infinitival complement-taking verbs as object control verbs follows Johnson 1988, 120.

6. I assume that the awkwardness of passives like ?*Dexter was seen eat a frog*, like that of ?*Dexter was promised to take a bath*, is semantic in origin, as the increased acceptability of *Dexter was seen eaten by frogs* and *Dexter was promised to be allowed to take a bath* suggests.

7. Koopman and Szabolcsi (1998), referring to Kenesei (1989), discuss data for Hungarian verb raising which appear to suggest that Hungarian has categories corresponding to both (9) and (12).

8. Again, I depart from the earlier analysis and follow Johnson (1988) in assuming an object control verb category for these verbs. VP_{-SUB} is an abbreviation for $S_{-SUB}^{-CP,-IP}\backslash NP$.

9. This is the first of two differently restricted rules of forward crossed composition—see the appendix to this chapter.

10. Increasingly higher-order composition rules are required to combine maximally incremental assembly of a derivation on the basis of the idealized German categories. (Four verbs requires $<B^3$, and so on.) Since there is a small finite bound on $<B^n$, this fact may suggest a reason why in reality German tends to adopt the Dutch tensed-first word order, and hence the Dutch categories, when more than two verbs are implicated in sentences of this kind, as Bech (1955) and Evers (1975) observed. It may also explain the somewhat greater difficulty of processing the "German" order found by Bach,

Brown and Marslen-Wilson (1986). See Joshi 1990 for discussion and an alternative explanation for the result.

11. This change improves upon the earlier analysis. However, there are Germanic languages in which these variants apparently *are* allowed with the corresponding verbs. (An example is the Zürich dialect of Swiss German, which allows the equivalent of (22–24), as well as the order that is allowed in Dutch. See Lötscher 1978, Shieber 1985, Cooper 1988.

12. In Steedman (1985) I reported some doubt among my informants as to whether it is actually ungrammatical. Seuren (1985) includes an extended discussion of these verbs in relation to the theory in Steedman 1985.

13. It seems possible however that there is a purely phonological solution to this problem.

14. The deeper the extracted NP, the more cumbersome the resulting sentence, presumably for pragmatic reasons. Where the semantics permits, as in the ambiguous (i), some informants will only accept the extraction of the higher complement NP:

(i) de jongen die ik het meisje zag kussen
the boy that I the girl saw kiss
'the boy that I saw kiss the girl/ the boy that I saw the girl kiss'

However, in examples like (35), where semantics will only permit the deeper extraction, they will accept it, so the limitation appears not to lie in the grammar.

15. Under certain circumstances topicalization of objects and the like appears to be allowed in subordinate clauses like (i) (Angeliek van Hout, personal communication):

(i) ... dat zulke lekkere chocola Hendrik altijd in zijn eentje probeert op te eten.
... that such delicious chocolate Hendrik always on his own tries up to eat
'... that Hendrik always tries to eat up such delicious chocolate on his own.'

I assume that examples like these in Dutch are handled by the same mechanism as relativization, discussed below.

16. This analysis is somewhat similar to the adjunction-based analyses proposed in Haegeman 1992, 193 (as opposed to the quite different analysis in Haegeman and van Riemsdijk 1986) and in Rambow 1994a. All of these analyses in turn represent a reversion to the adjunction analysis of Evers (1975).

17. "*" here means "bad under the reading indicated by the subscripts." Such scrambling is allowed under certain conditions in German, a fact that may excuse the use of this sort of stipulative fine-tuning of features to exclude it in Dutch.

18. Of course, the grammar correctly allows two further semantically distinct analyses for the example, which is ambiguous in Dutch. These further readings correspond to the English *the teacher who Jan saw t help Cecilia feed the hippopotamuses* and the semantically anomalous *the teacher who Jan saw Cecilia help the hippopotamuses feed t* (cf. note 14).

19. Maling and Zaenen (1978) note that Icelandic, which is sometimes claimed to be an SVO language, also fails to exhibit the subject/object extraction asymmetry. Elsewhere (Steedman 1996b, 55, example (100)) I show how this fact can be explained in present

terms by assuming (as Mailing and Zaenen do) that Icelandic is a V2 language, and (as I do in section 6.7 for Dutch and German main-clause orders) that V2 order arises from a VSO verb category via leftward-extraction.

