Chapter 5

The Combinatory Manifesto (Draft, November 1, 2000)

Mark Steedman

Informatics, University of Edinburgh

Combinatory Categorial Grammar (CCG, Dowty 1988, Steedman 1987), like other varieties of Categorial Grammar (CG, Bach 1976, 1988, Dowty 1978 Oehrle, Bach and Wheeler 1988, Buszkowski, Marciszewski and van Benthem 1988, Wood 1993), is a form of lexicalized grammar in which the application of syntactic rules is entirely conditioned on the syntactic type, or *category*, of its inputs. No rule is conditioned on structure or derivation.

Categories identify constituents as either *functions*, or *primitive categories* such as NP, PP, S, and so on. Primitive categories may be regarded as further distinguished by features, such as number, case, and the like. In the case of functions, the category identifies the syntactic type of the result and of the argument(s) (which may themselves be either functions or primitive categories), the order in which the arguments must combine, and whether they must occur to the right or the left of the functor. (In languages with free word order it may be convenient to schematize categories over a number of orders of combination and directionalities, but the version of CCG presented here follows Baldridge (1999) in assuming that such schemata cover only a finite number of deterministic categories.) Each syntactic category is associated with a logical form whose semantic type is entirely determined by the syntactic category.

5.1 The Question of Power

Pure or Adjukiewicz/Bar–Hillel CG (Ajdukiewicz 1935, Bar-Hillel 1953) limits syntactic combination to rules of functional *application* of functions to arguments to the right or left, a restriction which limits expressivity to the level of context-free grammar. CCG generalizes the context-free core by introducing further rules for combining categories which because of their strictly type-driven character and their semantic correspondence to the simplest of the combinators identified by Curry and Feys (1958) are called *Combinatory*

Rules, and which are the distinctive characteristic of CCG, giving it its name. These rules are strictly limited to certain directionally specialized instantiations of three basic operations, namely functional *Composition*, *Substitution* and *Type-raising*. Under the standard assumptions of CCG, these operations engender a grammar formalism which is provably of low "mildly context-sensitive" automata-theoretic power, weakly equivalent to Tree Adjoining Grammar (TAG, Joshi, Levy and Takahashi 1975; Joshi 1988), Head Grammar (HG, Pollard 1984), and Linear Indexed Grammar (LIG, Aho 1968—see Gazdar 1988 and Vijay-Shanker and Weir 1990, 1993, 1994).

The use of combinatory rules of functional composition and the like as primitive grammatical operations sets CCG apart from other generalizations of Categorial Grammar that are collectively referred to as Type Logical Grammar (TLG, Moortgat 1988, Morrill 1994, Oehrle 2000, this volume), based on the ideas of Lambek (1958, 1988) and van Benthem (1986), which axiomatize function-argument relations as implication in various non-classical or "substructural" logics (see Restall 2000).

Some of the combinatory rules of CCG emerge as theorems of some varieties of Type-Logical Grammar. However, even in the case of functional composition, to obtain as theorems the full range of rules that are used in CCG without engendering grammatical collapse into a permutation-complete calculus requires the "multi-modal" generalization of the Lambek-van Benthem discussed by Oehrle in this volume. For better or worse, that generalization is expressively extremely powerful—as powerful in fact as a Universal Turing Machine, the most powerful formalism there is (Carpenter 1995).

Expressive power is not intrinsically a bad thing. (If—as seems likely—the semantics is the ontogenetical and phylogenetical precursor and determiner of the syntactic system, then that fact will determine the automata-theoretic power of the syntax, which could in principle be quite high.) However, if the aim of linguistics is to provide an account of the observed universal regularities across syntax of human languages that is explanatory in psychological terms, then it is likely in view of the much more limited expressive power that is known to be adequate for programming languages, as well as some known cross-linguistic universal regularities discussed below, that this much power amounts to overkill in terms of degrees of freedom in the theory with comparison to those in the data. If it is possible to achieve linguistic coverage with low expressive power, then it is undoubtedly an advantage, and doing so has always been a key ingredient in the CCG program (see Ades and Steedman 1982), as in Generalized Phrase Structure Grammar (GPSG, Gazdar 1981), TAG, and HG, an ingredient that continues to distinguish this group of theories from the other formalisms considered in the present volume. We will return to the psychological implications below.

5.2 A Spectre is Haunting Linguistics

The involvement of combinatory rules, and in particular those of composition and type-raising, has the effect of greatly generalizing the notion of surface constituency, to the extent that most (though not all) contiguous substrings of sentences have the status of constituents, complete with a syntactic category and a logical form. This property has fortunate consequences for the linguistic analysis of certain phenomena such as coordination, parentheticalization, and intonation structure, all of which become one with the basic mechanism for handling relative clause formation or "*wh*-movement" (Steedman 2000b, hereafter *SP*), and all of which can be captured within a mildly context sensitive formalism.

Despite these manifest attractions, CCG remains a controversial and somewhat scary theory—almost the "bogeyman" that might have been a better translation than "spectre" in the original version of this section heading. Among the reasons for scepticism and dread have been: a) doubts about the propriety of flying in the face of so much received linguistic opinion about the nature of surface structure; b) confusion about how phenomena such as binding and control could be handled in a system in which derivation fails to preserve c-command; c) the technical difficulty of formulating language-specific grammars in CCG and consequent tendencies to overgeneration in early work; d) doubts about whether the rigidly ordered categories and combinatory rules that have been developed for English and Dutch could be generalized to apply to more freely ordered languages; e) the supposed difficulty of parsing in the face of "spurious ambiguity"—that is, the multiple derivations that the combinatory rules yield for any given logical form. In short, the theory has been criticized as unrealistic on psychological, linguistic, and computational grounds.

The present chapter argues that these criticisms are not well-founded, and that CCG is in fact the most linguistically, computationally, and psychologically plausible theory of grammar available. I will argue that traditional notions of surface structure in fact confound claims about surface derivation and constituency with perfectly correct claims about meaning representation or logical form. In particular, it is nowadays uncontroversial to assumethat c-command is a property of logical form, rather than surface syntactic structure. I will also show (following Komagata (1999), Hockenmaier, Bierner and Baldridge (2000) and others) that fragments capturing significant generalizations across different languages can be constructed within the framework, and (following Hoffman (1995a) and Baldridge (1998)) that there is a generalization to free word-order that leaves expressive power unchanged at the same mildly context-sensitive level as standard CCG. Finally, I will show, following Komagata (1999) and Hockenmaier, Bierner and Baldridge (2000), that the problem of parsing in the face of so-called spurious ambiguity is not only easily solvable with standard parsing methodologies. yielding processors which are not only of polynomial worst-case complexity and practicable average case complexity, as well as compatible with state-of-the-art probabilistic optimization, but also directly compatible under the most restrictive assumptions possible with what is known about human sentence processing, as discussed in the chapter by **?** in the present volume.

5.3 The Current Crisis in Syntactic Theory

The most obvious question to ask about the assumptions on which current linguistic theory rests is raised by the continuing need for volumes like the present one: why are there so many theories of grammar these days? Besides the theories discussed in this volume, there are others in active use, including TAG, which has already been mentioned, and Government-Binding theory itself (GB, a.k.a. Principles and Parameters, Chomsky 1981, 1995).

It is usual in science to react to the existence of multiple theories by devising a crucial experiment that will eliminate all but one of them. However, this tactic does not seem to be applicable to these proliferating syntactic theories. For one thing, in some respects they are all rather similar. Sometimes the similarities are disguised by the level of detail at which the grammar is presented—for example, TAG and CCG can be regarded as precompiling into lexical categories some of the feature-unification that goes on during derivations in LFG and HPSG and other attribute-value grammars. Nevertheless, thanks to Reinhart and Reuland (1991, 1993) and Pollard and Sag (1992), who clarified the descriptive account considerably, all of the theories under discussion including CCG and at least some varieties of GB have essentially the same binding theory, with a lexically defined domain of locality corresponding to the tensed clause, and a command or scope relation defined at some level representing predicate arguments structure, such as logical form. The mechanisms involved, even when couched in terms of transformations like "NP movement," seem to be of rather low expressive power—essentially context-free (CF) and "base generable," to use Brame's (1978) term. This covers many phenomena involving dependencies bounded by the tensed verbal domain, such as raising, control, passivization, reflexivization, etc.

It is only when we consider the unbounded dependencies that cross the bounds of the tensed clause in constructions such as the relative clause, various kinds of "reduced" coordinate structures, and other "stylistic" constructions, that the theories differ in important ways. However, in most cases, the apparatus that is added to the CF core is so powerful and expressive that it is impossible to falsify or to distinguish any of the alternatives on grounds of expressive power. "*Wh*-movement" or equivalent coindexing of traces in GB, "functional uncertainty" or the ability to define dependencies in terms of paths defined as regular expressions in LFG, set-valued "SLASH features" in HPSG, "assumption introduction" and withdrawal in TLG, are all examples of powerful mechanisms of this kind. It is a measure of their expressive power that they have to be attended by seemingly arbitrary constraints on their operation which are strikingly similar to one or another of the constraints that limited classical transformational rules, such as the Coordinate Structure Constraint (Ross (1967), the Fixed Subject Condition (a.k.a. the Empty Category Principle or "**that*-trace" filter, first identified by Bresnan 1972), and the Clause Non-final Incomplete Constituent Constraint (Kuno 1973).

Constraints on rules are not necessarily in themselves a sign of anything wrong with a theory of grammar. They can arise from all kinds of extragrammatical sources, such as the requirements of the semantics, the parser, or the language learner. (Island constraints like the Complex Nounphrase Constraint of Ross 1967 provide an example of a group of constraints that should probably be explained in terms of probabilistically or semantically guided parsing rather than in terms of grammar as such.)

However, when a constraint is observed to hold cross-linguistically, as in the case of certain restrictions discussed below which relate coordinate constructions to primary word-order, that fact calls for *some* kind of explanation. One way to provide that explanation is to show that the constraints stem from limitations in automata-theoretic power of the grammar itself. A theory that is incapable in the first place of expressing grammars for languages that violate the condition provides a very convincing explanation for why they hold. Such a theory of grammar may also bring beneficial complexity and learnability results (although such theoretical results do not necessarily tell us much about the actual difficulty of practical processing and language learning for realistic grammars).

The project of explaining constraints on observed grammars as arising in part from grammar formalisms of low expressive power was as noted earlier the impulse behind Generalized Phrase Structure Grammar. (Gazdar 1981; Gazdar et al. 1985), which tried to capture as much as possible within a strictly context-free formalism. While it was clear from the start that phenomena existed that were unlikely to be capturable in this way, the effects of seeing just how many linguistic generalizations *could* be captured in context-free terms, supporting a fully compositional semantics, was extremely salutory. Most of all, it focused attention on multiple long range dependencies, since these required generalization of the mediating SLASH feature to be either a stack- or set- valued feature. In particular the fact that multiple dependencies in English show a tendency to nest rather than cross, as evidenced by the following minimal pair, suggested that SLASH features should be stacks.

