Surface Structure

MS-CIS-92-51
LINC LAB 229

Mark Steedman

University of Pennsylvania
School of Engineering and Applied Science
Computer and Information Science Department
Philadelphia, PA 19104-6389

June 1992

Surface Structure*

Mark Steedman
University of Pennsylvania

Second Draft, July 27, 1993

Combinatory Categorial Grammar (CCG) was originally advanced as a theory relating coordination and relativisation.[1] The claim was that these constructions can be analysed at the level of surface grammar, without rules of movement, deletion, passing of slash-features, or the syntactic empty category *Wh*-trace. Instead, CCG generalises the notion of grammatical constituency to cover everything that can coordinate or result from extraction, via the use of a small number of operations which apply to adjacent lexically realised grammatical categories interpreted as functions. These operations over functions are related to certain primitive "combinators", such as func-

*Thanks to Elisabet Engdahl, Janet Fodor, Bob Frank, Mark Hepple, Michael Hegarty, Jim Higginbotham, Polly Jacobson, Nobo Komogata, Dick Oehrle, Jong Park, Matthew Stone, and Anna Szabolcsi, for comments and patient advice. An early version of this paper circulated as Technical Report MS-CIS-92-51, Dept. of CIS, University of Pennsylvania, and was presented in lectures to the Third European Summer School in Language, Logic, and Information, Saarbrucken, August 1991. The research was supported in part by NSF grant nos. IRI90-18513, IRI91-17110, and CISE IIP, CDA 88-22719, DARPA grant no. N00014-90-J-1863, and ARO grant no. DAAL03-89-C0031.

[1]See Ades & Steedman 1982; Dowty 1988; Jacobson 1990; Steedman 1985, 1987, 1990, 1991a; Szabolcsi 1989. CCG is related to but distinct from a number of other generalisations of the early categorial systems of Ajdukiewicz, Bar-Hillel, Lambek, Geach, Lewis, and Montague, including work by Cresswell 1973; Bach 1976, 1979, 1980; Shaumyan 1977; Keenan and Faltz 1978; von Stechow 1979; Dowty 1982; Flynn 1983; Zwarts 1986; van Bethem 1986; Uszkoreit 1986; Wittenburg 1986; Zeevat et al. 1987; Hoeksema 1989, Moortgat 1988b; Oehrle 1988; Wood 1988; Bouma 1987, Keenan 1988, Morrill 1988; Carpenter 1989; Karttunen 1989, Hepple 1990; Steele 1990; Barry 1991, and Reape 1991, among others.

tional composition. They appear to have attractive properties for a number of linguistic problem domains. Some of these results will be discussed in more detail below.

The inclusion of associative operations such as functional composition engenders very unorthodox derivational structures, which do not preserve traditional notions of dominance and command. It might appear therefore that the existence of phenomena like binding and control, which appear to depend on these relations, poses a threat to the theory. Part I of the present paper shows that the basic phenomena of binding and control can be captured straightforwardly in CCG, in much the same way that such phenomena are captured in other lexicalist grammatical frameworks that derive from the Montague tradition, including Montague Grammar itself (MG, Bach and Partee 1980), Head-driven Phrase Structure Grammar (HPSG, Pollard and Sag 1991, 1992, other versions of Categorial Grammar (Dowty 1982, Szabolcsi 1989, Jacobson 1987 and Hepple 1990), and more distant theoretical cousins like Lexical-functional Grammar (LFG, Bresnan 1982) and some versions of Government-binding theory (GB, Reinhart and Reuland 1991, Chomsky 1992). The aim is merely to formulate a binding theory for CCG that is adequate to sustain the argument in Part II, rather than to solve all problems in Binding Theory. Part II examines the interaction of this version of binding theory with the account of long range dependencies including "parasitic gaps" that is the distinctive contribution of CCG. The theory correctly explains a number of constraints on such constructions. These include a number of asymmetries with respect to extraction between subjects and other arguments, including "strong crossover" and certain phenomena that have been attributed to the Empty Category Principle, together with the equivalent of an "anti-*c*-command" restriction on parasitic gaps (cf. Taraldsen 1979). This part of the paper includes a number of refinements to the theory of parasitic gaps put forward in Steedman 1987.

PART I

§1 Combinatory Grammar

The material in this section constitutes a brief review of the theory presented in earlier papers, and an extension to incorporate binding and control. This extension is heavily indebted to work by other authors working in related frameworks, although it differs in several respects.[2]

¶1.1 The Categorial Base

Pure Categorial Grammar: CCG is a generalisation of the Categorial Grammars of Ajdukiewicz 1935 and Bar-Hillel 1953. Categorial grammars put into the lexicon most of the information that is standardly captured in context-free phrase-structure rules. For example, instead of using rules like 1 to capture the basic syntactic facts concerning English transitive sentences, we associate with English transitive verbs a category which we will usually write as in 2:

(1)
$$\begin{array}{lll} S & \rightarrow & NP \quad VP \\ VP & \rightarrow & TV \quad NP \\ TV & \rightarrow & \{eats, drinks, \ldots\} \end{array}$$

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

The category says that *eats* is a function that combines with an NP to its right to yield a predicate, which is itself a function bearing the category $S\backslash NP$, and which in turn combines with an NP to its left to yield an S. These two combinations take place via the following pair of rules of functional application, which in a pure categorial grammar are the *only* rules of combination:

(3) Functional Application:
$$\begin{array}{llllll} a. & X/Y & Y & \Rightarrow & X & (>) \\ b. & Y & X\backslash Y & \Rightarrow & X & (<) \end{array}$$

These rules have the form of very general binary PS rule schemata. Clearly what we have here is a context free grammar which happens to be written in the accepting, rather than the producing, direction, and in which there has been a 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 follows, they are clearly equivalent to the familiar trees (except that they are the the right way up):

[2] See the references in footnote 1, and citations in the text.

```
(4) a.  Keats     eats    apples.     b.  Keats  eats  apples
        ----- --------- ------             |     |      |
         NP   (S\NP)/NP   NP               NP    V      NP
              -------------->               \    \     /
                  S\NP                       \     VP
         -------------<                       \   /
              S                                 S
```

FUNCTION-ARGUMENT RELATIONS: While for many purposes this notation is quite sufficient, symbols like S, NP and $S\backslash NP$ can, and in fact must, be regarded as complex objects which include both major syntactic features, of the kind used in X-bar theory, and minor syntactic features like number, gender and person agreement. They also include semantic interpretations. The latter can be thought of either as purely model-theoretic objects, or as structures. For present purposes, it will be helpful to think of them as structures, writing the same category as follows:

(5) eats :- $(S : eat'\ np_2\ np_1 \backslash NP_{3s} : np_1)/NP : np_2$

In this notation, the elements that were written as S, NP, etc. are now written as complex terms. Such terms include a syntactic type, which can conveniently be abbreviated as S, NP_{3s}, etc, since the X-bar theory is not at issue here, and the precise implementation of minor features or feature bundles like agreement is of no particular importance. They also include an interpretation, associated with the syntactic type by a colon. Constants in the interpretation are distinguished from variables by primes. It is important to note that the interpretations obey a convention under which the application of a function (like eat') to an argument (like the variable np_2) is represented by concatenation (as in $eat'\ np_2$), where such expressions "associate to the left". The interpretation of the S result in the category above is thus equivalent to the following expression with the brackets suppressed by this convention:

(6) $(eat'\ np_2)\ np_1$

What is here called the "interpretation" is a level at which function-argument relations, are defined, together with the (interpretation of) the traditional VP and the relation of "command" between the (interpretations of) the arguments such as subject and object. This level of representation is assumed to be unordered. Interpretations therefore define binary "mobiles", capturing dependency but not linear order. The latter property is defined by the directional slashes in the syntactic category. Such categories can therefore be very directly compared with the lexical items in GB theories such as Zubizarreta's (1987), or the lexical entries in HPSG (cf. Pollard and Sag, 1987, or with the elementary trees of a "synchronous lexicalised" Tree Adjunction Grammar (TAG, cf. Joshi and Schabes 1992).

A word of caution is in order with respect to these representations. It might appear that the use of variables such as np_2 in expressions like 6 is equivalent to the involvement of empty categories. However the use of such variables is entirely "non-essential" – that is, we could do the same work in the lexicon without the use of variables, using combinator-based techniques of the kind that are used in syntax below. (See the references to Szabolcsi, Jacobson, and Dowty for examples of such an approach to the lexicon). The present use of structures including variables is merely an expository convenience.

Functions like 6 and arguments such as $NP : apples'$ can be regarded as *terms* or expressions in a logical language. Their combination can then be implemented via the device of *term unification*. For a full exposition of the concept of unification, the reader is directed to Shieber (1986).[3] Informally, unification can be regarded as merging or amalgamating terms that are "compatible", and as failing to amalgamate incompatible ones, via an algorithm that "instantiates" variables by substituting expressions for them in one or other of the expressions.[4]

[3]Unification-based grammar formalisms of the kind assumed here are also discussed in Pereira & Shieber 1987. The present categorial notation is discussed more fully in Steedman 1990 and 1991b.

[4]The result of unifying two compatible terms is the most general term that is an instance of both the original terms. For example, the following pairs of terms unify, to yield the results shown:

$$\text{(i)}\quad \begin{array}{llcl} x & a' & \Longrightarrow & a' \\ f'(g'a') & x & \Longrightarrow & f'(g'a') \\ f'x & f'(g'y) & \Longrightarrow & f'(g'z) \\ f'a'x & f'yy & \Longrightarrow & f'a'a' \end{array}$$

The following pairs of terms do not unify:

Under the unification interpretation, the identical functional application rules 3, in which the variables X and Y now range over the combined syntactic/semantic categories, now give rise to derivations written as follows:

(7)
```
      Keats                      eats                     apples
  -----------  --------------------------------  ----------
  NP3s:keats'  (S:eat' np2 np1\NP3s:np1)/NP:np2  NP:apples'
               ------------------------------------------->
                      S:eat' apples' np1\NP3s:np1
  ----------------------------------------------<
               S:eat' apples' keats'
```

In the first step of this derivation, the forward application rule 3 combines the verb and object via the unification of the term X/Y with the term $(S: eat'\ np_2\ np_1 \backslash NP_{3s} : np_1)/NP : np_2$, so that the subterm X is instantiated as $S : eat'\ np_2\ np_1 \backslash NP_{3s} : np_1$ and the subterm Y is instantiated as $NP : np2$. (The slashes and colons in categories like 6 can be regarded as function constants having the syntax of infix operators for the purposes of unification). The term Y is further instantiated by also being unified with the object $NP : apples'$, which instantiates the variable np_2 in the function and, crucially, its result X. (Thus unification simulates functional application or β-reduction.) The final step of the derivation illustrates the way in which minor features

---

(ii) $a' \quad b' \quad \Longrightarrow \quad fail$
$f'x \quad g'y \quad \Longrightarrow \quad fail$
$f'a'b' \quad f'yy \quad \Longrightarrow \quad fail$

It should be noticed that the unification of two variables is a "new" variable, distinct from either. It should also be noticed that nodes in trees, terms, and variables are under this interpretation pointers to data structures, and unification makes two pointers point to the identical data structure. Strictly, therefore, interpretation structures are directed acyclic graphs, rather than the trees that the present notation suggests. It is emphasised that the use of unification in the present theory is solely as a transparent implementation based on graph reduction for combinatory operations like functional application. Unification does absolutely no autonomous work in the theory of long-range dependencies. While the notation makes use of variables at the level of interpretation, the equivalence of variable-free combinatory systems to the λ-calculus implies that the use of variables for this purpose is not essential. We could accomplish the same effect (much less readably) with a variable-free combinatory semantics of the kind that has sometimes been advocated within CCG. (Cf. Szabolcsi 1989, Jacobson 1991, and the discussion in Dowty 1992. The proof of (weak) equivalence of variable-free combinatory systems to the λ-calculus is in Curry and Feys 1958, Ch. 6, who attribute the result to Rosser.)

capture number agreement. It is a convenience of the unification assumption that we can regard verbs like *ate* that do not happen to be inflectionally distinguished for number as bearing an "underspecified" value that will unify with any number. We shall exploit this notational convenience below.