20. As before, *S* abbreviates the category $S^{-CP,+IP}$ of an IP and *VP* abbreviates $S^{-CP,-IP}\backslash NP$.

21. The present analysis bears some resemblance to Hepple's (1990) account of German main-clause order, in that inverted order is taken as primary in main clauses, and SVX order is assumed to arise from topicalization, in line with the analysis in den Besten 1983. In other respects the analysis differs. A version of the present analysis first appeared as Steedman 1983 and was discussed by Hoeksema (1985), who proposes an alternative analysis (criticized by Hepple) involving multiple lexical entries for the verb.

22. I am grateful to Susanne Kronenberg, Anneke Neijt, and Angeliek van Hout for advice on this corner of the data.

23. Recall that the results of infinitival verbs (etc.) are underspecified with respect to the value of the feature $\pm SUB$. It follows that unification coerces a single lexical category to either value. Thus, this assumption does not entail any expansion to the lexicon. See note 11 for remarks about some Germanic dialects that appear not to draw this distinction.

24. In view of the fact that explicit case and so-called free word order are strongly correlated, we may further conjecture that certain languages with more elaborate case systems, such as Classical Latin and even other dialects of Germanic, may achieve such freedom by exploiting some of the further opportunities for composing cased categories that have been excluded with category-based restrictions on combinatory rules in the grammar of Dutch, as in (54), (60), (61), and (82).

25. Of course, with *non*configurational languages the whole question of which dependencies are crossed and which uncrossed becomes harder to answer. See note 24.

Chapter 7

1. This chapter is distantly based on some parts of Steedman 1990, which it revises and extends.

2. The adverbial inflection of the verb is required on the first conjunct for clausal coordination, but is not the source of the present anomaly. I am indebted to Nobo Komogata for providing these examples.

3. See Miyagawa 1997 for extended arguments in a different framework that such a base-generative account of scrambling is correct for Japanese, and for semantic arguments for distinguishing extraction from scrambling. See Komagata 1999 for a more extended discussion of CCG and Japanese including OS+OSV coordination, and see Hoffman 1995a,b and Baldridge 1998 for discussion of extended formalisms for unordered categories.

4. See McCloskey 1991. I am grateful to Jim McCloskey for help with these examples.

5. This consequence of the present analysis of the syntactic categories, whatever its

advantages for the theory of coordination, has led Dowty to abandon the idea entirely (David Dowty, personal communication).

6. The following analysis of SVO gapping entirely departs from that in Steedman 1990, in which the right conjunct was assembled by a very restricted form of crossed forward composition.

7. Examples like (28d) are standard in British English, although pseudogapping seems to be less productive in North American dialects (see Levin 1978 for an early discussion).

8. It is an assumption forced by the present theory that strings like *eighty percent* in these examples are complete subjects and that the gaps do *not* include anything corresponding to *of the population*, as the brackets indicate.

9. The analysis presented in this section differs in important respects from that in Steedman 1990, in which a related argument cluster arose from *crossed* composition of the SVO subject and complement.

10. The apparent possibility of examples like the following in dialects which allow pseudogapping (see note 7) suggests that such examples, like certain closely related "comparative subdeletion" examples, are anaphorically mediated, like VP ellipsis:

(i) Warren eats beans, but I doubt whether Dexter does potatoes.

11. To the extent that the NP in question is assumed to be nominative, the composition itself requires an instantiation forbidden by the restriction of type-raising to raising over categories that are permitted by the parametric specification of the lexicon under the Principle of Categorial Type Transparency.

12. The slight awkwardness of such examples is presumably related to that of multiple right node raising, which according to Abbott (1976) lies in the domain of pragmatics rather than syntax (see note 16 in chapter 3).