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

The two dependencies in (1a) must nest, rather than intercalate, as they would have to for (1b) to have a meaning to do with playing sonatas on violins (the asterisk here means "not allowed with the intended reading.")

However, the tendency to nest multiple dependencies is by no means universal. It was known that in certain Dutch constructions, multiple dependencies obligatorily intercalate (Huybregts 1976, 1984; Shieber 1985), as in the following example:[1]

(2) ... omdat ik Cecilia Henk de nijlpaarden zag helpen voeren.
... because I Cecilia Henk the hippopotamuses saw help feed

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

GPSG itself does not seem to have been particularly amenable to any restricted kind of generalization, and constraining automata-theoretic power ceased to be a major focus of concern during its evolution into HPSG. However, a number of other formalisms, including HG, TAG, and CCG, continued to explore the possibility of capturing human grammars using low-power formalisms. In particular, Ades and Steedman (1982, 522) suggested that the *same* stack might be implicated both in the push-down automaton (PDA) characteristic of context-free grammar and in mediating multiple unbounded dependencies. Vijay-Shanker and Weir (1990, 1993, 1994) subsequently showed that all three formalisms were weakly equivalent to LIG, and formed a new level in the Chomsky Hierarchy characterised by a generalization of the PDA called an Extended Push Down Automaton (EPDA) characterized by a single stack of stack-valued features. Subsequent explorations with TAG and CCG frameworks suggest that this level may be the lowest at which all syntactic phenomena of natural grammar can be captured.[2]

Among this set of weakly equivalent mildly context-sensitive formalisms, the claim of the present chapter is that CCG has distinct advantages for the analysis of certain phenomena of natural grammar that have resisted explanatory analysis in other frameworks—in particular, coordination and intonation structure. To the extent that these phenomena are subsumed under the same apparatus used for relativization and the bounded constructions, CCG has a claim to be considered as a promising theory of all human grammatical phenomena. To present this argument requires a much more radical rethinking of the nature of grammar than is required for the other theories considered in this volume.

5.3.1 Natural and Unnatural Grammar

In artificial languages like arithmetic, logic, and LISP, grammar is almost trivial. Syntax is mostly CF. The only important exception is variable binding (Floyd 1962), which motivated the introduction of the indexed languages to the Chomsky hierarchy (Aho 1968, 1969 and see Gazdar 1988), a level only "mildly" more expressive than CF. Rules and constituents of syntax and semantics are essentially related one-to-one in these artificial languages. What little ambiguity they permit is strictly local (usually resolved by the adjacent symbol.)

The fact that we would expect natural grammars to be the same is enshrined in "The Constituent Condition on Rules," a principle which has always constituted a primary assumption of any linguistic theory:

(3) The inputs to, and outputs from, all rules of syntax should be constituents.

Since constituents are merely the units to which the grammar assigns meaning, this amounts to saying that rules of grammar should have a compositional semantics.

On this view, it is important to know what it is legitimate to regard as a constituent. Interestingly, traditional linguistic tests for constituency are equivocal on this point. Chomsky (1955/1975, p.210-211) offers the following "criteria" (the scare quotes are his):

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

The implication is that rules designed to capture these phenomena should all conform to the constituent condition.

There has been some success in keeping rules within this prescription with the bounded constructions and the "*wh*-movement" family of unbounded constructions under heading (c). Such constructions (in English at least) generally concern relations between quite uncontroversial constituent types, such as NPs, PPs and verbs. It has been much harder to adhere strictly to the constituent condition in analyses of the remaining three phenomena, which appear to involve relations between fragmentary entities of various kinds that do not obviously correspond to any traditional type of constituent.

For example, at first glance coordination looks very akin to *wh*-movement, as Gazdar (1981) pointed out. Thus, "right node raising" is unbounded (*contra* Ross 1967):

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

It is also subject to similar, though not identical, "island conditions:"

(6) a. *How many mathematicians do you think in ten derive a lemma?
    b. *Three mathematicians in ten derive a lemma and in a hundred prove completeness.

However, the residues of coordination in general are much more diverse than those of *wh*-movement:

(7) a. I want to try to write, and *hope to see produced*, a musical about the life of Stafford Cripps.
    b. Give a dog a bone, and *a policeman a flower*.
    c. I want to try to write a novel, and *you, a screenplay*.

All of these constructions appear (at least at first glance) to violate the constituent condition on rules.

The reaction of the early transformational grammarians, including Chomsky (1957) and Ross (1967), was to regard such conjunctions as as arising from underlying or "deep structural" coordination of sentential constituents of type *S*, followed by processes of "deletion under coordination," delivering surface form. While this expedient restored the constituent condition on rules as far as the inputs to rules of coordination were concerned, and ensured interpretatibility at the underlying level of representation, the outputs of deletion processes remained in violation of the constituent condition. Such rules were also of a disconcertingly arbitrary and unmotivated character. While they generated intense interest in early transformational studies, they proved extremely resistant to explanatory formal analysis, and in fact coordination has received comparatively little attention in transformational studies since the mid-'70s.[3]

Although there is much less work on parentheticals (but see Emonds 1979 and Taglicht 1998), the intrusion of parentheticals seem to isolate similar nonconstituent fragments:

(8) a. "*Are you now or have you ever been*," demanded Harriet, "a member of the Friends of the Ukraine Film Society?"
b. *You could give*, he chortled, a policeman a flower.

They also obey similar restrictions to coordinates:

(9) *"Three mathematicians," expostulated Mexican Pete, "in ten derive a lemma."

Intonational phrases (as defined by for example "continuation rises" here marked informally by right parentheses) can correspond to non-constituents (small caps indicate nuclear stress):

(10) a. (The absent minded PROFESSOR)(was avidly READING)(about the latest BIOGRAPHY)(of Marcel PROUST).
b. (Marcel PROVED) (COMPLETENESS).

Phonological processes like American English flapping and the Rhythm Rule are sensitive to such intonational constituency. (In both (10a) and (b), the latter operates to move lexical stress from its normal position on the second syllable of *Marcel* to the first syllable, under the influence of the pitch accents on *PROUST* and *PROVED*.)

Intonational phrase boundaries obey similar restrictions to coordination, parentheticalization, and other "stylistic" constructions:

(11) *(Three MATHEMATICIANS)(in ten derive a LEMMA).

This observation gave rise to the "Sense Unit Condition" on prosodic phrases of Halliday (1967) and Selkirk (1984), according to which such phrases have to be in some sense semantically interpretable (Selkirk gives a definition in terms of headedness).

In view of the similarities noted above between the restrictions on coordination and those on intonation structure, it would have been possible to regard the latter as stemming from the same grammatical processes (of deletion, movement, or whatever) as the former.[4] However, Selkirk instead proposed incorporating an explicit level of Intonation Structure into a then-standard GB/P&P model of generative grammar, as shown in figure 5.3.1 (adapted from Selkirk 1984—cf. p.205).

In according first-class structural citizenship to the intonational phrase, this

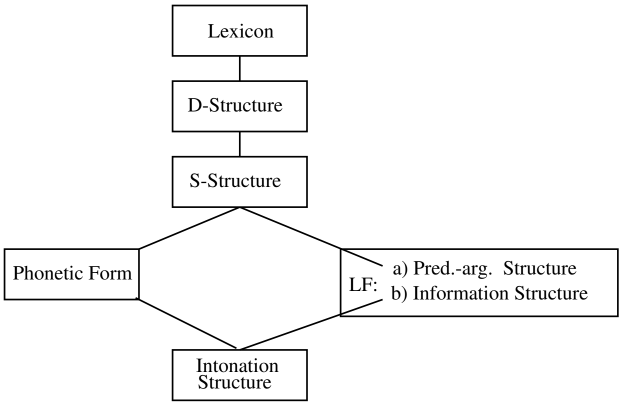


Figure 5.1
Architecture of a GB/P&P Generative Grammar

is an appealing move. However, the resultant theory has some worrying aspects. First, to be forced to stipulate a Sense Unit Condition is a massive failure for any theory of grammar. (Imagine the derision that would greet the proposal to apply this constraint to an otherwise overgenerating theory of traditional syntax.) What is a theory of grammar if not a theory of sense units?

Moreover, the theory now has *two* levels of structure mediating between sound and meaning. Since such structural levels are what we call syntax, the theory has in effect two orthogonal syntaxes. Since phonological form or PF also does significant strictly syntactic work in GB (Chomsky 1986 calls PF "Surface Structure"), dealing with coordination and parentheticalization, it duplicates much of the work of Intonation Structure.

It would be in keeping with other contemporary developments to reduce the number of modules in the theory. In view of the similarity in constraints upon relativization, coordination, parentheticalization, and intonation structure, it seems possible that one could do this by combining PF, S-structure, and Intonation Structure in a single module. To do this we will need to develop a theory that will stand traditional generative syntax on its head.

5.4 All That Is Solid Melts Into Air

A number of researchers have reacted to the above situation by abandoning traditional notions of constituency, allowing "flexible" syntactic structures that provide whatever is needed for the constituent condition on rules to hold. This move in turn requires lexicalizing the grammar. Many of them have turned to the original form of lexicalized grammar, the Categorial Grammars of Ajdukiewicz and Bar-Hillel, and its close relatives, the Dependency Grammars of Tesnière (E.g. Bach 1979; Dowty 1979; Ades and Steedman 1982; Hudson 1984; van Benthem 1986; Oehrle 1987; Szabolcsi 1989; Moortgat 1988; Hepple 1990; Jacobson 1990; Milward 1991; Pickering and Barry 1993; Wood 1993; Morrill 1994, among many others).

Many of these grammars give equal status to both traditional and nontraditional structures as in the following minimal example:

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

This chapter presents one of the alternatives, developed in work at Penn, Edinburgh, and elsewhere with Ades, Szabolcsi, Jowsey, Hepple, Wood, Joshi, Hoffman, Prevost, Park, Komagata, Bierner, Baldridge, Hockenmaier, and others.

5.4.1 Categorial Grammar

In CCG, as in other varieties of Categorial Grammar reviewed by Wood (1993) and exemplified in the bibliography below, syntactic information of the kind that can be captured for English in familiar context-free production rules like (13) is transferred to lexical entries like (14):

(13)
$$\begin{array}{lll} S & \rightarrow & NP \;\; VP \\ VP & \rightarrow & TV \;\; NP \\ TV & \rightarrow & \{proved, finds, \ldots\} \end{array}$$

(14) proved $:= (S\backslash NP)/NP$

This syntactic "category" identifies the transitive verb as a function, and specifies the type and directionality of its arguments and the type of its result. We here use the "result leftmost" notation in which a rightward-combining functor over a domain β into a range α are written α/β, while the corresponding leftward-combining functor is written $\alpha\backslash\beta$.[5] α and β may themselves be function categories.