The example shows that relations of dominance and command are represented in the interpretation of S. Their representation there is quite independent of the derivation, although in this case the structures are isomorphic. It follows that we are free to divorce syntactic dominance from dominance at the level of interpretation. In particular, we are free to assume, as many others have before, that the dominance relations at the level of interpretation (as distinct from those at the level of the derivation) reflect the "obliqueness hierarchy" upon grammatical relations. That is, the first argument of the intepretation of the verb is the most oblique, while the last argument, the subject, is the least, in a sense that has been variously captured in the thematic hierarchy of Jackendoff 1972, the relational hierarchy of Perlmutter and Postal (1977), the accessibility hierarchy of Keenan & Comrie (1977), argument-order in Montague Grammar (MG, Bach 1979, 1980, Dowty 1982, 1992, and Jacobson 1987, 1990, 1991), the control hierarchy in Lexical-Functional Grammar (LFG, Bresnan 1982), the SUBCAT order in Head-driven Phrase-structure Grammar (HPSG, Pollard and Sag 1987, 1991, 1992), and prominence order in Grimshaw 1990.

Such an assumption implies categories like the following for "dative alternation" verbs like *show* in VPs like *show the dog the rabbit* and *show the rabbit to the dog*:[5]

(8) a. showed $:= ((S : show1'\ x\ y\ z \backslash NP : z)/NP : x)/NP : y$
b. showed $:= ((S : show2'\ x\ y\ z \backslash NP : z)/PP : x)/NP : y$

[5]Although *show1* and *show2* are shown as distinct semantic functors, this is not to deny that they are related, say by a lexical rule. There are of course strong universal constraints on the ways categories may map surface structure arguments onto interpretations obeying the obliqueness hierarchy. There seem to be two forces at work constraining categories in this respect. For reasons spelled out below in the discussion of the binding theory, it appears that the obliqueness hierarchy is universally observed at the level of interpretation. However, at least for configurational languages, there seems to be a very strong tendency for the category to impose a linear order from least oblique to most oblique on string positions. This means that arguments that follow the verb tend to be reversed in order of obliqueness, and may explain the rarity of OS languages of all kinds.

Although these categories might appear to exploit variables to achieve the effect of empty categories, the use is again nonessential, and the same effect could be achieved (less transparently) by the use of a "wrap" operator (cf. Bach 1979, 1980) in the interpretation.

BINDING: As Bach and Partee 1980, Chierchia 1988, and others have pointed out, assuming a level of representation at which the obliqueness hierarchy is explicit may have advantages for specifying the theory of binding. For example, the following asymmetry in binding possibilities for reflexives and reciprocals is most naturally explained in terms of the claim that the binder must be lower on the obliqueness hierarchy than the anaphoric bindee:

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

(10) a. I showed the dogs to themselves/each other
b. I showed ?themselves/*each other to the dogs

One simple way to do so is as follows. First recall that the convention of left-associativity means that expressions like $show1'\ x\ y\ z$ are equivalent to $((show1'\ x)\ y)\ z$. Such expressions can therefore be mapped onto (binary, unordered,) trees in the obvious way. Since there are no unary branching nodes in such trees, a relation of *c*-command can be defined on these structures as follows:

(11) *C*-COMMAND: a term α in an interpretation *c*-commands another term β if the node immediately dominating α dominates β.

– where "dominates" is the transitive closure of "immediately dominates."

In defining *c*-command over structures that obey the obliqueness hierarchy, we are again following the lead of Bach and Partee 1980, Dowty 1982, 1992, Jacobson 1987, 1990, 1991, Hepple 1990 and others, who call a related notion "*F*-command", Pollard & Sag 1987, who call it "*o*-command", and Grimshaw 1990, who calls it "*a*-command".

We shall assume that the basic facts that must be accounted for by a theory of binding are as follows. First, reflexives and reciprocals must be bound to a less oblique argument of the same verb interpretation, unless they are "exempt", a term which applies to the reflexives in sentences like the following (cf. Jackendoff 1972; Higgins 1973; Kuno 1987):

(12) a. Chapman enjoyed the jokes about himself
b. A full-size portrait of himself playing the bongos is Chapman's most valued possession.
c. Chapman said that it might have been himself that broke the vase.
d. Chapman suspected that *The National Inquirer* would soon reveal those embarrassing pictures of himself at the Rickmansworth Young Conservatives Ball.

Kuno, 1987, Reinhart & Reuland 1991 and Pollard & Sag 1992 have suggested that binding of such apparently "long-range" anaphors is "logophoric", or discourse-pragmatically mediated. These authors point out that such exempt reflexives can often be replaced by pronouns like *him*, unlike true bound reflexives. However, they argue from minimal pairs like the following that the exempt reflexives must refer to the *experiencer* of the events under discussion:

(13) a. The pictures of himself/him in *Newsweek* embarrassed Chapman.
b. The pictures of *himself/him in *Newsweek* embarrassed Chapman's mother.

The second data-point is that pronouns must *not* be bound to other arguments of the same verb. They may be discourse-bound, by a process which is assumed to be entirely extrinsic to sentence grammar. We will keep an open mind as to whether the binding of pronouns by quantifiers is distinct from discourse-binding or not. If it is distinct then there are many reasons for believing that, like other aspects of quantification, it belongs outside the domain of syntax and argument structure. Sentences like the following suggest that quantifier binding is not limited by *c*-command, nor subject to the constraints like the ECP that limit extraction, and finally that binders and pronouns can "intercalate" their dependencies in a way that is orthogonal to even the liberal notion of surface constituency afforded by CCG:[6]

(14) a. Every man who owns a donkey_i feeds it_i
b. Every woman_i in this room says that she_i is a genius
c. Every man_i believes that some woman_j thinks that he_i loves her_j

[6]This may be too pessimistic. See Szabolcsi 1989 for a proposal to handle pronoun binding in a related Combinatory framework.

Third, full lexical NPs may *only* be bound by discourse. This means that to the extent that coreference in the following examples is possible at all, it cannot arise from syntactic binding:[7]

(15) a. #Keats$_i$ thinks that Keats$_i$ is a genius.
b. #He$_i$ thinks that Keats$_i$ is a genius.
c. *He$_i$ thinks that every man$_i$ is a genius.

Once syntactic binding is partitioned in this way, the conditions on binding of pronouns and reflexives can be captured in terms of the notion of *c*-command at the level of interpretation, a fact which is of course most naturally interpreted in terms of the notion "scope of a variable", with anaphors and pronouns respectively behaving somewhat like "statically bound" and "dynamically bound" variables in modern dialects of Lisp (cf. Abelson & Sussman 1985, pp.321-325).

To capture this idea, we shall first assume (inadequately) that pronouns and non-exempt reflexives bear categories exemplified as follows:

(16) Pronominal Categories (Preliminary Version):
a. *him* := $NP_{3sm} : pro\ x$
b. *himself* := $NP_{3sm} : PRO\ x$

(We shall see later how the non-nominative grammatical case of these items is handled, and how the exempt reflexives can be captured.) It will be convenient to refer to terms like $pro\ x$ and $PRO\ x$ as "pro-terms".[8]

We can now define binding in the following terms, which capture a crucial asymmetry between the conditions under which pro-terms can act as binders or bindees:

(17) A term in an interpretation is bound when it is identical to another term which *c*-commands it, or when it is identical to the argument of a pro-term which *c*-commands it.

[7]It has often been pointed out that in contexts such as the following, examples a and b are good. "Everyone thinks that Keats is a genius. Mary thinks Keats is a genius. Sally thinks Keats is a genius. Even *he/Keats thinks Keats is a genius.*" Of course, by definition, no such context can help example c.

[8]To be consistent with the rest of the notation, the function constants *pro* and *PRO* should be distinguished here with primes. For historical reasons, they are omitted.

"Identical" here refers to literal identity, in the sense of being *the same node, location, or address*, not mere structural isomorphism. Since unification of two nodes in an argument structure makes them point to the identical location, this statement can be viewed as defining binding in terms of unification. It is also weaker than the usual definition of binding in terms of coindexing or coreference. However, the property of being bound is stated as a condition on argument structures in general, not merely as a part of the binding theory.

We shall have more to say later about how pro-terms *get* bound, and in particular about the concept of a "local" binding domain. For the moment we will just give a preliminary definition of the binding theory informally as the following familiar triple of conditions, which in the present context apply to interpretations or argument structures. (It will be convenient to this purpose to refer to binding the argument in a pro-term as binding that pro-term, for short):

(18) BINDING THEORY (PRELIMINARY):

A: An anaphor-type pro-term $PRO\ x$ must be locally bound, if a local binder exists.
B: A pronoun-type pro-term $pro\ x$ must not be locally bound.
C: Nothing but (the argument of) a pro-term may be bound.

This is essentially the binding theory implicit in Pollard & Sag 1992, p.300, and Reinhart and Reuland 1991, modulo some definition of "local binding".

A few further comments are in order. First, the conditional appendix to Condition A permits the existence of "exempt" anaphors. The remainder of condition A captures the basic behaviour of reflexives and reciprocals, including the asymmmetries in 9 and 10. (We shall assume for the moment that at least some "pied-piping" reflexives, such as the one in *Sammy sent the parcel to himself*, are simply $NP : PRO\ x$, and that they can "see" the "immediately dominating" verb and its less oblique arguments, deferring discussion for the moment of how this can be, and how agreement in particular is to be handled.)

Condition B is the complementary condition to condition A. It is therefore narrower than the standard version, ruling out a, but not b or c, in the following example, parallel to 9:

(19) a. *I showed Every man_i/Keats_i him_i
b. *I showed him_i Keats_i
c. *I showed him_i every man_i

Examples b and c are ruled out instead by Condition C applying at the level of interpretation (at which obliqueness is observed), since the full NP would otherwise be bound at that level. Presumably, example b is permitted by discourse binding in the same kinds of loaded contexts as 15,b. But c is completely out, for the same reason as 15,c.

Condition C is therefore also weaker than the standard version. However, it has the important consequence that nothing except the variable in a pro-term may be bound by (that is, unified with) a *c*-commanding term of any kind. It therefore has the effect of imposing upon variables a version of the bijection principle that is characteristic of the the θ-criterion in GB. Nevertheless, it is weaker than the standard θ-criterion, since condition C only prohibits two thematic roles from projecting to a single argument *if they stand in a* c-*command relation.* This point will become important when we consider the anti-*c*-command condition on parasitic gaps in Part II below. It is also important in relation to the phenomenon of control, to which we turn next.