13. Byron's lines on the syphilitic Lord Elgin and the marbles, which include a gap parallel to (52a), must therefore be explained as having arisen in a poetic idiom in which *What hath wrought the pox?*, and hence OS order, was grammatical:

(i) *Noseless himself, he brings us noseless blocks,*
To show what time hath wrought, and what, the pox.

14. This observation also holds of the earlier analysis in Steedman 1990, based on crossed composition.

15. I will pass over the question of exactly how the anaphoric operator θ'' works, whether by directly applying higher-order unification to the types and Logical Forms, as in Dalrymple, Shieber and Pereira 1991 and Shieber, Pereira and Dalrymple 1996, or by one of the combinatory abstraction algorithms discussed in chapter 8, or by a more direct Alternative Semantics representation, of the kind discussed by Rooth (1985) and in chapter 5.

16. We could do this explicitly in the grammatical rule. However, this would make the grammar look as if it involved context-sensitive rules, which it does not.

17. This part of the account diverges from Steedman 1990.

18. Since chapter 5 treats discontinuous themes as multiple distinct presuppositions

about the alternative set, there is more to say here about how the corresponding category is generated. One possibility is to generate λ-terms corresponding to the individual presuppositions (in this case corresponding to "wants" and "to win") and to compose them, then generate a syntactic type.

19. In support of this view it is striking that examples parallel to (70) but involving verbs that do not have subject-extracting categories or "exceptional-case-marking" categories seem less good:

(i) #Dexter wonders whether Watford won, and Warren, Ipswich.

This may be because the corresponding contextual open questions are much harder to formulate and accommodate.

20. Steedman 1990 attempts to extend a related analysis to certain cases of "subject gapping" that are found in Germanic languages in sentences like the following (see Höhle 1983; Wunderlich 1988; Heycock 1991; Heycock and Kroch 1993, 1994):

(i) Toen kwam Jan binnen en dronk bier met ons.
Then came Jan in and drank beer with us
'Then Jan came in and drank beer with us'

A similar extension seems possible in the present framework, but I have not attempted it here.

21. I am grateful to Jason Baldridge for drawing my attention to this consequence of the present analysis, which does not emerge from the earlier analysis in Steedman 1990.

22. Stripping will of course potentially give rise to an extra derivation for right-node-raised sentences, as it will in any theory of grammar that covers the construction at all.

23. See Cormack 1984 and Jacobson 1991 for other categorially based proposals concerning VP ellipsis, and see Romero 1997 for an Alternative Semantics–based account of sluicing.

Chapter 8

1. Here I diverge from the argument on this point in Steedman 1988.

2. See Dowty 1993. However, see Szabolcsi 1989, Cresswell 1990 and Jacobson 1991, 1992a for treatments of such phenomena in variable-free combinatory terms.

3. If we are thinking of the typed calculi, then **I**, **K**, and **S** are schemata over infinite sets of typed combinators.

4. Note that I have not said how the recursive call to *fact* could work. Indeed such recursive definitions are only possible in the untyped calculi. I will return to this point.

5. For expository simplicity I have used the familiar λ-notation to identify the notion of abstraction, rather than Curry and Feys's (1958) "bracket abstraction" notation, which was introduced in order to make it clear that the algorithm constitutes a *definition* of λ-abstraction, rather than invoking it as a primitive.

6. **C** is in fact the combinator that is implicit in one of the Bach/Dowty family of "wrap" operations, which in the present grammar is confined to the lexicon.

7. Smullyan elegantly defines **C** as **BBT**(**BBT**)(**BBT**).

8. See Morrill and Carpenter 1990 for further discussion of the relation between grammar and the λ_1-calculus.

9. Strong generative power is the capacity to generate the trees that a compositional semantics requires, as opposed to the weak generative capacity to generate the set containing all and only the strings of the language. Grammars that generate the same trees for each string of a given language are said to be strongly equivalent, whereas those that assign different trees to the strings are merely weakly equivalent.

10. Interestingly, the parenthesis-free variant of CCG discussed by Friedman, Dai and Wang (1986) and Friedman and Venketasan (1986) is of greater power (see Weir 1988 for discussion).