In order to allow such functors to combine with their arguments, we need Combinatory Rules, of which the two simplest are the following functional application rules:

(15) *The functional application rules*

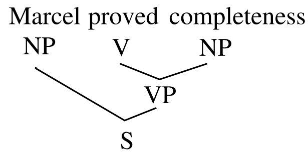
a. $X/Y \quad Y \quad \Rightarrow \quad X$ $\qquad (>)$

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

These rules have the form of very general binary PS rule schemata. In fact, "pure" categorial grammar limited to these these two rules alone is just context-free grammar written in the accepting, rather than the producing, direction, with a consequent transfer of the major burden of specifying particular grammars from the PS rules to the lexicon. While it is now convenient to write derivations as in a, below, they are equivalent to conventional phrase structure derivations b:

(16) a.

$$\begin{array}{ccc}
\text{Marcel} & \text{proved} & \text{completeness} \\
\hline NP & (S\backslash NP)/NP & NP \\
 & \multicolumn{2}{c}{\overline{\qquad S\backslash NP \qquad}{>}} \\
\multicolumn{3}{c}{\overline{\qquad\qquad S \qquad\qquad}{<}}
\end{array}$$

b. 

It is important to note that such tree-structures are simply a representation of the process of derivation. They are not structures that need to be built by a processor, nor do they provide the input to any rules of grammar.

Such categories can be regarded as encoding the semantic type of their translation. This translation can be made explicit in the following expanded notation, which associates a logical form with the entire syntactic category, via the colon operator, which is assumed to have lower precedence than the categorial slash operators. (Agreement features are also included in the syntactic category, represented as subscripts, much as in Bach 1983. The feature $3s$ is "underspecified" for gender and can combine with the more specified *3sm* by a standard unification mechanism that we will pass over here—see Shieber

1986.)[6]

(17) proved := $(S\backslash NP_{3s})/NP : \lambda x.\lambda y.prove'xy$

We must also expand the rules of functional application in the same way:

(18) *Functional application*

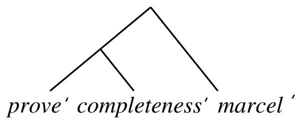
a. $X/Y : f \quad Y : a \quad \Rightarrow \quad X : fa$ $\quad (>)$

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

They yield derivations like the following:

(19)
$$
\begin{array}{ccc}
\text{Marcel} & \text{proved} & \text{completeness} \\
\hline
NP_{3sf} : marcel' & (S\backslash NP_{3s})/NP : \lambda x.\lambda y.prove'xy & NP : completeness' \\
\hline
 & \multicolumn{2}{c}{S\backslash NP_{3s} : \lambda y.prove'completeness'y \quad >} \\
\hline
\multicolumn{3}{c}{S : prove'completeness'marcel' \quad <}
\end{array}
$$

The derivation yields an S with a compositional interpretation, equivalent under a convention of left associativity to (20a):

(20) a. $(prove'completeness')marcel'$ b. 

Thus, the traditional subject-predicate structure reflecting c-command relations exhibited in (20b) is expressed at the level of propositional logical form or f-structure.

5.4.2 Coordination

Coordination is captured in CCG via the following combinatory rule allowing constituents of like type to conjoin to yield a single constituent of the same type:[7]

(21) *Simplified coordination rule* ($<\Phi>$)

$X \quad CONJ \quad X' \Rightarrow \quad X''$

It gives rise to derivations like the following:

(22)
$$\begin{array}{ccccc}
\text{Marcel} & \text{conjectured} & \text{and} & \text{proved} & \text{completeness} \\
\hline NP & (S\backslash NP)/NP & CONJ & (S\backslash NP)/NP & NP \\
 & \multicolumn{3}{c}{\underline{\qquad\qquad\qquad}{<}\Phi{>}} & \\
 & \multicolumn{3}{c}{(S\backslash NP)/NP} & \\
 & \multicolumn{4}{r}{\underline{\qquad\qquad\qquad}{>}} \\
 & \multicolumn{4}{c}{S\backslash NP} \\
\multicolumn{5}{c}{\underline{\qquad\qquad\qquad\qquad}{<}} \\
\multicolumn{5}{c}{S}
\end{array}$$

5.4.3 Composition

In order to allow coordination of contiguous strings that do not constitute constituents, CCG allows certain further operations on functions related to Curry's combinators (Curry and Feys 1958). For example, functions may nondeterministically *compose*, as well as apply, under the following rule:[8]

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

The operation of this rule in derivations is indicated by an underline indexed $>\mathbf{B}$ (because Curry called his composition combinator $\mathbf{B}$). Its effect can be seen in the derivation of sentences like (24), which crucially involves the composition of two verbs to yield a composite of the same category as a transitive verb. It is important to observe that composition also yields an appropriate interpretation for the composite verb *might prove*, as $\lambda x.\lambda y.might'(prove'x)y$, an object which if applied to an object *completeness* and a subject *Marcel* yields the proposition $will'(prove'completeness')marcel'$. The coordination will therefore yield an appropriate semantic interpretation.[9]

(24)
$$\begin{array}{cccccc}
\text{Marcel} & \text{conjectured} & \text{and} & \text{might} & \text{prove} & \text{completeness} \\
\hline NP & (S\backslash NP)/NP & CONJ & (S\backslash NP)/VP & VP/NP & NP \\
: marcel' & : conjecture' & : and' & : might' & : prove' & : completeness' \\
 & & & \multicolumn{2}{c}{\underline{\qquad\qquad\qquad}{>}\mathbf{B}} & \\
 & & & \multicolumn{2}{c}{(S\backslash NP)/NP} & \\
 & & & \multicolumn{2}{c}{: \lambda x.\lambda y.might'(prove'x)y} & \\
 & \multicolumn{4}{c}{\underline{\qquad\qquad\qquad\qquad}{<}\Phi{>}} & \\
 & \multicolumn{4}{c}{(S\backslash NP)/NP} & \\
 & \multicolumn{4}{c}{: \lambda x.\lambda y.and'(might'(prove'x)y)(conjecture'xy)} & \\
 & \multicolumn{5}{r}{\underline{\qquad\qquad\qquad\qquad\qquad}{>}} \\
 & \multicolumn{5}{c}{S\backslash NP} \\
 & \multicolumn{5}{c}{: \lambda y.and'(might'(prove'completeness')y)(conjecture'completeness'y)} \\
\multicolumn{6}{r}{\underline{\qquad\qquad\qquad\qquad\qquad\qquad}{<}} \\
\multicolumn{6}{c}{S : and'(might'(prove'completeness')marcel')(conjecture'completeness'marcel')}
\end{array}$$

5.4.4 Type-Raising

Combinatory grammars also include type-raising rules, which turn arguments into functions over functions-over-such-arguments. These rules allow arguments to compose, and thereby take part in coordinations as in (26).

(25) *Forward type-raising* ($>\mathbf{T}$)
$NP : a \quad \Rightarrow_{\mathbf{T}} \quad \mathrm{T}/(\mathrm{T}\backslash NP) : \lambda f.fa$

T is a metavariable over categories. If instantiated as S, it allows the following derivation:

(26)
$$\begin{array}{cccccc} \text{Marcel} & \text{proved} & \text{and} & \text{I} & \text{disproved} & \text{completeness}\\ \overline{NP} & \overline{(S\backslash NP)/NP} & \overline{CONJ} & \overline{NP} & \overline{(S\backslash NP)/NP} & \overline{NP} \\ \overline{S/(S\backslash NP)}^{>\mathbf{T}} & & & \overline{S/(S\backslash NP)}^{>\mathbf{T}} & & \\ \multicolumn{2}{c}{\overline{S/NP}^{>\mathbf{B}}} & & \multicolumn{2}{c}{\overline{S/NP}^{>\mathbf{B}}} & \\ \multicolumn{5}{c}{\overline{S/NP}^{<\Phi>}} & \\ \multicolumn{6}{c}{\overline{S}^{>}} \end{array}$$

Type-raising is restricted to primitive argument categories, NP, PP etc., and over primitive functors like verbs. It therefore resembles the traditional notion of *case*. Unlike the other combinatory rules, it can be regarded as a lexical or morphological-level process.[10]

5.4.5 Universal Grammar

Only combinatory rules which obey the following principles are allowed:

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

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

(29) *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 is merely the definition of combinators themselves. The other principles say that combinatory rules may not override, but must rather "project," the directionality specified in the lexicon. More concretely: the Principle of Consistency excludes the following kind of rule:

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

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

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

On the other hand, these principles do allow rules such as the following:

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

Such rules are not theorems of type calculi such as that of Lambek (1958) and its descendants, and in fact cause collapse of such calculi into permutation completeness if added as axioms Moortgat 1988). In CCG, individual languages are free to exclude such rules entirely (as English does rule (32)) or to limit them to specific categories (as Dutch does for the same rule in the fragment described in *SP*).[11]

5.4.6 Linguistic Predictions

Since complement-taking verbs like *think*, VP/S, can in turn compose via rule (23) with fragments like *Marcel proved*, S/NP, derived as in (26), we correctly predict the fact that right-node raising is unbounded, as in a, below, and also provide the basis for an analyis of the similarly unbounded character of leftward extraction, as in b (see Steedman 1996 for details, including discussion of ECP effects and other extraction asymmetries):

(33) a. [I conjectured]$_{S/NP}$ and [you think Marcel proved]$_{S/NP}$ completeness.
b. The result [which]$_{(N\backslash N)/(S/NP)}$ [you think Marcel proved]$_{S/NP}$.

This apparatus has also been applied to a wide variety of coordination phenomena, including English "argument-cluster coordination", "backward gapping" and verb-raising constructions in Germanic languages, and English gapping. The first of these is illustrated by the following analysis, from Dowty (1988), in which the ditransitive verb category$(VP/NP)/NP$ is abbreviated as

DTV, and the transitive verb category VP/NP is abbreviated as *TV*:[12]

(34) give a teacher an apple and a policeman a flower

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

Since we have already assumed Forward Type Raising and Composition, this construction is predicted to exist in English by arguments of symmetry.

Given independently motivated limitations on type-raising, examples like the following are still disallowed:[13]

(35) *Three mathematicians [[in ten]$_{PP}$ [derive a lemma,]$_{S\backslash NP}$] and [in a hundred prove completeness.]