CONTROL: Many authors who have appealed to versions of the obliqueness hierarchy in theories of binding have pointed out that the phenomenon of control, as exhibited by verbs like *persuade*, can be analysed in similar terms, and have argued that such bounded dependencies are base-generated and mediated in lexical semantics. A first approximation to the argument in present terms might be to stipulate a category $(S\backslash NP)/VP'$ for the subject control verb *tried* in sentences like *Keats tried to go.* However, if the control relation is to be defined at the level of the lexicon, then the interpretation of VP' must have a subject – that is, it must be a *property*, or function from entities into propositions. If so, then we should replace it with the following more explicit category:

(20) tried := $(S\backslash NP)/(S_{to-inf}\backslash NP)$

A first approximation to the semantics of this category might be the following:

(21) $(S : try'\ s\ x \backslash NP : x)/(S_{to-inf} : s \backslash NP : x)$

We may assume that the infinitival VP *to go* has the following category:

(22) $S_{to-inf} : go'\ y \backslash NP : y$

This category, like all other predicates, binds the subject of the interpretation to that of the syntactic subject. If the earlier category for *tried* combines with this infinitival, we would therefore get the following category for *tried to go.*

(23) $S : try'\ (go'\ z)\ z \backslash NP : z$

However, the interpetation $try'(go'\ z)\ z$ of the result is in violation of condition C of the Binding Theory, because the two instances of the variable z stand in a c-command relation. One possibility here is to follow Chierchia and Jacobson in adopting a property theory of VP, carrying out the binding of of the property extra-syntactically, via lexical entailments, without the essential use of variables. Alternatively, we can follow GB in adopting a PRO analysis, via the following alternative category for the control verb *tried:*[9]

(24) tried $:= (S : try'\ s\ x \backslash NP_{agr} : x)/(S_{to-inf} : s \backslash NP_{agr} : PRO\ x)$

There is one slight departure from the standard analysis. Rather than merely using a constant PRO to represent the controlled argument, leaving it to the binding theory or an autonomous module of control theory to establish the antecedent, we have made the lexical entry for the control verb do part of that work, by making the subject a proterm of exactly the same type $PRO\ x$ as an anaphor.[10] Note that the exemption clause in Condition A of the binding theory 18 allows *subject* PRO terms to be bound from outside their verb, since there is by definition no available c-commanding argument within.

[9] The notation s for the translation of the result of the infinitival argument of the control verb, rather than a term $iv\ (PRO\ x)$, leaves the specification of the control relation itself implicit. The abbreviation is possible because s will be coerced to such a term when the control verb applies to a suitable infinitival complement like *to go*, 22.

[10] The idea that control verbs in some sense involve implicit anaphors is as old as the equi transformation itself.

The category 24 combines with the infinitival 22, unifying the terms $PRO\ x$ in the former and y in the latter to yield the following category for the VP *tried to go*:

(25) $S : try'\ (go'\ (PRO\ z))\ z \backslash NP_{agr} : z$

The same arguments apply to the analysis of object control. The following is the full category of the verb *persuades* $((S\backslash NP)/(S\backslash NP))/NP$, reflecting the assumption that interpretations observe the obliqueness hierarchy:

(26) persuades $:= ((S : persuade'\ s\ x_1\ x_2 \backslash NP_{agr_2} : x_2)/(S_{to-inf} : s \backslash NP_{agr_1} : PRO\ x_1))/NP_{agr_1} : x_1$

The category above embodies a "wrap" analysis of object control verbs, akin to that proposed by Bach 1979, 1980, Dowty 1982, Szabolcsi 1989, and Jacobson 1987, 1990, albeit at the level of lexical interpretation rather than syntactic or phrasal derivation. That is, the command relation between the interpretation of the object NP and the predicate argument is *reversed* with respect to the derivation. Once again, the use of variables in this category is non-essential, as the comparison with these authors suggests.

When applied to an object like *Keats* and an infinitival like *to go*, 22 above, the category 26 gives rise to derivations like the following:[11]

(27)

```
                          persuades                              Keats        to go
----------------------------------------------------- --------- -----------------
((S:persuade' s x1 x2\NP:x2)/(Sto-inf:s\NP:PRO x1)/NP:x1 NP:keats' Sto-inf:go' y\NP:y
--------------------------------------------------------------->
       (S:persuade' s keats' x2\NP:x2)/(Sto-inf:s\NP:PRO keats')
       --------------------------------------------------------------------------->
                          S:persuade'(go' (PRO keats')) keats' x2\NP:x2
```

On the assumption that the binding of reflexives and their agreement with that binder can be handled in a way that we have yet to specify, such derivations will interact correctly with the binding theory, since they allow subject agreement and the local binding of the reflexive to be transmitted upward boundedly via the NP argument-term of the predicate and its PRO-term interpretation.

The above analysis combines many of the virtues of VP and S analyses of control, in that there is no surface category corresponding to the infinitival

[11] Agreement is suppressed.

subject, but the predicate category of the VP is captured at the level of interpretation via the category $S\backslash NP$. While we shall continue to use the categories VP, VP_{ing}, VP_{to-inf}, (etc.) as abbreviations wherever syntax alone is at issue, we shall always assume that the true categories are of the form $S_x\backslash NP$.[12]

While many questions concerning binding and control remain open at this point, we will postpone further discussion until the combinatory mechanism that handles unbounded dependencies and coordinate constructions has been introduced.

¶1.2 The Combinatory Generalisation

COORDINATION: To extend such grammars to cope with coordination we need a rule, or rather a family of rules, of the following form:[13]

(28) COORDINATION (<&>):

$$\begin{array}{llll} X : x_1 & conj \;\; X : x_2 & \Rightarrow_{\mathbf{\Phi 0}_{and'}} & X : and'\; x_2\; x_1 \\ X : x_1\backslash Y : y & conj \;\; X : x_2\backslash Y : y & \Rightarrow_{\mathbf{\Phi 1}_{and'}} & X : and'\; x_2\; x_1\backslash Y : y \\ (X : x_1\backslash Y : y)/Z : z & conj \;\; (X : x_2\backslash Y : y)/Z : z & \Rightarrow_{\mathbf{\Phi 2}_{and'}} & (X : and'\; x_2\; x_1\backslash Y : y)/Z : z \\ (etc.) & & & \end{array}$$

This rule captures the ancient intuition that *coordination is an operation which maps two constituents of like type (but different interpretations) onto a constituent of the same type.* Given such a rule or rules, derivations like the following are permitted, and yield semantically correct results (semantics

[12]The proposal that bounded phenomena like control are properties of the interpretation as defined in the lexicon, is by no means original. It is implicit in proposals by Brame 1976, and by Keenan & Faltz 1978 and Bach and Partee 1980, 1981, within a Montague framework. It is also the proposal that lies at the heart of the LFG account of binding and control, as in Bresnan et al. 1982, and has been extended recently by Pollard and Sag 1991.

[13]The rules as given are a simplification in two respects. They do not represent syntactically the "prepositional" or "proclitic" character of the English conjunctions, which associate to the right, (although this property is reflected at the level of interpretation). The rules also finess a number of well-known problems in specifying such rules in terms of unification. See Steedman 1990 for a more complete proposal. The $\mathbf{\Phi}^n$ family of combinators are closely related to **S** in the scheme of Curry & Feys (1958 discussed below), and correspond to the different instances of coordination discussed here.

is omitted from the body of the derivation to save space, as are the details of agreement):

(29)
```
       Keats   cooked   and    ate      apples
       ----- --------- ---- ---------- ------
        NP   (S\NP)/NP conj (S\NP)/NP    NP
             ------------------------<&>
                    (S\NP)/NP
                    ----------------------->
                               S\NP
---------------------------------------------<
S:and'(eat' apples' keats')(cook' apples' keats')
```

It would be nice if all coordination could be handled this simply, as constituent coordination, without movement or deletion. However sentences like the following appear to present difficulties for this proposal, since they show that substrings that are *not* normally regarded as constituents can, nevertheless, coordinate:

(30) a. Keats cooked, and *might eat*, some apples
b. Chapman cooked, and *Keats ate*, some apples
c. Keats will copy, and *file without reading*, some articles concerning Chapman.

The combinatory generalisation of categorial grammar adds exactly three further classes of combinatory rule to the context-free core. These rules have the effect of making such substrings into grammatical constituents in the fullest sense of the term, complete with an appropriate and fully compositional semantics, so that they too can coordinate without the intervention of movement or deletion. All of the rule-types adhere to the following restrictive assumption:

(31) THE PRINCIPLE OF ADJACENCY: Combinatory rules may only apply to entities which are linguistically realised and adjacent.

COMPOSITION – THE BLUEBIRD, **B**: The first such rule-type is motivated by examples like 30a, above. Rules of functional composition allow functional categories, like *might*, to combine with functions into their argument categories, such as *eat*, to produce non-standard constituents corresponding to such strings as *might eat.* The rule required here (and the most commonly used functional composition rule in English) is written as follows:

(32) FORWARD COMPOSITION (>B):
$X/Y \quad Y/Z \quad \Rightarrow_{\mathbf{B}} \quad X/Z$

The rule permits the following derivation for example 30a:

(33)
```
Keats    cooked    and   might      eat  some apples
----- --------- ---- ---------  ----- -----------
 NP   (S\NP)/NP conj (S\NP)/VP  VP/NP     NP
                     ---------------->B
                           (S\NP)/NP
      ------------------------------<&>
                   (S\NP)/NP
                   ---------------------------->
                             S\NP
-------------------------------<
               S
```

The combinatory rule and its application in the derivation are indexed as **B** because that is Curry's name for this combinator in the system of Combinatory Logic (Curry and Feys 1958).[14] The formalism immediately guarantees without further stipulation that this operation will compose the interpretations, as well as the syntactic functional types. For example, if we blow up the crucial step in the above derivation to show the interpretations, it appears to a first approximation as follows (recall that VP is an abbreviation for the category $S_{inf} \backslash NP_{agr}$):[15]

(34)
```
                        might                                         eat
-------------------------------------------------  --------------------------------------
(S:might' s np1\NPagr:np1)/(Sinf:s\NPagr:PRO np1)  (Sinf:eat' np2 np3\NPagr:np3)/NP:np2
-------------------------------------------------------------------------------------->B
                      (S:might'(eat' np2 (PRO np1)) np1\NP:np1)/NP:np2
```

The result of the composition has the same syntactic type as a transitive verb, so when it is applied to an object and a subject, it is guaranteed to

[14]Curry 1958, p.184, fn., notes that he called the operation **B** because that letter occurs prominently in the word "substitution", and because the names **S** and **C** were already spoken for. The operation is Smullyan's 1985 Bluebird.

[15]The modal is analysed as a subject-control verb, which binds the subject of the infinitival VP to a *PRO*-term. (The analysis has a precedent in the "auxiliaries as main verbs" analysis of McCawley, and in the LFG analysis of these phenomena – cf. Bresnan 1982 and widespread recent forpoals for "VP-internal" subjects in GB). As in the case of other control verbs like *try* 24 and *persuade* 26, subject agreement is transmitted via the feature *agr*.

yield exactly the same interpretation for the sentence *Keats might eat some apples* as would have been obtained without the introduction of this rule. The reader can easily satisfy themselves that this result preserves the canonical *c*-command relations that would result from a derivation using functional application alone.