11. Note also that the category X' in the LIG schemata will not in general be the same as the category X in the corresponding instantiation of the categorial rule. If X is a function category, then X' is its result and the stack ... is its argument(s). In terms of the earlier \$ convention, X is of the form $X'\$$.

12. The above remarks are confined to the intuitive demonstration that any CCG can be translated into a strongly equivalent LIG. It follows that the combinatory categorial languages (CCLs) are a subset of the linear indexed languages (LILs). It does not follow that any LIG can be realized as a CCG. The weak equivalence of CCG and LIG is proved by Vijay-Shanker and Weir (1990, 1994), who show that the set of tree-adjoining languages (TALs) is a subset of the set of CCLs, and that the set of LILs is a subset of that of TALs, from which it follows that the three language classes are identical.

13. The question of the existence of a class of automata that imposes the correct restriction should not be confused with the question of how to naturally constrain the more powerful automaton to capture performance limitations—say via limitations on stack depth, as in much work following Yngve 1960 and Miller and Chomsky 1963, such as Stabler 1994, Gibson 1996, and Lewis 1993.

14. See earlier chapters on the necessary restrictions on these rules.

15. The temptation to write context-sensitive combinatory rules, as in a momentary aberration I proposed elsewhere (Steedman (1987), 413) should also be resisted.

16. This notation is a variant of those proposed by Lyons (1968) and Huck (1985).

17. This device is of course borrowed from Prolog. As in Prolog, different occurrences of the underline symbol represent *distinct* irrelevant variables.

18. Declarativizing positions like this may seem laborious, but we will later be able to borrow the DCG trick of encoding such positions implicitly in difference-lists, so that we will be able to forget about absolute positions entirely.

19. When the rules are further specified in this way, as they must be for real grammars, it is desirable for reasons of efficiency and internal theoretical consistency to further stipulate that L_z be distinct from R_z. This prevents temporary spurious coordination of categories with different directionality.

20. The rules of functional substitution (9) of chapter 4 are also correctly limited to two general and four specific instances by the theory.

21. See below for discussion of how this property interacts with coordination.

22. The categories required for less rigidly configurational languages than English have been investigated by Kang (1988) for Korean, Hoffman (1995a,b) and Bozsahin (1998) for Turkish, Komagata (1999) for Japanese, Baldridge (1999) for Tagalog, and Trechsel (to appear) for Tzotzil. Komagata shows that the underspecification of raised categories can be transmitted under functional composition by similarly underspecified composition rules to capture scrambling in a strictly verb-final language. Baldridge suggests that freely ordering verb categories of the kind proposed by Hoffman should be regarded as schematizing over finitely many directional categories, and that composition and type-raising should be order-preserving.

23. The above argument translates directly into unification-based frameworks such as PATR or Prolog. A small Prolog program is given in Steedman 1991c which can be used to exemplify and check the argument. The program is based on the simple shift-reduce parser/recognizer discussed together with other more practical algorithms in chapter 9 below, and uses difference-list encoding of string position (see Stirling and Shapiro 1986, 239; Pereira and Shieber 1987; Gerdeman and Hinrichs 1990). This program is not proposed as a practical or efficient CCG parser. Although a few English lexical categories and an English sentence are given by way of illustration in Steedman 1991c, the very general combinatory rules that are included there will require further constraints if they are not to overgenerate with larger fragments. (For example, $>\mathbf{B}$ and $>\mathbf{B}_{\times}$ must be distinguished as outlined above, and the latter must be greatly constrained for English.) One very general constraint, excluding all combinations with or into *NP*, is included in the program, in order to force type-raising and exemplify the way in which further constrained rule-instances may be specified.

The 1991 paper includes a claim that the right conjunct in English gapping under an earlier analysis of the phenomenon than the one presented in chapter 7 has the interesting property of refusing to combine with a "real" SVO verb, because the direction features have the effect of paradoxically requiring that verb to be to the right of the subject and to the left of the object. This claim does not apply to, and is not necessary for, the present analysis.