5.4.7 "Mild Context Sensitivity"

The availability of *crossed* composition (32) to the grammar of Dutch under the Principles of Consistency and Inheritance allows crossed dependencies:

(36)

$$\begin{array}{cccccc}
\text{dat} & \text{ik} & \text{Cecilia} & \text{de nijlpaarden} & \text{zag} & \text{voeren} \\
 & \overline{NP_1} & \overline{NP_2} & \overline{NP_3} & \overline{((S_{+SUB}\backslash NP_1)\backslash NP_2)/VP_{-SUB}} & \overline{VP\backslash NP_3} \\
 & & & & \multicolumn{2}{c}{\overline{((S_{+SUB}\backslash NP_1)\backslash NP_2)\backslash NP_3}^{>\mathbf{B}_\times}} \\
 & & & \multicolumn{3}{c}{\overline{(S_{+SUB}\backslash NP_1)\backslash NP_2}^{<}} \\
 & & \multicolumn{4}{c}{\overline{S_{+SUB}\backslash NP_1}^{<}} \\
 & \multicolumn{5}{c}{\overline{S_{+SUB}}^{<}}
\end{array}$$

We pass over here the somewhat technical question of how *non*-crossing orders like (37) are excluded, merely noting that they depend on treating categories like *VP* as bundles of features or attribute-value pairs, and that such fine tuning is justified by the fact that other verbs, such as *proberen* (to *try*) *do* permit analogous orders (see *SP*, 145–153).

(37) *dat ik Cecilia zag de nijlpaarden voeren.

The possibility of crossing dependencies immediately implies that CCG may be trans-context-free in terms of generative capacity. CCG is is in fact provably weakly equivalent to Tree-adjoining Grammar (TAG), Head Grammar (HG), and Linear Indexed Grammar (LIG), a group constituting the least expressive natural generalization of CFG that has so far been identified. This equivalence gives rise to a polynomial time worst-case complexity result (Vijay-Shanker and Weir 1990, 1993, 1994).

5.5 Gapping and the Order of Constituents

The phenomenon of "argument cluster coordination" illustrated in (34) is an example of a much broader cross-linguistic generalization due to Ross (1970), concerning a relation between basic word-order parameters such as verb finality and constraints on deletion or gapping under coordination. While Ross originally framed his generalization in terms of a then-orthodox deep-structural word order, and on that basis it has been challenged, when reformulated in terms of surface constituent order, it appears to hold.

Ross (1970) noticed that direction of gapping (leftward or rightward) depends on basic constituent order:

(38) a. SOV: *SOV and SO, SO and SOV
b. VSO: VSO and SO, *SO and VSO
c. SVO: SVO and SO, *SO and SVO

Apparent exceptions such as Zapotec (traditionally VSO) and Dutch (traditionally SOV) which allow both leftward and rightward gapping also have suspiciously mixed word-order, rendering the identification of "basic" word-order moot.

5.5.1 Gapping and SOV Word Order

On the assumption that type raising is order-preserving, and defined over the Japanese SOV lexical type for transitive verbs like (39), the subject and object NP can not only combine with the verb by forward application, but also forward-compose, as in (40)

(39) tazuneta := $(S\backslash NP_{nom})\backslash NP_{acc} : \lambda x.\lambda y.visit'xy$

(40) a. Ken-ga Naomi-o tazuneta.
Ken-NOM Naomi-ACC visit-PAST.CONCL
'Ken visited Naomi.'

b.

$$
\begin{array}{ccc}
\text{Ken-ga} & \text{Naomi-o} & \text{tazuneta} \\
\overline{S/(S\backslash NP_{nom})}^{>\mathbf{T}} & \overline{(S\backslash NP_{nom})/((S\backslash NP_{nom})\backslash NP_{acc})}^{>\mathbf{T}} & \overline{(S\backslash NP_{nom})\backslash NP_{acc}} \\
\multicolumn{2}{c}{\overline{\qquad\qquad\qquad\qquad\qquad\qquad}^{>\mathbf{B}}} & \\
\multicolumn{2}{c}{S/((S\backslash NP_{nom})\backslash NP_{acc})} & \\
\multicolumn{3}{c}{\overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}^{>}} \\
\multicolumn{3}{c}{S}
\end{array}
$$

The resulting nonstandard constituent *Ken-ga Naomi-o* can therefore conjoin, in a mirror image of the English derivation (34):

(41) [Ken-ga Naomi-o], [Erika-ga Sara-o] tazuneta.
$S/((S\backslash NP_{nom})\backslash NP_{acc})$ $S/((S\backslash NP_{nom})\backslash NP_{acc})$ $(S\backslash NP_{nom})\backslash NP_{acc}$
Ken-NOM Naomi-ACC Erika-NOM Sara-ACC visit-PAST.CONCL
'Ken visited Naomi, and Erika, Sara.'

Ditransitives similarly allow larger argument clusters (to save space the intransitive category is abbreviated *VP*, the transitive *TV*, and the ditransitive *DTV*):

(42) a. Kyooju-ga komonjo-o gakusee-ni kasita.
Professor-NOM manuscript-ACC student-DAT lent-PAST.CONCL
'The professor lent the manuscript to the student.'

b.

$$
\begin{array}{cccc}
\text{Kyooju-ga} & \text{komonjo-o} & \text{gakusee-ni} & \text{kasita.} \\
\overline{S/VP}^{>\mathbf{T}} & \overline{VP/TV}^{>\mathbf{T}} & \overline{TV/DTV}^{>\mathbf{T}} & \overline{DTV} \\
\multicolumn{2}{c}{\overline{\qquad\qquad\qquad}^{>\mathbf{B}}} & & \\
\multicolumn{2}{c}{S/TV} & & \\
\multicolumn{3}{c}{\overline{\qquad\qquad\qquad\qquad\qquad}^{>\mathbf{B}}} & \\
\multicolumn{3}{c}{S/DTV)} & \\
\multicolumn{4}{c}{\overline{\qquad\qquad\qquad\qquad}^{>}} \\
\multicolumn{4}{c}{S}
\end{array}
$$

In this case there is another derivation for the argument cluster:

(43)

$$
\begin{array}{cccc}
\text{Kyooju-ga} & \text{komonjo-o} & \text{gakusee-ni} & \text{kasita.} \\
\overline{S/VP}^{>\mathbf{T}} & \overline{VP/TV}^{>\mathbf{T}} & \overline{TV/DTV}^{>\mathbf{T}} & \overline{DTV} \\
 & \multicolumn{2}{c}{\overline{\qquad\qquad\qquad}^{>\mathbf{B}}} & \\
 & \multicolumn{2}{c}{VP/DTV} & \\
\multicolumn{3}{c}{\overline{\qquad\qquad\qquad\qquad\qquad}^{>\mathbf{B}}} & \\
\multicolumn{3}{c}{S/DTV} & \\
 & \multicolumn{3}{c}{\overline{\qquad\qquad\qquad\qquad}^{>}} \\
 & \multicolumn{3}{c}{S}
\end{array}
$$

Again, both derivations are guaranteed to yield identical logical forms, and all non-standard argument cluster constituents formed in both derivations can coordinate.

Ross's generalization that SOV verbs gap on the left conjunct is therefore captured: the Principles of Adjacency, Consistency, and Inheritance, together with the order-preserving constraint on type-raising that is the sine qua non of an order-dependent language, permit any raised categories or rules of composition that would produce a *leftward*-looking function. "Forward Gapping" is therefore disallowed in any language with a pure verb-final lexicon:

(44) *Ken-ga Naomi-o tazunete, Erika-ga Sara-o
Ken-NOM Naomi-ACC visit-PAST.ADV Erika-NOM Sara-ACC
'Ken visited Naomi, and Erika, Sara.'

Dutch which is regarded as an SOV language *does* allow coordinations on the above pattern in subordinate-clause conjunctions, in apparent exception to Ross' generalization:

(45) ...dat Maaike aardappels eet en Piet bonen
...that Maaike potatoes eats and Piet beans
'...that Maaike eats potatoes and Piet beans.'

However, this fact is clearly related to the fact that Dutch has VSO/SVO word order as well as SOV. In CCG terms, this corresponds to the fact that its lexicon is *not* purely SOV, and that main verbs must be assumed to bear VSO categories. Ross' generalization should therefore be rephrased as in *SP*, in terms of surface order, i.e available lexical type(s), not a single underlying order. (In fact, it should be clear by now that in CCG there is no such thing as underlying word-order.)

This observation is relevant to the fact that Japanese also allows OSV word order, as in (46):

(46) Naomi-o Ken-ga tazuneta.
Naomi-ACC Ken-NOM visit-PAST.CONCL
'Ken visited Naomi.'

OS order can also give rise to constituent cluster coordination parallel to (41), as in (47):[14]

(47) [Naomi-o Ken-ga,] [Sara-o Erika-ga] tazuneta
$S/((S\backslash NP_{acc})\backslash NP_{nom})$ $S/((S\backslash NP_{acc})\backslash NP_{nom})$ $(S\backslash NP_{acc})\backslash NP_{nom}$
Naomi-ACC Ken-NOM, Sara-ACC Erika-NOM visit-PAST.CONCL
'Ken visited Naomi and Erika, Sara.'

The natural response is to regard so-called scrambling constituent orders in Japanese as lexically specified, either via multiple verb categories or via explicitly unordered leftward verb categories, in keeping with the observation that scrambling (as distinct from true extraction) is clause-bounded. (See Miyagawa 1997 on base-generation of scrambling in Japanese. See Komagata 1999, Hoffman 1995a,b, and Baldridge 1998 for discussion of extended formalisms for unordered categories. Baldridge argues that free word-order phenomena should be handled by schematizing lexical categories over finite numbers of standard CCG categories like the above Japanese verbs, strictly limiting the corresponding generalized combinatory rules to ones whose results correspond to the set of categories that would result from standard CCG rules applied to (the cross-product of) the relevant sets of standard categories. Such a generalization preserves the original weakly TAG-equivalent level of mildly context-sensitive expressive power.)

Unlike the Hoffman/Komagata generalization, Baldridge's Set-CCG does not allow the following pattern of coordination to be captured:

(48)

?[Naomi-o Ken-ga,]	[Erika-ga Sara-o]	tazuneta
$S/((S\backslash NP_{acc})\backslash NP_{nom})$	$S/((S\backslash NP_{nom})\backslash NP_{acc})$	$(S\backslash NP_{nom})\backslash NP_{acc}$
Naomi-ACC Ken-NOM,	Erika-NOM Sara-ACC	visit-PAST.CONCL

'Ken visited Naomi and Erika, Sara.'

The status of such coordinations in Japanese is not entirely clear. They tend to be judged less acceptable than the earlier alternatives, and seem to be pragmatically marked. (Interestingly, the similarly problematic SO&OSV pattern seems to be even less acceptable than the OS&SOV pattern in (48)). But the consensus seems to be that they are grammatical.

We will return briefly to this problem in section 5.5.3.