Moreover, this non-standard verb *might eat* is now a consituent in every sense of the word. It can therefore coordinate with other transitive verbs like *cooked* and take part in derivations like 33. Since this derivation is in every other respect just like the derivation in 29, it too is guaranteed to give a semantically correct result, and to preserve canonical *c*-command relations in that result.

As a result, it interacts correctly with the binding theory, transmitting the binding of anaphors via the PRO-terms clause-boundedly, as in the following examples:

(35) a. Sammy might find himself/*herself.
b. Sammy might try to find himself/*herself.
c. *Sammy said Rosie might try to find himself.

It is important to realise that these examples depend on *himself* being bound to the (PRO-term) subject of *find himself*, which in turn is controlled by the subject of *might*, to yield a term like the following for the VP *might find himself*:

(36) $S : might'(find'(PRO\ (PRO\ np1))(PRO\ np1))\ np1\backslash NP_{3sm} : np1$

The clause-boundedness of this type of reflexive anaphora is thereby guaranteed.

TYPE-RAISING – THE THRUSH, **T**: The second novel kind of rule that is imported under the combinatory generalisation is motivated by examples like 30b above, repeated here:

(37) Keats cooked, and Chapman ate, some apples

If we are to maintain the assumption that everything that can coordinate is a constituent formed without deletion or movement, then *Keats* and *cooked* must also be able to combine to yield a constituent of type S/NP, which can combine with objects to its right. The way this is brought about is by adding instances of rules of type-raising to the system, including some of the following general form:[16]

(38) FORWARD TYPE-RAISING (>T):
$$Y \Rightarrow_{\mathbf{T}} X/(X\backslash Y)$$

One instance of this rule makes the subject NP into a function over predicates. Subjects can therefore compose with functions *into* predicates – that is, with transitive verbs, as in the following derivation for 37:[17]

(39)
```
  Keats    cooked   and  Chapman    ate    some apples
-------- --------- ---- -------- --------- -----------
   NP    (S\NP)/NP conj    NP    (S\NP)/NP     NP
-------->T               -------->T
S/(S\NP)                 S/(S\NP)
------------------->B    ------------------->B
        S/NP                   S/NP
        ---------------------------<&>
                   S/NP
                   ------------------------------->
                               S
```

The combinatory rule and its application in the derivation are indexed as **T**.[18] The semantics of the type-raised subject in the derivation is as in the following expanded derivation for the substring *Keats cooked*:

(40)
```
      Keats                           cooked
   ------------      ----------------------------------
   NP3sm:keats'      (S:cook' np1 np2\NPagr:np2)/NP:np1
---------------------->T
S:s/(S:s\NP3sm:keats')
---------------------------------------------------------->B
               S:cook' np3 keats'/NP:np3
```

[16]Again this is a schema, not a single rule. We shall see below that Y is limited to categories which constitute arguments of verbs.

[17]Agreement is ignored as usual.

[18]The rule was called $\mathbf{C}_*$ by Curry, and is Smullyan's Thrush. Type-raising is of course widely used in Montagovian semantics. We shall see below that this rule, unlike the other combinators, should probably be regarded as operating pre-syntactically, to a fixed set of argument types such as NP in the lexicon.

The result is therefore guaranteed to be a function which when it reduces with an object *some apples* will yield the same result that we would have obtained from the traditional derivation shown in 7, namely the following:

(41) $S : cook'\ apples'\ keats'$

(And of course, the same facts guarantee that the coordinate example 37 will deliver an appropriate interpretation.) As in the earlier example, it is important to notice that it is at the level of the interpretation of this S that traditional constituents like the VP, and relations such as *c*-command (or equivalently, *F*-command), continue to be embodied. This is an important observation, since as far as surface structure goes, we have now compromised both, and it is now *only* at the level of interpretation that the binding theory can apply.

SUBSTITUTION – THE STARLING, **S**: The third and final variety of combinatory rule is motivated by examples involving "parasitic" dependencies like 30c, repeated here:

(42) Keats will copy, and file without reading, some articles concerning Chapman

Under the simple assumption with which we began, that only like *constituents* can conjoin, the substring *file without reading* must be a constituent formed without movement or deletion. What is more, it must be a constituent of the same type as a transitive verb, VP/NP, since that is what it coordinates with. It follows that the grammar of English must include the following operation, which is of a class first proposed by Szabolsci (1983, 1989):[19]

(43) BACKWARD CROSSED SUBSTITUTION (<Sx)
$Y/Z \quad (X\backslash Y)/Z \quad \Rightarrow_{\mathbf{S}} \quad X/Z$
where $Y = S_x\backslash NP$

(The restriction on the rule is discussed below.)

[19]The name "substitution" was proposed for the combinator **S** in homage to Curry's explanation (see note 14) of his choice of the name **B** as deriving from this word, and because **S** is the general form of the operation of which **B** is a special case. Schönfinkel 1924 called it *Verschmeltzung*, or "fusion". Kaplan 1975 called it "composition" (!), and Szabolcsi 1983, 1989 calls it "connection". It is Smullyan's Starling.

This rule permits the following derivation for the sentence:[20]

(44)

```
Keats will  copy  and  file    without     reading,  some articles
---------- ----- ---- ----- ------------- --------  -------------
   S/VP    VP/NP conj VP/NP (VP\VP)/VPing VPing/NP       NP
                            ----------------------->B
                                  (VP\VP)/NP
                      ------------------------<Sx
                                VP/NP
           ---------------------------<&>
                       VP/NP
    ------------------------->B
                S/NP
                ------------------------------------------->
                                  S
```

It is important to notice that the crucial rule resembles a generalised form of functional composition, but that it *mixes* the directionality of the functors, combining a leftward functor over VP with a rightward function into VP. We must therefore predict that other combinatory rules, such as composition, must also potentially have such "crossed" instances.

The restriction on rule 43 permits only categories of the form $S_x\backslash NP$ (equivalently, VP_x) to unify with the variable Y. It has the effect of excluding the analogous derivation of

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

Similar examples are discussed by Frank 1991. He contrasts them with examples like the following, which seem much better:

(46) ?A building which the mayor denounced the landlord's [neglecting]$_{VP_{ing}/NP}$ [after promising to repair]$_{(VP_x\backslash VP_x)/NP}$

The latter will be allowed by the rule on the assumption that gerunds are $S_{ing}\backslash NP$ – here abbreviated VP_{ing} –, and that possessives are functions of type NP/VP_{ing}, a proposal which is implicit in the analysis of Abney 1987.

20 Infinitival and gerundival predicate categories are abbreviated as VP and VP_{ing}, and the NPs are shown un-type-raised, for ease of reading.

Like the other combinatory rules, the substitution rule combines the interpretations of categories as well as their syntactic categories. Thus, the function *without reading* is assembled by functional composition as follows:[21]

(47)

```
                         without                                         reading
-------------------------------------------------------  ------------------------------
((Sx:without' s1 s2\NP:x3)\(Sx:s2\NP:x3))/(Sing:s1\NP:x3)  (Sing:read' x5 x4\NP:x4)/NP:x5
----------------------------------------------------------------------------------------->B
              ((Sx:without'(read' x5 x3) s2\NP:x3)\(Sx:s2\NP:x3))/NP:x5
```

The phrase *file without reading* is then derived as follows using the Backward Crossing Substitution rule:

(48)

```
           file                                     without reading
------------------------------  ------------------------------------------------------
(Sinf:file' x2 x1\NP:x1)/NP:x2  ((Sx:without'(read' x5 x3) s2\NP:x3)\(Sx:s2\NP:x3))/NP:x5
-----------------------------------------------------------------------------------------<Sx
              (Sinf:without'(read' x2 x1)(file' x2 x1)\NP:x1)/NP:x2
```

It follows that if this consituent *file without reading* is combined with an object *some articles* on the right, and then combined with the S/VP *Keats will*, it will yield the following interpretation, which correctly entails that Keats will file the articles without Keats reading the articles:

(49) $S : will'(without'(file'\ articles'\ (PRO\ keats'))(read'\ articles'\ (PRO\ keats'))\ keats')$

It follows as usual that a similarly correct interpretation will be produced for the coordinate sentence 42.

¶1.3 Some Consequences of the Generalisation

CCG and Universal Grammar: These three classes of rule – composition, type-raising, and substitution – constitute the entire inventory of combinatory rule-types that CCG adds to pure categorial grammar. The earlier papers show that the three types of rule are limited by two principles, over and above the Principle of Adjacency 31.[22] The principles are the following:

[21]As usual, VP and VP_{ing} are an abbreviation for $S_{inf}\backslash NP$ and $S_{ing}\backslash NP$. As usual, S_x is an S which is underspecified on the relevant feature(s). Agreement is omitted. It is worth noting that while nothing stops us from following the mainstream in assuming that the subject of the adjunct is a PRO-term, rather than a simple variable, nothing forces this assumption either: the relevant variable x_1 conforms to Condition C.

[22]See Steedman 1987, 1990 and especially 1991b where the principles are formalised in unification-based terms.

(50) THE PRINCIPLE OF DIRECTIONAL CONSISTENCY: All syntactic combinatory rules must be consistent with the directionality of the principal function.

(51) THE PRINCIPLE OF DIRECTIONAL 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)

Together they amount to a simple statement that *combinatory rules may not contradict the directionality specified in the lexicon.* In Steedman 1991b, I argue that this simply reflects the fact that directionality is a property of *arguments* in the functor types.

The principles permit the following instances of the two syntactic combinatory rule-types:

(52) FUNCTIONAL COMPOSITION:

a. $X/Y \quad Y/Z \quad \Rightarrow_{\mathbf{B}} \quad X/Z \quad (>\mathsf{B})$

b. $X/Y \quad Y\backslash Z \quad \Rightarrow_{\mathbf{B}} \quad X\backslash Z \quad (>\mathsf{Bx})$

c. $Y\backslash Z \quad X\backslash Y \quad \Rightarrow_{\mathbf{B}} \quad X\backslash Z \quad (<\mathsf{B})$

d. $Y/Z \quad X\backslash Y \quad \Rightarrow_{\mathbf{B}} \quad X/Z \quad (<\mathsf{Bx})$

(53) FUNCTIONAL SUBSTITUTION

a. $(X/Y)/Z \quad Y/Z \quad \Rightarrow_{\mathbf{S}} \quad X/Z \quad (>\mathsf{S})$

b. $(X/Y)\backslash Z \quad Y\backslash Z \quad \Rightarrow_{\mathbf{S}} \quad X\backslash Z \quad (>\mathsf{Sx})$

c. $Y\backslash Z \quad (X\backslash Y)\backslash Z \quad \Rightarrow_{\mathbf{S}} \quad X\backslash Z \quad (<\mathsf{S})$

d. $Y/Z \quad (X\backslash Y)/Z \quad \Rightarrow_{\mathbf{S}} \quad X/Z \quad (<\mathsf{Sx})$

Any language is free to restrict these rules to certain categories, as we did with the backward crossed substitution rule 43, or even to entirely exclude a given rule type. But the above is the entire catalogue of rule-types.[23]

[23] I have not here discussed the generalisation of these rule-types to cover rules corresponding to combinators like $\mathbf{B}^2$, and $\mathbf{S}^2$. The generalisation has the effect of permitting composition into all lexical verb types, and is crucial to the analysis of Dutch verb-raising (Steedman, 1985), and of strings like the following:

(i) a man to whom I will show, and may give, my Swiss Army knife

We will assume that the principles of Consistency and Inheritance further restrict the rules of type-raising to the following two "order-preserving" cases.[24]

(54) TYPE-RAISING:

a. $X \Rightarrow_{\mathrm{T}} \mathrm{T}/(\mathrm{T}\backslash X)$ $(>\mathtt{T})$

c. $X \Rightarrow_{\mathrm{T}} \mathrm{T}\backslash(\mathrm{T}/X)$ $(<\mathtt{T})$

The earlier papers show that all four types of composition are implicated in the grammars of various languages. Dowty 1988 and Steedman 1985, 1990 show that the existence of numerous so-called "gapped" constructions and their cross-linguistic dependence on base constituent order follow immediately.