Chapter 9

1. The division of processing labor between a nondeterministic algorithm and an oracle is not always made explicit, particularly in implementations. However, all processors can be viewed in this way. A more extensive survey of the parsing literature in terms of a similar tripartite architecture can be found in Kay 1980.

2. These remarks should not be taken to imply a stand on either side of the question of whether the appearance of rule learning can or cannot be achieved by probabilistic mechanisms such as parallel distributed processors (Rumelhart, Hinton and McClelland 1986) or neural networks. The fact that many actual rule-like systems such as English past tense morphology include exceptions means that in practice rule-based and probabilistic theories have each come to embody elements of the other (see Prasada and Pinker 1993). To the extent that such hybrid systems embody structure, whether distributed in representation or not, it is likely that these remarks about incremental refinement apply to them as well. In fact, to the extent that connectionists deny the idea

of *any* kind of grammar, let alone a covering grammar, they apply even more strongly.

3. Natural processors might conceivably require grammars to be in some normal form. However, provided that the normal form is a class of grammars of the same automata-theoretic power that the semantics of the language requires (and therefore of the same power as the competence grammar), we would expect that normal form to simply be a characteristic of the grammars we actually observe. In other words, we would view it as a (processing-based) constraint on the form of the competence grammars that actually exist. If on the other hand we are to entertain the possibility that the requirement for a covering grammar might arise from the fact that the mechanisms that have access to the outside world are for some reason of a lesser automata-theoretic power than the competence grammar, then the evolutionary claims become even more farfetched. It seems inevitable that the mapping between analyses under the two grammars must become more complex. The problem of incremental language learning becomes correspondingly more complex. So of course does the problem mentioned above, of evolving the two systems in lockstep. Indeed, we have to ask ourselves how these two systems, which by assumption have completely different automata-theoretic character, could begin to talk to one another in the first place. We have to ask ourselves whether it would not be simpler for evolution to bring the processor more in line with the requirements of competence grammar—after all, it has already come up with such a mechanism once, in the form of the interpreter for the competence semantics.

4. The Strict Competence Hypothesis is somewhat unhelpfully referred to in Steedman 1989 simply as the "strong competence hypothesis." The version defined there and here is stricter than Bresnan and Kaplan's version.

5. The CKY algorithm requires grammars in Chomsky Normal Form, where all productions are either binary productions with two nonterminals as daughters, or unary productions replacing a nonterminal with a terminal. All the combinatory rules except the unary type-raising rules and the coordination rule are binary. I will defer the question of what to do about these exceptions, merely noting that everything else about CCG is in Chomsky Normal Form.

6. Strictly speaking, even ordinary CFPSG potentially gives rise to multiple equivalent derivations. However, in contrast to those of CCG, all of the derivation trees are isomorphic, and the standard algorithms like the bottom-up algorithms discussed below are guaranteed to find exactly one derivation for each reading.

7. The term is misleading in the sense that *any* theory of grammar that actually covers the range of coordination, intonation, and extraction phenomena addressed by CCG must necessarily include the same ambiguities and encounter the same nondeterminism in parsing them. The ambiguity is there in the language itself.

8. Eisner (1996) has proposed a related technique that annotates constituents in the chart in such a way that multiple copies are never built, avoiding the need for the unification-based matching check. The technique is related to similar proposals for normal form parsing by Hepple and Morrill (1989), Hendriks (1993), König (1994), and Hepple (1999) for the Lambek calculi.

9. This particular algorithm is based on the incremental version 12.4.2 of the CKY algorithm in Harrison 1978, 433. See Briscoe 1984, Pulman 1986, Hausser 1986, Wirén

1992, van der Linden 1993, and Milward 1994, 1995 for alternative approaches to incremental parsing with CG and related grammars.

10. This profusion of ambiguity suggests that there is a pressure to keep words short. Such a pressure presumably reflects the obvious need to keep utterances from taking too long in comparison to the accompanying thought processes.

11. The even earlier fragment *The flowers sent* ... probably does not of itself provide enough information to eliminate the spurious main-clause analysis, because of the potential availability of continuations like ... *the patient into raptures*. This observation is in line with Trueswell, Tanenhaus, and Garnsey's (1994) results from eye-tracking experiments on similar sentences in null contexts, which suggest that reanalysis is triggered at the preposition, rather than at the preceding verb.