5.5.2 Gapping and VSO Word Order

It is obvious on arguments from symmetry that the VSO subject and object must be assumed to raise over VSO verbs like the transitive verb like *chonaic* ("saw"), (49), and to compose with each other and with adjuncts in an order-preserving way to yield a single function, as in (50)

(49) chonaic := $(S/NP)/NP : \lambda x.\lambda y.see'yx$

(50) a. Chonaic Eoghan Siobhán.
saw Eoghan Siobhán
'Eoghan saw Siobhán.'

b.
$$
\begin{array}{ccc}
\text{Chonaic} & \text{Eoghan} & \text{Siobhán} \\
\hline
(S/NP)/NP & \overline{(S/NP)\backslash((S/NP)/NP)}^{<\mathbf{T}} & \overline{S\backslash(S/NP)}^{<\mathbf{T}} \\
 & \multicolumn{2}{c}{\overline{S\backslash((S/NP)/NP)}^{<\mathbf{B}}} \\
\multicolumn{2}{c}{\overline{\qquad\qquad S \qquad\qquad}^{<}} &
\end{array}
$$

Since such clusters necessarily bear leftward functor categories, when they coordinate they give the appearance of rightward gapping, in line with Ross' generalization.

(51) Chonaic [Eoghan Siobhán] agus [Eoghnaí Ciarán].
$(S/NP)/NP$ $S\backslash((S/NP)/NP)$ $S\backslash((S/NP)/NP)$
saw Eoghan Siobhán and Eoghnaí Ciarán
'Eoghan saw Siobhán, and Eoghnaí, Ciarán.'

Again the three principles exclude the "backward gapping" construction that Ross (1970) held to be generally disallowed in strictly verb-initial languages:

(52) *[Eoghan Siobhán] agus chonaic [Eoghnaí Ciarán].
$S\backslash((S/NP)/NP)$ $(S/NP)/NP$ $S\backslash((S/NP)/NP)$
Eoghan Siobhán and saw Eoghnaí Ciarán

As in the case of SOV languages, allegedly VSO languages exist that allow leftward gapping as well as rightward. Zapotec (Rosenbaum 1977) is a standard example. However, like Dutch, Zapotec has mixed word order. It is therefore again consistent with a version of Ross' generalization formulated as in *SP* in terms of surface orders and available lexical verb categories rather than "deep" or underlying word-order.

5.5.3 Gapping and SVO Word Order

The fact that gapping in English and other verb-medial languages is rightward, on the VSO pattern, needs further apparatus. since the NPs in the ungapped conjunct are separated by the SVO verb.

(53) Marcel proved completeness, and Gilbert, soundness.

Comparison with example (34) and the fact that the English lexicon contains a limited class of VSO verbs already suggests that an explanation is not far away.

SP, following Steedman 1990, proposes a class of "decomposition" rules which map the left conjunct onto a virtual VSO verb and a virtual SO argument cluster with the same category as the adjacent gapped right coordinate. The latter can coordinate with the virtual cluster and apply to the verb (whose interpretation has to be obtained contextually, possibly by processes akin to VP anaphora). This augmentation of CCG makes a number of correct predictions about the possibility of "Stripping" constructions in English. Karamanis (2000) and Bozsahin (2000) have used this apparatus to capture the kind of mixed-order gapping illustrated for Japanese in (48) in Greek and Turkish. However, it remains controversial and is passed over here.

5.6 Intonation Structure and Parentheticals

In order to capture coordination with rules adhering to the constituent condition, CCG generalizes surface constituency to give substrings like *Marcel proved* and even *a policeman a flower* the full status of constituents.

But if they are constitutents of coordinate constructions, they must be possible constituents of ordinary non-coordinate sentences as well. This is exactly what we need in order to bring intonation structure and the related phenomenon of parentheticalization into the grammar. (Because of restrictions of space, this part of the account is sketched in less detail than that on coordination. For details the reader is referred to Prevost (1995) and Steedman (2000a).)

5.6.1 Intonation and Information Structure

Consider the following minimal pair of dialogs, in which intonational tunes are indicated both informally via parentheses and small capitals as before, and in the standard notation of Pierrehumbert (1980) and Pierrehumbert and Beckman (1989), in which prosodic phrases are specified solely in terms of two kinds of elements, the pitch accent(s) and the boundary:

(54) Q: I know who proved soundness. But who proved COMPLETENESS?

A: (MARCEL) (proved COMPLETENESS).
H* L L+H* LH%

(55) Q: I know which result Marcel PREDICTED. But which result did Marcel PROVE?

A: (Marcel PROVED) (COMPLETENESS).
L+H*LH% H* LL%

In (54A), there is a prosodic phrase on MARCEL including the sharply rising pitch accent that Pierrehumbert calls H*, immediately followed by an L boundary, perceived as a rapid fall to low pitch. There is another prosodic phrase having the somewhat later-rising and (more importantly) lower-rising pitch accent called L+H* on *COMPLETENESS*, preceded by null tone (and therefore interpolated low pitch) on the word *proved* and immediately followed by an utterance-final rising boundary, written LH%.

In (55A) above, the order of the two tunes is reversed: this time, the tune with pitch accent L+H* and boundary LH% occurs on the word *PROVED* in one prosodic phrase, *Marcel PROVED*, and the other tune with pitch accent H* and boundary LL% is carried by a second prosodic phrase *COMPLETENESS*.

The intuition that these tunes strongly convey systematic distinctions in discourse meaning is inescapable. For example, exchanging the answer tunes between the two contexts in (54) and (55) yields complete incoherence. Prevost and Steedman (1994) claim that the tunes L+H*LH% and H*L (or H* LL%) are respectively associated with the "theme" and "rheme" of the sentence, where these terms are used in the sense of Mathesius (1929), Firbas (1964, 1966), and Bolinger (1989), and correspond roughly to a generalization of the more familiar terms "topic" and "comment", which however are generally restricted by definition to traditional constituents.

Informally the theme can be thought of as corresponding to the content of a contextually available *wh*-question, which may be explicit, as in (54) and (55), or implicit in other discourse content. The position on the pitch accent, if any, in the theme, distinguishes words corresponding to "focused" elements of the content which distinguish this theme from other contextually available alternatives. The rheme can then be thought of as providing the answer to the implicit *wh*-question, with the pitch accent again marking focused words which distinguish this answer semantically from other potential answers. The system comprising the oppositions of theme/rheme and focus/background is known as information structure. Steedman (2000b) provides a more formal definition in terms of the "alternative semantics" of Rooth (1985, 1992), and the related “structured meanings” of Cresswell (1973, 1985) and others.[15]

Since alternatives like the following are equally valid surface derivations in CCG, it will be obvious that CCG provides a framework for bringing intonation structure and its interpretation information structure into the same syntactic system as everything else:

(56)
$$\begin{array}{ccc}
\text{Marcel} & \text{proved} & \text{completeness} \\
\hline
NP : marcel' & (S\backslash NP)/NP : prove' & (S\backslash NP)\backslash((S\backslash NP)/NP) \\
\overline{S/(S\backslash NP) : \lambda f.f\ marcel'}^{\,>\mathbf{T}} & & : \lambda p.p\ completeness' \\
 & \multicolumn{2}{c}{\overline{S\backslash NP : \lambda y.prove'completeness'y}^{\,<}} \\
\multicolumn{3}{c}{\overline{S : prove'completeness'marcel'}^{\,<}}
\end{array}$$

(57)
$$\begin{array}{ccc}
\text{Marcel} & \text{proved} & \text{completeness} \\
\hline
NP : marcel' & (S\backslash NP)/NP : prove' & S\backslash(S/NP) : \lambda p.p\ completeness' \\
\overline{S/(S\backslash NP) : \lambda f.f\ marcel'}^{\,>\mathbf{T}} & & \\
\multicolumn{2}{c}{\overline{S/NP : \lambda x.prove'x\ marcel'}^{\,>\mathbf{B}}} & \\
\multicolumn{3}{c}{\overline{S : prove'completeness'marcel'}^{\,<}}
\end{array}$$

Crucially, these alternative derivations are guaranteed to yield the same predicate argument relations, as exemplified by the logical form that results from (57) and (56), *prove′completeness′marcel′*. However, the derivations do so via different routes that construct lambda terms corresponding semantically to the theme and rheme. In particular the derivation (56) corresponds to the information structure associated with the intonation contour in (54), while derivation (57) corresponds to the that in (55).

We will pass over the details of exactly how this observation is captured in CCG, referring the reader to Prevost (1995) and to Steedman 2000b. The latter paper generalizes this approach to the full range of tunes identified by Pierrehumbert, including those with multiple pitch accents and multiple or disjoint themes and rhemes.

5.6.2 Parentheticals

While we will not discuss parentheticals in any detail here, it seems likely that they too should be defined in terms of information structural units. In most cases, the parenthetical intrusion itself appears at the boundary between theme and rheme, hence its subjection to the same constraints as intonational phrase boundaries:

(58) a. Marcel proved, so he claimed, the crucial property of completeness.
b. *"Three mathematicians," expostulated Mexican Pete, "in ten derive a lemma."

5.7 Implications for Performance: the Strict Competence Hypothesis

Psycholinguistic evidence concerning performance can never prove or disprove a competence theory. However, all things being equal, standard assumptions about the desirability of parsimony in scientific theories would lead us to expect that the performance mechanism would require the minimum of specialized apparatus over and above the competence grammar that supports and is presumably determined by the semantics. To that extent, it is relevant to consider the implications for processing of the present theory of grammar, despite the fact that like any competence theory, it stands or falls on linguistic grounds.

The minimum apparatus besides competence grammar that is required for processing consists of the characteristic automaton for the relevant class of grammars, a minimal algorithm for applying the rules, and some memory for building interpretable structure. Any extra apparatus such as rule-orderings or "strategies," covering grammars, and the like, is otiose. To the extent that such extra stipulations are cross-linguistically universal, they complicate the problem of explaining language evolution. To the extent that they are language-specific, they complicate the problem of explaining child language acquisition.

The most restrictive hypothesis of all is that the processor involves no resources at all beyond the minimum specified above. Such processors are incapable of building intermediate structures other than the constituents defined by the competence grammar, and for this reason the hypothesis that the human processor has this character is called the "strict competetence" hypothesis (SCH).

One very simple processor adhering to this principle is based on the left-to-right version of the Cocke-Kasami-Young (CKY) parser (see Harrison 1978), a bottom-up parser which fills the cells of an $n \times n$ table or half-matrix t representing all spans between positions (i, j) in a string of n words.

The associativity of functional composition in interaction with type-raising potentially creates exponentially many multiple derivations for any given constituent for a given span with a given sense or interpretation (the so called "spurious ambiguity" problem). It follows that such a parser will have exponential computational costs *unless* we include a check that a newly-derived category spanning (i, j) *including its normalized logical forms* is not already on the list in $t(i, j)$ before appending it, a suggestion first made by Karttunen (1989).