For example, by including the backward composition and type-raising rules 52c and 54b, as well as the forward versions that were seen earlier, the existence of the following coordinate construction is captured without further stipulation, as noted by Dowty 1988:[25]

(55)

```
   give        Deadeye Dick     a sugar-stick  and   Mexican Pete       a bun
---------- -----------------<T ------------<T ---- -----------------<T --------<T
(VP/NP)/NP (VP/NP)\((VP/NP)/NP)  VP\(VP/NP)  conj (VP/NP)\((VP/NP)/NP) VP\(VP/NP)
           ------------------------------<B       ---------------------------<B
                 VP\((VP/NP)/NP)                            VP\((VP/NP)/NP)
                 ------------------------------------------------------------<&>
                                     VP\((VP/NP)/NP)
-------------------------------------------------------------<
                      VP
```

This and other related examples, which notoriously present considerable problems for other grammatical frameworks (cf. Hudson 1982), are extensively discussed elsewhere, and constitute the primary reason for taking type

This detail is discussed in the earlier papers, where there is also some discussion of whether non-order-preserving instances of type-raising are permitted in principle. See Joshi et al. 1991 for discussions of power and complexity of CCG, including a polynomial time worst-case parsing result.

[24]The way in which the principles impose this restriction, at least in configurational languages, is discussed at length in Steedman 1991b.

[25]Throughout the paper, raised types are shown fully instantiated to aid readability, although we assume that they can be schematised, as in the rules 54.

raising and composition as the primitives of Combinatory Categorial Grammar.

The ubiquity of type-raised arguments in explaining the above phenomena is our first hint that argument categories should be regarded as *lexically* raised, as proposed by Karttunen 1989. The implication is that determiners and prepositions are functions *into* raised categories. Since such categories become rather unreadable, it will often be convenient to abbreviate them as $NP^{\uparrow}/N$ and $PP^{\uparrow}/NP^{\uparrow}$.[26]

Earlier work in the present framework shows that the inclusion of this particular set of operations makes a large number of correct predictions concerning extraction. Most basically, the analysis immediately entails that the dependencies engendered by coordination will be unbounded, and free in general to apply across clause boundaries. Thus all of the following examples, parallel to the triple 30 with which we began the section, are immediately accepted as well, without any further addition to the grammar whatsoever:

(56) a. Keats cooked, and *suspects that Chapman will eat*, the apples.
b. Keats cooked, and *I suspect that Chapman will eat*, the apples.
c. Keats wrote, and *I suspect that Chapman will file without reading*, an article on the habits of nightingales.

It should also be obvious that we have here everything that we need in order to capture *leftward* extraction in the related relative clauses, which are of course similarly unbounded:

(57) a. a man who [suspects that Chapman will eat the apples]$_{S\backslash NP}$
b. the apples that [Keats suspects that Chapman will eat]$_{S/NP}$
c. some articles that [Keats suspects that Chapman will file without reading]$_{S/NP}$

We can do so simply by assuming that relative pronouns are functions from $S\backslash NP$ and S/NP into noun modifiers $N\backslash N$. We therefore predict considerable symmetry between the scope of relativisation and right-node raising. We return to this question, together with the whole issue of the origin of certain well-known constraints on leftward and rightward extraction (and certain asymmetries between the two) in part II below.

[26]This use of $\uparrow$ is unrelated to that of Moortgat 1988b.

BINDING THEORY REVISITED: We now have all the pieces in place that are needed to fill in some of the details in the sketch of a binding theory for CCG that we left unspecified earlier. We will follow Szabolcsi 1989 in assuming that reflexives bear *only* type-raised categories. Most basically, anaphors must bear the following "accusative" category:[27]

(58) $(S : tv\ (PRO\ y)\ y\backslash NP_{3sm} : y)\backslash((S : tv\ (PRO\ y)\ y\backslash NP_{3sm} : y)/NP_{3sm} : PRO\ y)$

This category permits bounded anaphora in sentences like *Keats must shave himself.* It can only apply to a lexical verb, because only lexical verbs have an interpretation of the form $tv\ x\ y$. Its binding to the subject of that verb is passed up through the auxiliary by the usual control mechanism. This mechanism correctly fails to permit the unbounded equivalent **Sammy says that Rosie must shave himself*, in accordance with Condition A.

To capture sentences like *Keats showed himself a movie* we need a further "dative" type-raised category, as follows:

(59) $((S : tv\ (PRO\ y)\ y\backslash NP_{3sm} : y)/X)\backslash(((S : tv\ (PRO\ y)\ y\backslash NP_{3sm} : y)/X)/NP_{3sm} : PRO\ y)$

The extra argument X is an argument of any type and interpretation (such as NP : *a movie*), the interpretation appearing in the expression tv as a less oblique argument of the verb.[28] The category permits derivations like the following:

(60)
```
 Keats        showed                himself                  a movie
-------- -------------- -------------------------- ------------------
S/(S\NP) ((S\NP)/NP)/NP ((S\NP)/X)\(((S\NP)/X)/NP) (S\NP)/((S\NP)/NP)
         -----------------------------------------<
                      (S\NP)/NP
                      ---------------------------------------------------<
                                        S\NP
----------------------------------------->
                    S
```

[27]Szabolcsi handles the semantics differently, using combinators rather than variables.

[28]This is a clumsiness in the present notation, since this is still essentially a type-raised category. Presumably, the subject controlled anaphor categories 58 and 59 can be schematised over.

The reader will easily be able to satisfy themselves that the result carries the correct interpretation $show1'(a'\ movie')(PRO\ keats')keats'$, and that the category 59 also accepts *Sammy showed himself to Rosie* and *Gilbert persuaded himself to like George.* We need a few more related categories to capture families of sentences including *Sammy showed Rosie himself*, and *Keats bet himself a bottle of champagne that it was Thursday.*[29]

Since the maximum number of arguments a verb can have is four, and since condition A limits us to categories in which the binder is less oblique than the bindee, there is a maximum of ten fully instantiated categories needed to capture all bounded reflexive NP anaphora. However, if we attempt to capture the non-subject-controlled binding in 9a, *I showed the dogs (to) themselves*, via the following category for the reflexive, we encounter a problem.

(61) $(S : dtv\ (PRO\ x)\ x\ y \backslash NP : y) \backslash ((S : dtv\ (PRO\ x)\ x\ y \backslash NP : y)/NP_{3pl} : PRO\ x)$

The category is semantically impeccable, but fails to enforce number agreement syntactically. We must therefore either assume that this particular type of agreement is semantic, or we must assume that the non-subject controlled anaphor is a homophone bearing a different category. The latter assumption is supported by the observation that non-subject controlled coreference can take non-*c*-commanding antecedents in English, as in the following:

(62) a. A picture of Chapman by himself adorned the mantlepiece.
b. A picture of himself by Chapman adorned the mantlepiece.
c. The pictures of himself at the Rickmansworth Young Conservatives' Ball embarrassed Chapman.

The implication that the homophony of these two reflexives in English is accidental is supported by the observation that in Dutch, the true subject-controlled anaphor *zich* is lexically distinct from another reflexive *zichself*, which can be bound to non-subjects and non-*c*-commanders (cf. Koster 1987, p.326). [30]

[29] Again, we can presumably schematise over such subject-controlled anaphors

[30] One candidate for a bounded non-subject controlled anaphor category is to make it a function that maps type-raised objects like *the dogs* onto a non-standard constituent *the*

As noted earlier, the present paper follows Pollard and Sag (and diverges from Szabolcsi) in assuming that the so-called pied-piping of reflexives is more restricted than that of *wh* and that all anaphora in NPs like *a picture of themselves/each other* comes under the exempt heading, however that is mediated. Furthermore, to capture the fact that some PPs like *to themselves/each other* behave as as non-exempt bounded anaphors, we will follow Pollard and Sag 1992, p.286 in treating those prepositions which do pied-pipe reflexives as semantic *identity functions*, This means that the lexically raised category that was earlier assigned to prepositions as functions from raised NP categories to raised PP categories $PP^{\uparrow}/NP^{\uparrow}$, can be written as the following schema, in which T is a variable:

(63) to := $(\mathrm{T}\backslash(\mathrm{T}/PP : y))/(\mathrm{T}\backslash(\mathrm{T}/NP : y))$

If this function is applied to the accusative reflexive category 58, it permits the following derivation (in which this schema is instantiated for ease of reading):

(64)
```
Keats       talks                          to                              himself
-------- ---------- ------------------------------------------ ------------------
S/(S\NP) ((S\NP)/PP ((S\NP)\((S\NP)/PP))/((S\NP)\((S\NP)/NP)) (S\NP)\((S\NP)/NP)
                    ------------------------------------------------------------->
                                           (S\NP)\((S\NP)/PP)
           -----------------------------------------------------<
                                S\NP
------------------------------------->
                S
```

Since the variable y in category 63 gets bound to the term $PRO\ y$ in 58, the sentence *Keats talks to himself* will end up meaning the correct thing, namely *talks-to′* (*PRO keats′*) *keats′*, on the assumption that the verb has the following category:

(65) talks := $(S : \textit{talks-to}'\ x\ y\backslash NP : y)/PP : x$

dogs themselves of essentially the same type that was obtained via composition of type raised arguments in example 55 above. This analysis is compatible with the treatment of reflexive "pied-piping" below. The specification of this category, which is subtly different from that proposed by Szabolcsi 1989, p.307, is suggested as an exercise.

The schema 63 will similarly capture sentences like *Chapman read a poem to himself.*

The fact that both control and (non-exempt) anaphor binding have now been brought entirely within the lexicon suggests that we can simplify the Binding Theory in a number of ways. Most obviously, condition A can now be entirely excluded from syntax proper. This is in fact almost a forced move in the present theory. Consider the following sentences:

(66) a. Chapman is easy to please
b. Chapman tries to be easy to please

It is reasonable to assume that the interpretation of a, above, is something like *easy′(please′ chapman′ x)*, where the object of *please* is the surface subject. If so, then the interpretation of b must be something like the following, in which this argument becomes a controlled PRO-term inherited from the control verb via the syntactic subject of *to be easy to please.*

(67) *try′(easy′(please′ (PRO chapman′) x))chapman′*

This expression does not obey Principle A as it was defined earlier, since the PRO-term is not bound to an available local less oblique argument. Unless we exclude Principle A from syntax, we shall have to make otherwise unmotivated assumptions about the "arbitrarily interpreted" subject of *please* – for example, that it is expletive – or abandon the assumption that both control and anaphora are mediated by PRO-terms.[31]

However, once Principle A is restricted to the lexicon, we can simplify it still further. The exemption clause is no longer required, since we have assumed that the subjects of lexical infinitives are not PRO-terms, but are rather variables for which syntactic combination with a control verb is required for instantiation as a PRO-term. Principle A therefore reduces to a requirement that PRO-terms be bound in the lexicon, where lexical binding is defined as follows:

[31]Similar conclusions follow from the interaction of passives and control in sentences like *John tries to be loved*, and the assumption that *John is loved* has the interpretation *love′ john′ x*. I am indebted to Mark Hepple for discussions and advice about this problem.