12. See Steedman 1985. While I will not attempt to explore this claim fully here, it seems to be supported by experimental observations by Inoue and Fodor (1995), Konieczny et al. (1997), and others, who show that interpretable structure is built in advance of encountering the verb in verb-final constructions in Japanese and German.

13. The notation is adapted. Pulman also proposes an operation corresponding to type-raising as a rule of processing, under the name "Invoke."

14. See Shieber and Johnson 1993 for an extensive critique of this proposal.

15. Stabler's note 3 seems to confirm that he does not see this as the central issue.

16. See Haddock 1987, 1989, Dale and Haddock 1991, Stone and Doran 1997; Stone and Webber 1998 and Stone 1998, 1999 for proposals for efficient constraint-satisfaction-based mechanisms for incremental semantic evaluation of this kind.

17. Komagata's English lexicon contains 1174 words with 272 distinct categories. The Japanese lexicon contains around 200 words. The average number of lexical entries per word averaged over the 3517 words (197 sentences) parsed with the English grammar was 3.8. That for the 431 words (22 sentences) parsed using the Japanese grammar was 4.5. Average case parsing times were better than n^3.

Chapter 10

1. I have borrowed this interpretation of the GB notion of projection from Szabolcsi 1992.

2. The only obvious alternative seems to be for the child to compute the respective effects of the candidates on the existing grammar—say, by computing the number of unattested alternatives to the sentences that the child has actually encountered that is predicted by each. On this measure the most restrictive alternative would be correctly chosen.

3. One would of course predict a quite different course of acquisition for subject extraction in (for example) Italian. According to the account proposed in Steedman 1996b, Italian subject relatives arise from VOS verbs for which the child will have encountered nonextracted sentences. There is therefore a strong prediction that embedded subject relatives should be acquired much earlier in Italian.

References

Aarts, Erik. 1995. *Investigations in Logic, Language, and Computation*. Ph.D. thesis, Universiteit Utrecht.

Abbott, Barbara. 1976. "Right Node Raising as a Test for Constituenthood." *Linguistic Inquiry*, 7, 639–642.

Ades, Anthony, and Mark Steedman. 1982. "On the Order of Words." *Linguistics and Philosophy*, 4, 517–558.

Aho, Alfred, and S .C. Johnson. 1974. "LR parsing." *Computing Surveys*, 6, 99–124.

Ajdukiewicz, Kazimierz. 1935. "Die syntaktische Konnexität." In Storrs McCall, ed., *Polish Logic 1920–1939*, 207–231. Oxford: Oxford University Press. Translated from *Studia Philosophica*, 1, 1–27.

Alshawi, Hiyan, and Richard Crouch. 1992. "Monotonic Semantic Representation." In *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics*, 32–38. San Francisco, CA: Morgan Kaufmann.

Altmann, Gerry. 1985. *Reference and the Resolution of Local Syntactic Ambiguity*. Ph.D. thesis, University of Edinburgh.

Altmann, Gerry. 1988. "Ambiguity, Parsing Strategies, and Computational Models." *Language and Cognitive Processes*, 3, 73–98.

Altmann, Gerry, and Mark Steedman. 1988. "Interaction with Context During Human Sentence Processing." *Cognition*, 30, 191–238.

Aoun, Joseph, Norbert Hornstein, David Lightfoot, and Amy Weinberg. 1987. "Two Types of Locality." *Linguistic Inquiry*, 18, 537–578.

Asher, Nicolas, and Tim Fernando. 1997. "Labelling Representations for Effective Disambiguation." In *Proceedings of the 2nd International Workshop on Computational Semantics, Tilburg*. Tilburg: Katholieke Universiteit Brabant.

Bach, Emmon. 1976. "An Extension of Classical Transformational Grammar." In *Problems in Linguistic Metatheory: Proceedings of the 1976 Conference at Michigan State University*, 183–224. Lansing: Michigan State University.