Such parsers have been shown by Komagata (1999) and Hockenmaier, Bierner and Baldridge (2000) to be of better than cubic observed time complexity in the length of the sentence for reasonable-sized hand-built grammars, and to be practicable as a basis for large-scale semiautomatically learned grammars induced from corpora.

Cubic time costs are still prohibitive for really large volume parsing and unrealistic as a model of the human parser, which appears to be linear time or better. For large volume parsing of text corpora, statistical optimization techniques integrating probabilistic head-dependency with competence-based grammar of the kind proposed by Collins (1998) and Charniak, Goldwater and Johnson (1998) are particularly well-adapted to CCG, and are under investigation by Hockenmaier, Bierner and Baldridge (2000).

For modeling human parsing, there is every indication that something even more restrictive is needed. Bever's (1970) observation that naive subjects typically fail to find any grammatical analysis at all for "garden path" sentences like (59a) shows that the human processor is "incomplete":

(59) a. The doctor sent for the patient arrived.
     b. The flowers sent for the patient arrived.

The fact that (as Bever also noticed) the same subjects typically judge the isomorphic sentence (59b) as grammatical suggests that the human sentence processor prunes the search space on the basis either of the relative likelihood of nounphrases like *the doctor* or *the flower* being dependent in relations like subject or object on verbs like *send for*, or the relative likelihood of the various logical form corresponding to entire prefixes such as *the flowers/doctor sent for* in a particular context. In the case of (59a) this will cause the only analysis compatible with the rest of the sentence to be rejected, causing the garden path. Crain and Steedman (1985) and Altmann and Steedman (1988) showed that manipulating the context for related sentences in such a way as to pragmatically support the modifier reading eliminates the classic garden path effect. This fact suggests that the latter alternative is at work, rather than (or perhaps as well as) the former purely statistical mechanisms.

They proposed a modification of the basic parser according to which as each word was processed in a left-to-right traverse of the sentence, rival analyses developed in parallel could immediately be pruned according to the following principle:

(60) *The Principle of A Priori Plausibility*
The analysis whose interpretation is most plausible in the current context will be preferred.

Since plausibility could depend on reference to context, this was expanded as including the following special case.

(61) *The Principle of Parsimony*
The analysis whose interpretation carries fewest unsatisfied but accommodatable presuppositions will be preferred.

(Since restrictive reduced relative NP interpretations for phrases like *the doctor sent for the patient* carry presuppositions such as that there is more than one doctor in the discourse contest, in neutral contexts they will be dispreferred, leading to garden paths like the one in (59).)

In terms of traditional grammar, both probabilistic and semantic interpretations of the plausibility effect on garden paths present a problem for the strict competence hypothesis (SCH). If the parser is to take account of the incompatibility of *flowers* and the subject slot of the tensed verb reading of *sent for*, this information must become available before *the patient* is integrated. (Otherwise the processor would be able to "see" the incompatible verb "arrived," and avoid the garden path in (59a).)

This means that the parser must implicitly or explicitly have access to the interpretation or partial structure corresponding to the prefix *The flowers sent for* .... But this substring is not a legal constituent according to standard grammars. So SCH appears to be breached: the parser has built or thought about building a relation that the grammar does not recognize via constituency.

This may not seem to be a very serious problem in English, where the subject and verb are immediately adjacent and could be related by other means, albeit in violation of SCH. However in verb final languages characterized by constructions like the Dutch example (2), in which arbitrarily many arguments can be separated from their verbs by long-distance dependencies similar effects would be much more problematic, in effect requiring the parser to to have sophisticated predictive mechanisms and to build explicit or implicit partial structures corresponding to non constituent fragments.

Dutch, German and Japanese speakers greet with derision the suggestion that their languages prohibit any analysis until the verb group (in the Dutch bare infinitival construction, the *entire* verb group) has been processed. Moreover, there are a number of experimental results which are claimed to show such effects. In particular, Gorrell (1995b,a); Inoue and Fodor (1995); Mazuko and Itoh (1995); Sturt and Crocker (1996); Kamide and Mitchell (1999) show that Japanese speakers are committed to one analysis of an ambiguity arising from the possibility of null anaphora in complex argument sequences, as revealed by garden path effects when a verb incompatible with the preferred analysis is encountered. Konieczny et al. (1997) claim to show a similar early committment for German. All authors relate these effects to availability of *case* information in these languages, a phenomenon whose resemblance to type-raising has already been noted.

For example, Kamide and Mitchell (1999) note that in ambiguous sentences like the following, in which the dative object of either *kasita* ("lent") or *miseta* ("showed") could be a null anaphor, there is a bias (once the intrinsic preferences of the verbs are controlled for) in favor of the former reading—that is, for the explicit NP meaning the student to be attached to the higher verb (HA) rather than the lower (LA):

(62) Kyooju-ga gakusee-ni toshokansisho-ga kasita mezurasii komonjo-o miseta
[Professor-NOM {student-DAT] librarian-NOM} lent$_\uparrow$ unusual manuscript-ACC showed$_\uparrow$
'the professor showed [HA: the student] the unusual manuscript which the librarian had lent [LA: the student].'

Introducing disambiguating transitives at $\uparrow$ shows early commitment to [ ] over { } causing increased processing load as measured by selfpaced reading time at the disambiguation point *tore* in (64):

(63) Kyooju-ga gakusee-ni toshokansisho-ga yabutta mezurasii komonjo-o miseta
[Professor-NOM student-DAT] librarian-NOM tore$_\uparrow$ unusual manuscript-ACC showed
'the professor showed the student the unusual manuscript which the librarian had torn.'

(64) Kyooju-ga gakusee-ni toshokansisho-ga kasita mezurasii komonjo-o yabutta
Professor-NOM {student-DAT librarian-NOM} lent unusual manuscript-ACC tore$_{\uparrow}$
'the professor tore the unusual manuscript which the librarian had lent the student.'

Kamide and Mitchell claim that these results show that the processor is committed to [ ] over { } *before* encountering the subordinate clause verb, since the verbs alone are on average unbiased. Under standard grammatical assumptions this would imply that the parser is in violation of SCH (since neither [Professor-NOM student-DAT] nor {student-DAT librarian-NOM} are constituents under these assumptions). However, both are available as constituents complete with interpretations under the various alternative derivations that CCG allows (cf. (42)), and can therefore be compared without violating SCH. (Such a comparison is likely to favor the former if only on probabilistic grounds, since the order DAT-NOM is non-canonical in Japanese.)

It would be pleasant to be able to take this result at face value, and claim it in support of a model of human sentence processing using CCG as the competence grammar under the strict competence hypothesis. We have already seen that both [Professor-NOM student-DAT] and {student-DAT librarian-NOM} are available as argument cluster constituents complete with logical forms under alternative CCG derivations illustrated in (42) and (43). CCG therefore provides everything that is needed for the parser to compare the analyses either in probabilistic or semantic/pragmatic terms under the principle of a priori plausibility and its variants.

However, such a claim would be premature at this point. Kamide and Mitchell's interpretation of their result is based on a somewhat restrictive view of the possible parsing architectures. Under more liberal assumptions, a processor using a standard right branching grammar could produce these effects without making any committment until the subordinate verb is actually encountered. If the processor can at that point without violating SCH temporarily build interpretable structure correponding to both null dative anaphor and lexicalised dative argument and compare the corresponding propositions—"librarian lent *pro trace*" vs. "librarian lent student *trace*'—choosing the most likely, then we cannot predict the preference from the off-line comparison of "librarian lent *pro* manuscript" vs. "librarian lent student manuscript," which Kamide and Mitchell show to be balanced overall.

The reason we cannot predict it is that (63), unlike the offline comparison of simplex clauses in isolation, provides a really salient antecedent for the pronominal null anaphor *pro* at the level of logical form, namely the professor.[16] Moreover, (63) does not offer *any* explicit possible antecedent for the matrix clause null anaphor. The preference for the interpretation of the relative clause under these conditions might well swing towards "librarian lent *pro trace*"—or in CCG terms the logical form $\lambda x.lend'(pro'professor')x\ librarian'$.

The only way to decide whether there really is the even earlier committment hypothesised by Kamide and Mitchell is to do an experiment in the style of Crain and Altmann, controlling the contextual appropriateness of the [Professor-NOM student-DAT] and {student-DAT librarian-NOM} cluster constituents via contextual narratives or the Japanese equivalent of questions such as *What manuscript did the professor show the student?* vs. *What manuscript did the professor tear?*.

If such an early committment by the processor is observed, it will be of interest that CCG allows such processors to adhere rigorously to the strict competence hypothesis, even for verb-final languages.

5.8 The Workers Have Nothing to Lose But Their Chains

Because of its very literal-minded adherence to the constituent condition on rules, and the consequent introduction of composition and type-raising, which project directionality specifications and other information from the lexicon subject to principles of directional inheritance and consistency, Combinatory Categorial Grammar abandons traditional notions of surface constituency in favor of "flexible" surface structure, in which most contiguous substrings of a grammatical sentence are potential constituents, complete with a compositional semantic interpretation, for the purposes of the application of grammatical rules.

The benefits of this move are the following.

1. Coordination, Parentheticalization, and Intonation Structure can all be handled with the same apparatus that is required for "*wh*-movement" constructions such as relativization, using purely type-driven syntactic rules that strictly adhere to the Constituent Condition on Rules.

2. Efficient processing including weakly semantically interactive incremental parsing remains possible and is compatible with rigorous ob-

servation of the strict competence hypothesis, even for head-final languages.

3. Everything that depends on relations of "c-command" (e.g. binding and control, quantifier scope) must be dealt with at the level of logical form (cf. Bach 1980; Lasnik and Saito 1984), with a consequent transfer of responsibility for the grammar of bounded constructions to the lexicon.

4. The modules of Phonological Form, S-Structure, and Intonational Structure are unified into a single surface derivational module.

The architecture of the theory that ensues can be summarized in the version of the standard T- or Y- diagram shown in figure 5.8, in which the markers θ and ρ denote features marking elementary categories which permit the combinatory rules to project theme- and rheme- marking from accented words onto phrasal constituents. (The markers θ' and ρ' in the logical forms denote the corresponding semantic functions at the level of information structure, while asterisks in logical forms mark focussed or contrastive semantic elements deriving from words bearing pitch-accents.

While information structure further reduces to a fairly tradition (albeit unordered) predicate-argument structure by β-normalization (essentially, functional application of the λ terms), such a reduction is unnecessary both for the theory of grammar and for the purposes of comprehension and production. It follows that information structure is the only level of logical form that is required and the only true representational level in CCG: since all rules are purely type-driven and monotonic, and no rule depends upon derivational history, surface structures need not be built as explicit representations. CCG is therefore a strictly monostratal theory that fully integrates coordination and intonation structure into surface grammar.