(68) A pro-term with argument x in an interpretation in a lexical entry is lexically bound *either*: a) if the pro-term is c-commanded by an identical term x, *or*; b) if the pro-term is the interpretation of an argument of a function the interpretation of whose result r is c-commanded by an identical term x.

This definition sounds more complicated than it is. Case (a) is subsumed by the definition 17 of binding. Since all occurences of the term $PRO\ y$ are identical in the basic bounded accusative anaphor category 58, repeated here, and since at least one occurrence is bound/c-commanded by y, the (argument of the) PRO-term is lexically bound:

(69) $(S : tv\ (PRO\ y)\ y \backslash NP_{3sm} : y) \backslash ((S : tv\ (PRO\ y)\ y \backslash NP_{3sm} : y)/NP_{3sm} : PRO\ y)$

The second disjunct in 68 says that the argument of a pro-term is lexically bound if it is in a controlled complement whose result is c-commanded by the binder. Control is therefore an example of lexical binding, since the result s of the infinitival complement $S_{to-inf} : s\backslash NP_{agr_1} : PRO\ x_1$ is c-commanded by x_1 in the category 26, repeated here:

(70) persuades $:= ((S : persuade'\ s\ x_1\ x_2 \backslash NP_{agr_2} : x_2)/(S_{to-inf} : s\backslash NP_{agr_1} : PRO\ x_1))/NP_{agr_1} : x_1$

Such binding is not necessarily to a clausemate, but this definition embodies a form of subjacency restriction, since only an argument of a directly subcategorised function can be controlled.

This move in turn suggests a further simplification. Coordinate sentences parallel to 55 suggest that ordinary pronouns, like reflexives and other NPs, can also bear type-raised categories. The schema a, when instantiated by appropriate verbs, gives rise to categories parallel to 58 and 59, differing only in the lack of any binding:

(71) a. him $:= \mathrm{T}\backslash(\mathrm{T}/NP_{3sm} : pro\ y)$
b. him $:= (S : s\backslash NP : x)\backslash((S : s\backslash NP : x)/NP_{3sm} : pro\ y)$
c. him $:= ((S : s\backslash NP : x)/X)\backslash(((S : s\backslash NP : x)/X)/NP_{3sm} : pro\ y)$

Since these categories do not bind the pro-term, any pronominal binding is either of the discourse- or quantificational variety. We have conjectured that neither is mediated in syntax at all. Condition B therefore also belongs outside syntax, partly in the lexicon and partly in the theory of quantification and pragmatics. The latter systems presumably impose the condition by default, maybe because they lack any way of inducing local bindings in the first place.[32]

We might therefore consider reformulating the binding theory 18 as follows:[33]

(72) BINDING THEORY:

A: An anaphor-type pro-term $PRO\ x$ must be lexically bound.
B: A pronoun-type pro-term $pro\ x$ must not be lexically bound.
C: Nothing but (the argument in) a pro-term may be bound.

The only component of the binding theory thus stated that operates at all in syntax is Condition C. Since the only true binders are lexical, the condition as it applies to syntactic derivation entails that *no combinatory rule may unify two terms one of which* c*-commands the other.*

Within the degrees of freedom that we have exploited in choosing categories for subject-controlled anaphors and pronouns, a number of other proform categories can be specified. Condition A is no longer limited to bounded anaphora, so we can in principle define unbounded subject-controlled anaphors, as in the following accusative instances, in which the PRO-term is lexically bound under the definition 68. These differ from the bounded cases 58 etc only in lacking the clausemate restriction on the form of the resulting proposition:

(73) a. himself $:= (S : iv\ x\backslash NP : x)\backslash((S : iv\ x\backslash NP : x)/NP_{3sm} : PRO\ x)$
b. himself $:= ((S : iv\ x\backslash NP : x)/X)\backslash(((S : iv\ x\backslash NP : x)/X)/NP_{3sm} : PRO\ x)$

Such categories could be used to capture anaphors of the kinds that appear to exist in languages like Dutch, Norwegian and Icelandic (cf. Koster, 1987, Ch. 6, and Hellan 1988, Ch. 2, esp. p.87 ff. and references therein). The differences among the long-range anaphors available in such languages presumably arise from subtle differences in the specification of the antecedent via such discourse-related features as "experiencer" and "topic".

[32]This suggestion is supported by the fact that **Every man$_i$ likes him$_i$* is so much worse than **He$_i$ likes him$_i$*. The latter is possible in contexts which support a mutual discourse referent – but such a referent cannot by definition corefer with the bound variable in a quantifier.

[33]Binding is defined as in 17, lexical binding as in 68.

English subject-controlled "exempt anaphors" might also be captured in such terms. This proposal would immediately explain the sensitivity of exempt anaphora to the island constraints discussed in Part II, as illustrated below:

(74) a. Who did Sammy write a novel about?
b. Sammy wrote a novel about himself/*him.

(75) a. *Who did Sammy like Rosie's novel about?
b. Sammy liked Rosie's novel about *himself/him.

It would also correctly permit the following variety of "exempt anaphora", on the assumption that *realised that pictures of* is a composable constituent, and can act as the argument of category 73,b:[34]

(76) Chapman realised that pictures of himself were on sale in the foyer.

Certain other non-subject controlled varieties of English exempt anaphor, involved in sentences like the following, might be defined as pronouns, including in their interpretations not *PRO*-terms but *pro*-terms, bound outside syntax, but with similar restrictions to the experiencer:

(77) A picture of himself at the Rickmansworth Young Conservatives' Ball hung above the mantlepiece in Chapman's apartment.

However, all such finer points of the binding theory and their bearing on CCG must await discussion on another occasion, for we have accomplished the immediate goal of showing that CCG is compatible with a straightforward binding theory, and the diverse characteristics of anaphors in even those languages most closely related to English suggests that this discussion is unlikely to be brief.

¶1.4 Interim Summary

Because of the associativity of certain combinatory rules, CCG differs from most other theories in assigning many alternative surface structures for any

[34]See below for discussion of why the Subject condition holds notwithstanding.

given reading of a sentence, in some of which the object may structurally command the subject (or even a subject in a higher clause), as in the following minimal example:

(78) a.
```
mary     prefers   corduroy
-------- --------- --------
S/(S\NP) (S\NP)/NP T\(T/NP)
------------------>B
       S/NP
       ---------------------<
                S
```

b.
```
mary     prefers   corduroy
-------- --------- --------
S/(S\NP) (S\NP)/NP T\(T/NP)
         -------------------<
                   S\NP
-------------------->
          S
```

Such structures do not in general represent traditional notions of dominance and command. However, we have seen that, at the level of the interpretation, all these derivations yield the *same* function-argument structure, in which the relations of dominance and command over subjects and other elements hold in pretty much their traditional form, except that the obliqueness hierarchy is observed. It is at this level that the binding theory is defined.

The second part of the paper will show that a theory of this kind explains a number of well-known constraints on relativisation, and that the interactions between these two systems can in all cases attributed to the effect of the binding theory applying to the interpretations or function-argument structures that are delivered by surface structure derivations, not to the derivations themselves.

PART II

§2 Relativisation without ECP

We can assume on the basis of their semantics that nominative and accusative relative pronouns respectively bear the categories $(N\backslash N)/(S\backslash NP)$ and $(N\backslash N)/(S/NP)$. That is simply to say that they are functions from predicates to noun modifiers that can be written in full as follows:[35]

(79) a. *who,that,which* $:= (N : s\&(n\ x)\backslash N : n\ x)/(S : s\backslash NP : x)$
b. *who(m),that,which* $:= (N : s\&(n\ x)\backslash N : n\ x)/(S : s/NP : x)$

These categories accept the relative clauses in 57, repeated below:

(80) a. a man who *(suspects that Chapman) will eat the apples.*
b. the apples that *Keats (suspects that Chapman) will eat.*
c. some articles that *Keats (suspects that Chapman) will file without reading*

Note that the variable x in the interpretations will be carried through to the interpretation of nouns like *man who walks* and *man who Mary loves* , where it is implicitly λ-bound, as in the following:

(81) a. $N : (walks'\ x)\&(man'\ x)$
b. $N : (likes'\ x\ mary')\&(man'\ x)$

Such N-interpretations (which obey Principle C of the binding theory) are therefore properties that can be used directly to identify appropriate individuals in the model.[36]

¶2.1 Pied-Piping:

[35] A further relative pronoun category, which is required by examples like *packages [which I sent and which you carried]$_{(N\backslash N)/PP}$ to Philadelphia*, is ignored here. The details of long-range agreement in the relative pronoun categories 79 are discussed in the section on subject extraction below.

[36] It should be noted that we are assuming a further stage of semantic interpretation, which we might or might not want to identify with a structural level of "logical form", at which matters such as quantifier scope are further resolved. The present interpretation structures are thus reminiscent of Webber's 1978 and Schubert & Pelletier's 1982 "semantically ambiguous" translations. See Park ? for a possible realisation of this idea in CCG.

In order to capture pied-piping of relatives, we must assume, following Szabolcsi 1989, that the *wh*-relative pronouns have a further pair of type-raised categories that allow them to combine with functions over NPs such as *on*, *the covers of*, and *on the covers of*, to yield the usual relative pronoun categories. We can provisionally write the pied-piping categories as follows:[37]

(82) a. *who(m),which* $:= ((N : s\&(n\ y)\backslash N : n\ y)/(S : s\backslash NP : x))\backslash(NP : x/NP : y)$
b. *who(m),which* $:= ((N : s\&(n\ y)\backslash N : n\ y)/(S : s/PP : x))\backslash(PP : x/NP : y)$
c. *who(m),which* $:= ((N : s\&(n\ y)\backslash N : n\ y)/(S : s/NP : x))\backslash(NP : x/NP : y)$

We are assuming here that *unraised* prepositions and determiners are available for composition into entities like *the covers of*, and *on the covers of*, bearing the following categories:

(83) a. $NP : \mathit{the}'(\mathit{covers}'(\mathit{of}'\ y))/NP : y$
b. $PP_{on} : \mathit{the}'(\mathit{covers}'(\mathit{of}'\ y))/NP : y$

(This assumption is justified below, in the section on islands.) They yield the following categories for *the covers of which* and *on the covers of which*:

(84) a. $(N : s\&(n\ y)\backslash N : n\ y)/(S : s/NP : \mathit{the}'(\mathit{covers}'(\mathit{of}'\ y)))$
b. $(N : s\&(n\ y)\backslash N : n\ y)/(S : s/PP_{on} : \mathit{the}'(\mathit{covers}'(\mathit{of}'\ y)))$

These can combine with fragments like *Keats (expects that Chapman) will design/speak* in the usual way, to accept relative clauses like the following:

(85) a. A report the cover of which Keats (expects that Chapman) will design.
b. A subject on which Keats (expects that Chapman) will speak.