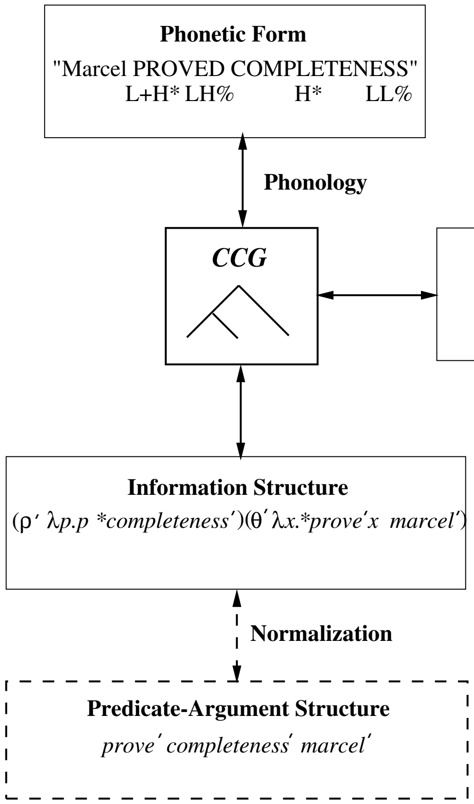


Figure 5.2
Architecture of CCG

Notes

Chapter 5

1. 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.

2. This conjecture has been challenged by Rambow (1994) and defended by Joshi, Rambow and Becker (to appear).

3. Notable exceptions are Goodall (1983, 1987) and Moltmann (1992).

4. I believe that this was actually proposed by Mark Liberman at one stage.

5. There is an alternative "result on top" notation due to Lambek Lambek (1958), according to which the latter category is written $\beta\backslash\alpha$.

6. This notation follows Steedman (1987). Another notation, used in Steedman Steedman (1990), associates a unifiable logical form with each primitive category, so that the same transitive verb appears as follows:

(i) proved := $(S : \mathit{prove}'xy\backslash NP_{3s} : y)/NP : x$

The advantage is that the predicate-argument structure is built directly by the unification, and that the combination rules need no further modification. Otherwise the choice is largely a matter of notational convenience.

7. The semantics of this rule, or rather rule schema, is somewhat complex, and is omitted here. The rule is also simplified syntactically in several respects for the present purpose.

8. Combinatory rules like functional composition resemble a highly restricted (because type-driven rather than structure-dependent) class of "generalized" or "double-based" transformations of the kind proposed in Chomsky 1957.

9. The analysis begs some syntactic and semantic questions about the coordination rule. See Steedman (1990, 1996) for more complete accounts.

10. To the extent that both type-raising and case associate entities with roles in actions, they are both akin to the psychologists notion of the "affordance" of a percept, as that word is used by Gibson (1966) and his followers.

11. The fact that CCG allows category-specific restrictions on combinatory rules is a point of difference from the Lambek and Type-logical tradition discussed in the chapter by by Oehrle in this volume, and means that the CCG slash cannot in any simple sense be axiomatized as implication in any logic. The difference should not, however, be overstated. There is an obvious duality between restricting rules as to the categories that they may apply to, and restricting the categories themselves by distinguishing different modal varieties of slash associated with different axiom sets generating different sets of rules as theorems.

12. In more recent work, Dowty has disowned CCG in favour of TLG, because this analysis apparently entails an "intrinsic" use of logical form to account for binding phenomena. This issue is discussed in Steedman (1996)

13. This appears to offer an advantage over non-type-raising accounts using the product operator • of Lambek (Pickering and Barry 1993; Dowty 1996).

14. This fact precludes any attempt to account for (46) in terms of forward crossed composition, as in the following derivation:

(i)
$$
\begin{array}{ccc}
\text{Naomi-o} & \text{Ken-ga} & \text{tazuneta} \\
\overline{(S\backslash NP_{nom})/((S\backslash NP_{nom})\backslash NP_{acc})}^{>\mathbf{T}} & \overline{S/(S\backslash NP_{nom})}^{>\mathbf{T}} & \overline{(S\backslash NP_{nom})\backslash NP_{acc}} \\
 & \multicolumn{2}{c}{\overline{\qquad\qquad\qquad\qquad\qquad\qquad}^{*>\mathbf{B}_{\times}}} \\
 & \multicolumn{2}{c}{S\backslash NP_{acc}} \\
\multicolumn{3}{c}{\overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}^{>}} \\
\multicolumn{3}{c}{S}
\end{array}
$$

15. The much-abused term "focus" is used in CCG strictly in the "narrow" or phono-

logical sense of the term, to refer to the effects of contrast or emphasis on a word that ensues from the presence of a pitch-accent.

16. Another is the student, but the accessibility hierarchy and extrapolation from English suggests that the matrix subject will be the default antecedent for the null anaphor.

References

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

Aho, Alfred. 1968. "Indexed Grammars—An Extension of Context-Free Grammars." *Communications of the Association for Computing Machinery*, 15, 647–671.

Aho, Alfred. 1969. "Nested Stack Automata." *Communications of the Association for Computing Machinery*, 16, 383–406.

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.

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

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.

Bach, Emmon. 1979. "Control in Montague Grammar." *Linguistic Inquiry*, 10, 513–531.

Bach, Emmon. 1980. "In Defense of Passive." *Linguistics and Philosophy*, 3, 297–341.

Bach, Emmon. 1983. "Generalized Categorial Grammars and the English Auxiliary." In Frank Heny and Barry Richards, eds., *Linguistic Categories: Auxiliaries and Related Puzzles, II*, 101–120. Dordrecht: Reidel.

Bach, Emmon. 1988. "Categorial Grammars as Theories of Language." In Richard T. Oehrle, Emmon Bach, and Deirdre Wheeler, eds., *Categorial Grammars and Natural Language Structures*, 17–34. Dordrecht: Reidel.

Baldridge, Jason. 1998. *Local Scrambling and Syntactic Asymmetries in Tagalog*. Master's thesis, University of Pennsylvania.

Baldridge, Jason. 1999. "A Categorial Account of Extraction Asymmetries in Tagalog and Toba Batak." Ms., University of Edinburgh.

Bar-Hillel, Yehoshua. 1953. "A Quasi-Arithmetical Notation for Syntactic Description." *Language*, 29, 47–58.

van Benthem, Johan. 1986. *Essays in Logical Semantics*. Dordrecht: Reidel.

Bever, Thomas. 1970. "The Cognitive Basis for Linguistic Structures." In John Hayes, ed., *Cognition and the Development of Language*, 279–362. New York: Wiley.

Bolinger, Dwight. 1989. *Intonation and Its Uses*. Stanford, CA: Stanford University Press.

Bozsahin, Cem. 2000. "Ordered Set-CCG." Ms. Middle East Technical University, Ankara.

Brame, Michael. 1978. *Base Generated Syntax*. Seattle, WA: Noit Amrofer.

Bresnan, Joan. 1972. *Theory of Complementation in English Syntax*. Ph.D. thesis, MIT.

Buszkowski, Wojciech, Witold Marciszewski, and Johan van Benthem. 1988. *Categorial Grammar*. Amsterdam: John Benjamins.

Carpenter, Bob. 1995. "The Turing-Completeness of Multimodal Categorial Grammars." Ms., `http://macduff.andrew.cmu.edu/carpenter/research.html`.

Charniak, Eugene, Sharon Goldwater, and Mark Johnson. 1998. "Edge-Based Best-First Chart Parsing." In *Proceedings of the 6th Workshop on Very Large Corpora, Montreal, August*, 127–133.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 1975. *The Logical Structure of Linguistic Theory*. Chicago: University of Chicago Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of Language*. New York: Praeger.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Collins, Michael. 1998. *Head-Driven Statistical Models for Natural Language Parsing*. Ph.D. thesis, University of Pennsylvania.

Crain, Stephen, and Mark Steedman. 1985. "On Not Being Led up the Garden Path: the Use of Context by the Psychological Parser." In Lauri Kartunnen David Dowty and Arnold Zwicky, eds., *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*, 320–358. Cambridge: Cambridge University Press.

Cresswell, M.J. 1973. *Logics and Languages*. London: Methuen.

Cresswell, M.J. 1985. *Structured Meanings*. Cambridge, MA: MIT Press.

Curry, Haskell B., and Robert Feys. 1958. *Combinatory Logic: Vol. I*. Amsterdam: North Holland.

Dowty, David. 1978. "Governed Transformations as Lexical Rules in a Montague Grammar." *Linguistic Inquiry*, 9, 393–426.

Dowty, David. 1979. "Dative Movement and Thomason's Extensions of Montague Grammar." In Steven Davis and Marianne Mithun, eds., *Linguistics, Philosophy, and Montague Grammar*, 153–222. Austin, TX: University of Texas Press.

Dowty, David. 1988. “Type-Raising, Functional Composition, and Nonconstituent Coordination.” In Richard T. Oehrle, Emmon Bach, and Deirdre Wheeler, eds., *Categorial Grammars and Natural Language Structures*, 153–198. Dordrecht: Reidel.

Dowty, David. 1996. “Nonconstituent Coordination, Wrapping, and Multimodal Categorial Grammars: Syntactic Form as Logical Form.” In Kees Doets, ed., *Proceedings of the Tenth International Congress of Logic, Methodology, and Philosophy of Science, 1995*. Amsterdam: North Holland.

Emonds, Joseph. 1979. “Appositive Relatives Have No Properties.” *Linguistic Inquiry*, 10, 211–243.

Firbas, Jan. 1964. “On Defining the Theme in Functional Sentence Analysis.” *Travaux Linguistiques de Prague*, 1, 267–280.

Firbas, Jan. 1966. “Non-thematic Subjects in Contemporary English.” *Travaux Linguistiques de Prague*, 2, 229–236.

Floyd, Robert. 1962. “On the Nonexistence of a Phrase-Structure Grammar for ALGOL 60.” *Communications of the Association for Computing Machinery*, 5, 483–484.

Gazdar, Gerald. 1981. “Unbounded Dependencies and Coordinate Structure.” *Linguistic Inquiry*, 12, 155–184.

Gazdar, Gerald. 1988. “Applicability of Indexed Grammars to Natural Languages.” In Uwe Reyle and Christian Rohrer, eds., *Natural Language Parsing and Linguistic Theories*, 69–94. Dordrecht: Reidel.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwell.

Gibson, James. 1966. *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton-Mifflin Co.

Goodall, Grant. 1983. “A Three-dimensional Analysis of Coordination.” In *Proceedings of the 19th meeting of the Chicago Linguistic Society*. Chicago: University of Chicago, Chicago Linguistic Society.

Goodall, Grant. 1987. *Parallel Structures in Syntax*. Cambridge: Cambridge University Press.

Gorrell, Paul. 1995a. “Japanese Trees and the Garden Path.” In Reiko Mazuka and Noriko Nagai, eds., *Japanese Sentence Processing*, 331–350. Hillsdale, NJ: Erlbaum.