We also predict that pied piping will be sensitive to subjacency within the *Wh*-element in the same ways as extraction itself. Thus we expect the following asymmetry:

(86) a. (reports) [[the height of the lettering on the covers of]$_{NP/NP}$ which]$_{NP}$ the government prescribes.
b. *(reports) [[the woman that wrote]$_{*NP/NP}$ which]$_{NP}$ Keats met.

[37]Again, presumably such category sets can be schematised. We pass over the question of pied piping of possessives like *(a man) the latchet of whose shoes I am not worthy to stoop down and unloose*, merely noting that they can be handled on the assumption that *whose* is a function from N into the categories given here.

¶2.2 Subject/Object Asymmetries

Strong Crossover: The relative pronoun categories 79 capture without further stipulation the asymmetry between subjects and other arguments that gives rise to "Strong Crossover" phenomena, as exhibited in the following examples:

(87) a. *(a) man who(m)$_i$ he$_i$ thinks that Mary likes
b. (a) man who$_i$ thinks that Mary likes him$_i$

As in many other theories, this result follows from Condition C, which forbids the variable x in the following translations for the relevant nouns from being bound (see discussion of example 81):[38]

(88) a. $*N : (\mathit{thinks}'(\mathit{like}'\ x\ \mathit{mary}')\ (\mathit{pro}\ x))\&(n\ x)$
b. $N : (\mathit{thinks}'(\mathit{like}'\ (\mathit{pro}\ x)\ \mathit{mary}')\ x)\&(n\ x)$

[38]Cf. Szabolcsi, 1989, p.315 for a related proposal in a variable-free CCG. The present theory has nothing to say about *weak* crossover, the condition that is sometimes invoked in connection with the asymmetry between pairs of sentences like the following:

(i) a. who$_i$ loves his$_i$ mother?
b. ?who(m)$_i$ does his$_i$ mother love?

This phenomenon has been used to argue for the distinction between $A/\overline{A}$ positions, and hence for the distinction between D-structure (and, implicitly, the present notion of interpretation), and S-structure. However, it is far from clear that weak crossover is a phenomenon of the same system as *Wh*-movement. It is well-known that weak-crossover effects are not found in relative constructions (Chomsky 1982, p.93), or in topicalisations:

(ii) a. The boy who(m)$_i$ his$_i$ mother truly loves.
b. This boy$_i$, his$_i$ mother truly loves!

The present theory implies that ib, and iic and d, are equally well-formed syntactically, since all three conform to the Binding Conditions. However, it seems likely that weak crossover is a property neither of the systems responsible for *Wh* constructions, nor of the Binding Theory. It has often been noted (e.g. by Jackendoff 1972) that the unacceptability of ib seems to be of the same type as the unacceptability of the following with the bound variable reading:

(iii) His$_i$ mother loves every boy$_i$

It therefore seems likely that weak crossover belongs at the same level as quantifier scoping – perhaps at LF, as in Chomsky 1982.

The fact that interpretations respect obliqueness also correctly predicts the following strong crossover effect:

(89) a. *(a) man who(m)$_i$ I told him$_i$ that Mary liked
b. (a) man who$_i$ I told that Mary liked him$_i$

The interpretations are as follows:

(90) a. $*N : (tell'(like'\ x\ mary')\ (pro\ x)\ i')\&(n\ x)$
b. $N : (tell'(like'\ (pro\ x)\ mary')\ x\ i')\&(n\ x)$

On the assumption that pied-piping of relative pronouns is handled as in 82, we also exclude the following example (related to ones discussed by Safir 1986):

(91) *a man pictures of who(m)$_i$ he$_i$ (thinks that Mary)likes

On the assumption that *in situ Wh*-items also include a variable in their translation – so that for example the category of unrelativised *who(m)* is $NP : x$ or the type-raised equivalent – we also exclude the following:

(92) *He$_i$ thinks Mary likes who(m)$_i$?

EXTRACTION OF AND OUT OF SUBJECTS: Extraction out of subjects is not generally possible, even when similar extractions out of non-subjects are allowed, as shown by the following sentences, which are of a type that motivated the Subject Condition of Chomsky 1970:

(93) a. a man who(m) I like every friend of
b. *a man who(m) every friend of likes me.

Even on the assumption that type-raising is lexical, and that subjects like *every friend of* therefore have the category $(S/(S\backslash NP))/NP$, such extraction would require either an otherwise unmotivated additional lexical category for the relative pronoun, or type-raising of the predicate *votes Republican*, as the following blocked derivation reveals:

(94) ...*[who(m)]$_{(N\backslash N)/(S/NP)}$ [(I think) every friend of]$_{(S/(S\backslash NP))/NP}$ [votes Republican]$_{S\backslash NP}$

However, we have already seen that tensed predicates cannot type-raise. If they could then they would not only permit the above derivation, but also the corresponding Right Node Raising and Heavy NP Shift, as in example 115b.

A number of further constraints on long range dependencies that are asymmetrical with respect to subjects and objects, and which have been argued to stem from the Empty Category Principle, again arise in present terms because the categories reflect the different directionality of the subject and object arguments of the SVO verb. This ingredient of the theory captures very directly the concept of "canonical government configuration" or "direction of government" (cf. Kayne 1984, pp.167-169, Pesetsky 1982 and Koster 1987, p.19). In present terms, this principle is an inevitable consequence of the Principle of Inheritance and the feature-based analysis of directionality.

For example, as has been noted before, the theory predicts the following familiar asymmetry in extractability of English subjects and objects:

(95) a. (a man whom) [I think]$_{S/S'}$ [that]$_{S'/S}$ [Keats likes]$_{S/NP}$
b. *(a man whom) [I think]$_{S/S'}$ [that]$_{S'/S}$ [likes Keats]$_{S\backslash NP}$

Such asymmetries have been attributed to the Empty Category Principle (ECP) of Chomsky 1981. According to the present theory, they are possible in languages like English which have SVO lexicons, because the crucial compositions that potentially permit them require different instances of the composition rules. The non-extractability of the subject in a strongly configurational SVO language like English is, furthermore, a forced move, because a subject extraction like 95b would require the addition of the "forward crossing" composition rule 52b in order to compose the categories S/S and $S\backslash NP$:

(96) $X/Y \quad Y\backslash Z \;\Rightarrow\; X\backslash Z$

While such rules are, as we have seen, permitted (and in fact predicted) by the theory, a language like English cannot possibly allow such a rule to apply to the categories in b above. If it did so, then another distinguishing property of English, namely its configurationality, would be lost, for word order would immediately collapse entirely, allowing examples like the following:

(97) *Keats I [thinks (that) went home]$_{(S\backslash NP)\backslash NP}$

Thus the theory predicts that asymmetries in extractability for categories which are arguments of the same verb depend upon asymmetries in the directionality of those arguments. The fact that this particular asymmetry tends to be characteristic of configurational SVO languages and constructions therefore follows without the stipulation of any "empty category principle" and without any distinction of subject and object function argument relations in terms of "properness" of government or the $A/\overline{A}$ distinction (Chomsky, 1981.

However, this observation leaves unexplained the fact that English subjects can be extracted from bare complements:

(98) a. a man who(m) I think likes Keats
b. a man who(m) I think Keats likes

We cannot include such sentences by allowing a rule of crossing forward composition, no matter how restricted. Such a mechanism would immediately cause overgenerations parallel to 97. The only degree of freedom that remains within the present theory is to assume that this phenomenon arises in the lexicon. We must assume that verbs like *think* bear, in addition to categories like VP/S' and VP/S a special subject-extracting category of the following form:

(99) *think* $:= (VP/NP_{+wh,agr})/(S\backslash NP_{agr})$

In essence this category embodies the GPSG analysis proposed by Gazdar 1981, as modified by Hepple 1990 within a different categorial framework, and by Pollard and Sag, forthcoming. We shall see that there are some advantages to the present version.[39] The NP argument of this category bears a feature $+wh$, which prevents this argument from being saturated by

[39]More precisely, the relation that such categories bear to the basic VP/S categories is a first cousin to the "Slash Termination Metarule 2" of Gazdar et al (1985 – cf. Hukari & Levine 1987 for relevant discussion). The analysis differs from that presented in the published version of Steedman 1987, although it is essentially that presented in a widely circulated draft – see discussion by Bouma 1987, and Oehrle et al. 1990, whose objections to that proposal are met by Hepple's proposal and the version given here.

anything but a relative pronoun. The feature is in every respect exactly like the agreement features discussed earlier. Indeed, it must be accompanied by a number agreement feature, since *likes Keats* must agree with the head noun of the relative clause. We can conveniently write the argument in question as $NP_{+wh,agr}$. These features work in the following way.

The subscript *agr* is an abbreviation for a feature-value pair whose value is an as yet uninstantiated variable on a feature *AGR*. The variable may become instantiated by unification with a particular predicate. When it is so instantiated, the same variable on the "extracting" NP argument will be bound to the same number and person. This completely standard device excludes the following:

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

More interestingly, the subscript $+wh$ on the extracting NP argument is an abbreviation for a feature-value pair consisting of the value + on a feature WH which is an (optional) element of the category NP. This feature on the argument means that this function cannot combine with lexically realised NPs, all of which are assumed to be distinguished by the value − on this feature. (Like most such minor features, it will be omitted from the notation by convention whenever it is not directly relevant to the discussion. Nevertheless it is assumed to be present on *all* non-*wh* nominal categories). Thus the following derivation is blocked:

```
(101) *    I                   think                likes Keats  the man
      ---------- ------------------------------ ----------- ---------
      S/(S\NP1s) ((S\NP1s)/NP+wh,agr)/(S\NPagr)   (S\NP3s)   NP-wh,3s
                 ----------------------------------------->
                             (S\NP1s)/NP+wh,3s
      ------------------------------------------>B
                              S/NP+wh,3s
                              ------------------------------------*
```

By contrast, the object relative pronoun category does not restrict the NP argument in its argument S/NP on the feature *wh*. The syntactic category 79b, including the agreement features which were earlier suppressed, is written in full as follows:[40]

[40]The pied-piping categories 82 should also be modified to include agreement. We

(102) *who(m),that,which* := $(N_{agr}\backslash N_{agr})/(S/NP_{agr})$

The argument S/NP_{agr} of this relative pronoun category can unify with any argument term in a functor whether or not it is specified for agreement, and whether or not it is specified via the value + on the feature *wh*.[41]

The derivation of 98a is therefore allowed as follows:[42]

(103)
```
... man             who(m)               I think        likes Keats
    --- --------------------- ------------------------ -----------
    N3s (Nagr\Nagr)/(S/NPagr) (S/NP+wh,agr)/(S\NPagr)    S\NP3s
        ----------------------------------------------->B
               (Nagr\Nagr)/(S\NPagr)
               -------------------------------------------------->
                                   N3s\N3s
    ---------------------------------------<
                      N3s
```

In GB terms, what we have done is to distinguish the extracting subject, via the lexical entry for the verb, as an argument that can only be "antecedent-governed". In G/HPSG terms, we have defined a verb with a SLASH argument in the absence of any verb with a corresponding SUBCAT argument. However, we have accomplished this effect without invoking an empty category, without distinguishing between A and $\overline{A}$ positions, and using only the apparatus responsible for the equivalent of head government. We have also avoid the equivalent distinction between extracted and in situ arguments that is implicit in the latter theory's slash-feature-percolation apparatus.

continue to suppress the obvious but tedious details of exactly how *agr* can be partially specified for +/−animacy in the case of *who(m)* and *which*.