Gorrell, Paul. 1995b. *Syntax and Parsing*. Cambridge: Cambridge University Press.

Halliday, Michael. 1967. *Intonation and Grammar in British English*. The Hague: Mouton.

Harrison, Michael. 1978. *Introduction to Formal Language Theory*. Reading MA: Addison-Wesley.

Hepple, Mark. 1990. *The Grammar and Processing of Order and Dependency: A Categorial Approach*. Ph.D. thesis, University of Edinburgh.

Hockenmaier, Julia, Gann Bierner, and Jason Baldridge. 2000. "Providing Robustness for a CCG System." In *Proceedings of the ESSLLI workshop on Linguistic Theory and Gr ammar Implementation*. `http://grok.sourceforge.net/`.

Hoffman, Beryl. 1995a. *Computational Analysis of the Syntax and Interpretation of "Free" Word-Order in Turkish*. Ph.D. thesis, University of Pennsylvania. IRCS Report 95-17.

Hoffman, Beryl. 1995b. "Integrating Free Word Order, Syntax, and Information Structure." In *Proceedings of the 7th Conference of the European Chapter of the Association for Computational Linguistics, Dublin*, 245–252. San Francisco, CA: Morgan Kaufmann.

Hudson, Richard. 1984. *Word Grammar*. Oxford: Blackwell.

Huybregts, Riny. 1976. "Overlapping Dependencies in Dutch." *Utrecht Working Papers in Linguistics*, 1, 24–65.

Huybregts, Riny. 1984. "The Weak Inadequacy of Context-free Phrase-structure Grammars." In Ger de Haan, Mieke Trommelen, and Wim Zonneveld, eds., *Van Periferie naar Kern*. Dordrecht: Foris.

Inoue, A., and Janet Dean Fodor. 1995. "Information-paced parsing of Japanese." In R. Masuka and N. Nagai, eds., *Japanese Sentence Processing*. Hillsdale, NJ: Erlbaum.

Jacobson, Pauline. 1990. "Raising as Function Composition." *Linguistics and Philosophy*, 13, 423–476.

Joshi, Aravind. 1988. "Tree Adjoining Grammars." In David Dowty, Lauri Karttunen, and Arnold Zwicky, eds., *Natural Language Parsing*, 206–250. Cambridge: Cambridge University Press.

Joshi, Aravind, Leon Levy, and M. Takahashi. 1975. "Tree-adjunct Grammars." *Journal of Computer Systems Science*, 10, 136–163.

Joshi, Aravind, Owen Rambow, and Tilman Becker. to appear. "Complexity of Scrambling—A New Twist to the Competence-Performance Distinction." In Anne Abeille and Owen Rambow, eds., *Proceedings of the 3rd TAG+ workshop, Jussieu, 1994*. Stanford, CA: CSLI Publications.

Kamide, Yuki, and Don Mitchell. 1999. "Incremental Pre-head Attachment in Japanese Parsing." *Language and Cognitive Processes*, 11, 631–662.

Karamanis, Nikiforos. 2000. "Ordered Set-CCG." Ms. University of Edinburgh.

Karttunen, Lauri. 1989. "Radical Lexicalism." In Mark Baltin and Anthony Kroch, eds., *Alternative Conceptions of Phrase Structure*. Chicago: University of Chicago Press.

Komagata, Nobo. 1999. *Information Structure in Texts: A Computational Analysis of Contextual Appropriateness in English and Japanese*. Ph.D. thesis, University of Pennsylvania.

Konieczny, Lars, Barbara Hemforth, Christoph Scheepers, and Gerhard Strube. 1997. "The Role of Lexical Heads in Parsing: Evidence from German." *Language and Cognitive Processes*, 12, 307–348.

Kuno, Susumu. 1973. "Constraints on Internal Clauses and Sentential Subjects." *Linguistic Inquiry*, 4, 363–386.

Lambek, Joachim. 1958. "The Mathematics of Sentence Structure." *American Mathematical Monthly*, 65, 154–170.

Lambek, Joachim. 1988. "Categorial and Categorical Grammars." In Richard T. Oehrle, Emmon Bach, and Deirdre Wheeler, eds., *Categorial Grammars and Natural Language Structures*, 297–317. Dordrecht: Reidel.

Lasnik, Howard, and Mamoru Saito. 1984. "On the Nature of Proper Government." *Linguistic Inquiry*, 15, 235–289.

Mathesius, Vilem. 1929. "Functional Linguistics." In Josef Vachek, ed., *Praguiana: Some Basic and Less Well-known Aspects of the Prague Linguistics School*, 121–142. Amsterdam: John Benjamins. Publ. 1983.

Mazuko, Reika, and Kenji Itoh. 1995. "Japanese Trees and the Garden Path." In Reiko Mazuka and Noriko Nagai, eds., *Can Japanese Speakers be Led Down the Garden Path*, 295–329. Hillsdale, NJ: Erlbaum.

Milward, David. 1991. *Axiomatic Grammar, Non-Constituent Coordination, and Incremental Interpretation*. Ph.D. thesis, University of Cambridge.

Miyagawa, Shigeru. 1997. "Against Optional Scrambling." *Linguistic Inquiry*, 28, 1–25.

Moltmann, Friederike. 1992. *Coordination and Comparatives*. Ph.D. thesis, MIT.

Moortgat, Michael. 1988. *Categorial Investigations*. Ph.D. thesis, Universiteit van Amsterdam. Published by Foris, Dordrecht, 1989.

Morrill, Glyn. 1994. *Type-Logical Grammar*. Dordrecht: Kluwer.

Oehrle, Richard T. 1987. "Boolean Properties in the Analysis of Gapping." In Geoffrey Huck and Almerindo Ojeda, eds., *Syntax and Semantics, Vol. 20: Discontinuous Constituency*, 201–240. Orlando, FL: Academic Press.

Oehrle, Richard T. 2000. "Multi-modal Type Logical Grammar." In Robert Borsley and Kersti Borjars, eds., *Non-Transformational Syntax: A Guide To Current Models*, in preparation. Oxford: Blackwell.

Oehrle, Richard T., Emmon Bach, and Deirdre Wheeler, eds. 1988. *Categorial Grammars and Natural Language Structures*. Dordrecht: Reidel.

Pickering, Martin, and Guy Barry. 1993. "Dependency Categorial Grammar and Coordination." *Linguistics*, 31, 855–902.

Pierrehumbert, Janet. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. thesis, MIT. Distributed by Indiana University Linguistics Club, Bloomington.

Pierrehumbert, Janet, and Mary Beckman. 1989. *Japanese Tone Structure*. Cambridge, MA: MIT Press.

Pollard, Carl. 1984. *Generalized Phrase Structure Grammars, Head Grammars, and Natural Languages*. Ph.D. thesis, Stanford University.

Pollard, Carl, and Ivan Sag. 1992. "Anaphors in English and the Scope of Binding Theory." *Linguistic Inquiry*, 23, 261–303.

Prevost, Scott. 1995. *A Semantics of Contrast and Information Structure for Specifying Intonation in Spoken Language Generation*. Ph.D. thesis, University of Pennsylvania.

Prevost, Scott, and Mark Steedman. 1994. "Specifying Intonation from Context for Speech Synthesis." *Speech Communication*, 15, 139–153.

Rambow, Owen. 1994. *Formal and Computational Aspects of Natural Language Syntax*. Ph.D. thesis, University of Pennsylvania.

Reinhart, Tanya, and Eric Reuland. 1991. "Anaphors and Logophors: An Argument Structure Perspective." In Jan Koster and Eric Reuland, eds., *Long Distance Anaphora*, 283–322. Cambridge: Cambridge University Press.

Reinhart, Tanya, and Eric Reuland. 1993. "Reflexivity." *Linguistic Inquiry*, 24, 657–720.

Restall, Greg. 2000. *An Introduction to Substructural Logics*. London: Routledge.

Rooth, Mats. 1985. *Association with Focus*. Ph.D. thesis, University of Massachusetts, Amherst.

Rooth, Mats. 1992. "A Theory of Focus Interpretation." *Natural Language Semantics*, 1, 75–116.

Rosenbaum, Harvey. 1977. "Zapotec Gapping as Counterevidence to Some Universal Proposals." *Linguistic Inquiry*, 8, 379–395.

Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Ph.D. thesis, MIT. Published as *Infinite Syntax!*, Ablex, Norton, NJ, 1986.

Ross, John Robert. 1970. "Gapping and the Order of Constituents." In Manfred Bierwisch and Karl Heidolph, eds., *Progress in Linguistics*, 249–259. The Hague: Mouton.

Selkirk, Elisabeth. 1984. *Phonology and Syntax*. Cambridge, MA: MIT Press.

Shieber, Stuart. 1985. "Evidence against the Context-Freeness of Natural Language." *Linguistics and Philosophy*, 8, 333–343.

Shieber, Stuart. 1986. *An Introduction to Unification-Based Approaches to Grammar*. Stanford, CA: CSLI Publications.

Steedman, Mark. 1987. "Combinatory Grammars and Parasitic Gaps." *Natural Language and Linguistic Theory*, 5, 403–439.

Steedman, Mark. 1990. "Gapping as Constituent Coordination." *Linguistics and Philosophy*, 13, 207–263.

Steedman, Mark. 1996. *Surface Structure and Interpretation*. Cambridge, MA: MIT Press.

Steedman, Mark. 2000a. "Information Structure and the Syntax-Phonology Interface." *Linguistic Inquiry*, 34. (to appear).

Steedman, Mark. 2000b. *The Syntactic Process*. Cambridge, MA: MIT Press.

Sturt, Patrick, and Matthew Crocker. 1996. “Monotonic Syntactic Processing: a Cross-Linguistic Study.” *Language and Cognitive Processes*, 11, 449–494.

Szabolcsi, Anna. 1989. “Bound Variables in Syntax: Are There Any?” In Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, eds., *Semantics and Contextual Expression*, 295–318. Dordrecht: Foris.

Taglicht, Josef. 1998. “Constraints on Intonational Phrasing in English.” *Journal of Linguistics*, 34, 181–211.

Vijay-Shanker, K., and David Weir. 1990. “Polynomial Time Parsing of Combinatory Categorial Grammars.” In *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics, Pittsburgh*, 1–8. San Francisco, CA: Morgan Kaufmann.

Vijay-Shanker, K., and David Weir. 1993. “Parsing Some Constrained Grammar Formalisms.” *Computational Linguistics*, 19, 591–636.

Vijay-Shanker, K., and David Weir. 1994. “The Equivalence of Four Extensions of Context-Free Grammar.” *Mathematical Systems Theory*, 27, 511–546.

Wood, Mary McGee. 1993. *Categorial Grammar*. London: Routledge.