[41]Formally, the property of being unmarked or unspecified on the feature *wh* can be regarded simply as having a value on this feature which is "underspecified", and may unify with either + or −. In this case we must regard the feature with the underspecified value as being suppressed by convention in the present notation. Alternatively, we can represent the unmarked property as the complete absence of this feature-value pair from the syntactic category. In the latter case, we must use the PATR generalisation of term unification, which allows categories that do bear a given feature-value pair to unify with categories which lack the corresponding feature-value pair entirely (cf. Shieber, 1986).

[42]See Steedman 1987 for further details of nominative and accusative relative pronouns, including certain dialectal variations. The analysis interacts correctly with the theory of control, for example predicting the well-formedness of *a man whom it is easy to believe might win*. I am grateful to Polly Jacobson for this observation.

Thus the relative pronoun category 102 is *also* free to unify with functions whose argument is entirely unspecified on either feature. Since all non-subject verbal arguments in English are unspecified in this way, most relative clause residues are of this type, S/NP. An example is the following unchanged derivation for 98b, which depends on the ordinary bare complement category VP/S for the verb *think*:

(104) ... $[\text{who(m)}]_{(N_{agr}\backslash N_{agr})/(S/NP_{agr})}$ $[\text{I think Keats likes}]_{S/NP}$

Similarly, such underspecified functors can still combine with normal NPs, which we noted are always marked as $-wh$, as in the right node raised construction below:[43]

(105) $[\text{I think Keats likes, but you say he detests,}]_{S/NP}$ $[\text{the man in the grey flannel suit}]_{NP_{-wh}}$

Under the conventions just set out, all previous examples of relative clause derivations go through unchanged with the new version of the object relative category, and the ECP and the Subject Condition continue to be respected, without the stipulation of such conditions (or related devices like the Generalised Left Branch Condition of GPSG or the Trace Condition of HPSG) in the theory itself.

OTHER ASYMMETRICAL EXTRACTING SUBJECTS: In support of the above analysis of English subject extraction, it is vital to show that the degrees of freedom exploited in its specification are indeed degrees of freedom that are exploited in all available alternative ways by other constructions and other languages. It is assumed in the above analysis that sentences including bare complements with the subject NP *in situ* and those with an extracted subject involve different lexical categories for the verb. It follows that we must predict the occurence of verbs which only bear *one* of the two categories, and forbid one or the other analysis. Kayne 1984, p.111 points out that the French verb

[43]While sentences with multiple *wh*-elements, like *Which man gave what to whom?* are not discussed here, it is clear that such elements also have the category NP_{-wh} (or the order-preserving type-raised equivalent), just like full NPs. Otherwise they would permit multiple *Wh* questions like the following:

(i) *Who thinks likes Keats who?

Similarly, while topicalisation is not discussed here, the treatment in Steedman 1987 goes through with the category $S_t/(S/NP)$ for topics.

croire and the English verb *assure* are cases in point of verbs that *only* have the extracting category, and forbid a lexical NP in canonical position, citing examples like the following:[44]

(106) a. A man who(m) I assure you to be a genius
b. *I assure you Keats to be a genius

PSEUDOSUBJECTS: There is one final peculiarity of subjects that should be mentioned here. The fragment *I think that Keats* can be assembled into a constituent of type $S/(S\backslash NP)$, thereby permitting right node raising of the tensed VP, as in a, below, in an manner identical to subject coordination, as in b:[45]

(107) a. [You doubt that Keats,]$_{S/(S\backslash NP)}$ but [I wonder whether Chapman,]$_{S/(S\backslash NP)}$ walks.
b. [Keats]$_{S/(S\backslash NP)}$ and [Chapman]$_{S/(S\backslash NP)}$ walk.

The fact that such non-standard constituents have the same syntactic type as a type-raised subject threatens to allow illegal coordinations like the following:

(108) *[Keats,]$_{S/(S\backslash NP)}$ but [I wonder whether Chapman,]$_{S/(S\backslash NP)}$ walks.

However, it is clear from example 107 that "pseudo-subjects" like *I wonder whether Chapman* must differ categorially from true subjects, since coordination of two singular subjects changes the agreement to plural, but coordination of pseudo-subjects does not. The two must therefore be categorially distinct.

There is one crucial semantic difference between all type-raised nominal categories and all other categories of type $S/(S\backslash NP)$. The two categories in example 108 can be written in full as follows:

[44] I am assuming that the extracting argument is in addition in both cases marked for a $+wh$ argument, so that neither construction permits rightward movement, as in *?I assure you to be as sane as the next man the person who stands accused of this horrible crime.* If either *croire* or *assure* (or both) can in fact support rightward movement, then the analysis still goes through on the assumption that the extracting argument is unmarked on the feature *wh*. In any case, the existence of such verbs in SVO languages is predicted, and appears to offer a natural explanation for the conspiracy between subject extraction and subject inversion in Italian (Rizzi 1982, p.147).

[45] In 1987, I claimed that 107,a was bad. While the construction is clumsy, this now seems an aberration.

(109) a. $S : s/(S : s\backslash NP : keats')$
b. $S : wonder'(whether's)i'/(S : s\backslash NP : chapman')$

The nominal subject, unlike the pseudo-subject, has an identical interpretation for its own result and that of its complement predicate, a property which it is tempting to identify with that of bearing *case*. We may therefore assume that the instance of the coordination schema which applies to subjects (and changes agreement properties) only applies to categories that have this property. while the instance that applies to pseudo-subjects (which does not change agreement) cannot.[46]

¶2.3 ISLANDS:

The unboundeness of the dependencies involved in relaticisation is notoriously limited by "island constraints", which have been related to the principle of "subjacency" and the concept of "barrier", and to asymmetries between arguments and adjuncts. These have been discussed in categorial terms by Szabolcsi and Zwarts 1990, 1993, and Hepple 1990, and will be ignored here apart from the following general remarks.

ADJUNCT ISLANDS: The fact that both adjuncts and relative clauses are in general islands follows in English from the assumption that they are backward modifiers, as can be seen from the categories in the following unacceptable examples:

(110) a. * a book $[\text{which}]_{(N\backslash N)/(S/NP)}$ $[\text{I will}]_{S/VP}$ $[\text{walk}]_{VP}$ $[\textit{without reading}]_{(VP\backslash VP)/NP}$
b. * a book $[\text{which}]_{(N\backslash N)/(S/NP)}$ $[\text{I met}]_{S/NP}$ $[\text{a}]_{NP/N}$ $[\text{woman}]_{N}$ $[\textit{who wrote}]_{(N\backslash N)/NP}$

However, such examples are only blocked on the assumption that verbs like *walk* cannot type-raise over VP adjuncts, to become $VP/(VP\backslash VP)$, and that nouns like *woman* cannot raise over N adjuncts, to become $N/(N\backslash N)$. If they could acquire these categories, they could compose into the adjunct, allowing the extraction. This observation confirms the earlier assumption that only certain *argument* categories can type-raise, and that such type-raising is a *lexical* process.

[46]The best way to realise this proposal in unification terms is a question of some technicality, and will be discussed elsewhere. See Henderson 1990, p.81 for a proposal to reflect this distinction in syntax proper.

This view of the categories of arguments and adjuncts has the following further consequences. First, all lexically raised non-subject arguments become backward functors. The tendency of NPs to be "islands" with respect to extraction is thereby explained on the same grounds as that of adjuncts. So is the possibility of nominative vs non-nominative case on pronouns.

Second, the possibility of exceptions to the island status of NPs and adjuncts, and their equally notorious dependence on lexical content and such semantics-related properties as definiteness and quantification, must be explained either on the assumption that verbs can be selectively type-raised over such adjuncts, and lexicalised, or on the grounds that the exceptional items also have the unraised types. In particular, the fact that PPs *aren't* islands in English can be explained on the assumption that prepositions have both a raised category $PP^{\uparrow}/NP^{\uparrow}$ and an unraised category PP/NP_{+wh}, the latter permitting composition and therefore *Wh*-extraction.[47] The reason for restricting the category to relatives, via the feature-value $+wh$, is that preposition stranding rightward movement is not in general possible, a question to which we return in the next subsection.

Certain quantifier determiners, such as *every*, may have similar stranding categories, as is suggested by examples like the following:

(111) ?A woman whom I met every friend of.

On the other hand, the exceptional lexical sensitivity exhibited by "picture NP" extractions like the following suggests that even when stranding categories exist, the compositions that would allow extractions can be disfavoured on semantic grounds.[48]

(112) a. the man who(m) Chapman wrote a/?the/*my book about
b. ?the man who(m) Chapman burned a book about.

[47] Semantically, this category must, like its non-stranding relative, be an identity function, $PP : y/NP_{+wh} : y$.

[48] The suggestion that exceptions to subjacency are lexical in origin is closely related to the unification of notions of subjacency and government via the notion of "barrier" in Chomsky 1986b, p.10-16. The related idea that island phenomena are related to semantic interpretation goes back at least to Oehrle 1974 and Cooper 1982, and has more recently been advocated in Steedman 1987, and by Szabolcsi and Zwarts 1990, 1993.

Finally, we make an obvious prediction concerning the island status of adjuncts in languages where they are *forward*-looking functors. We will return to this prediction in the discussion of Dutch parasitic gaps below.

Certain finer details of preposition stranding in English and French can be treated using the apparatus introduced for subject extraction. We assumed above that the rather unusual possibility of preposition stranding in English required the non-raised category PP/NP_{+wh}. The fact that preposition stranding in French is restricted to rightward extraction, but prohibited in leftward extraction, can be be captured within the same degrees of freedom. We can assume that the raised category $PP^{\uparrow}/NP^{\uparrow}$ again makes PP an island to extractions in general, while the stranding category is PP/NP_{-wh}.[49]

This analysis in turn suggests an explanation for an even more curious asymmetry in English preposition stranding with respect to *rightward* movement, first noted by Ross 1967.

PREPOSITION STRANDING AND RIGHTWARD EXTRACTION: Heavy NP shift (and in fact all extraction of "non-peripheral" arguments other than subjects), depends on the involvement of a suitably constrained version of the backward crossed composition rule, $< Bx$, as in the following non-coordinate examples:

(113)
```
 I will     buy  tomorrow   an engraving by Rembrandt
--------   ----- --------   -------------------------
 S/VP      VP/NP   VP\VP                NP
           -------------<Bx
               VP/NP
  ------------------>B
        S/NP
        --------------------------------->
                     S
```

[49] An essentially equivalent proposal has been independently made by Pollard and Sag forthcoming.

```
(114)  I will       give    to Chapman    an engraving by Rembrandt
      --------   ---------- ------------   -------------------------
        S/VP     (VP/PP)/NP  VP\(VP/PP)               NP
                 ----------------------<Bx
                         VP/NP
        -----------------------<B
              S/NP
              --------------------------------->
                             S
```

In the first of these derivations, an adjunct is composed with a transitive verb. In the second, a type-raised argument PP is composed with a ditransitive verb.[50]

This move correctly excludes heavy shift of subjects (a), and out of subjects (b):

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

(The exclusion of the latter follows from the assumption that type-raising is an obligatory lexical operation upon arguments of verbs, since it follows that it cannot apply to tensed predicates. They therefore cannot be heavy shifted over).

Moreover, since this is the only fully general way in which both leftward and rightward extraction of non-peripheral arguments is permitted in the present theory, we are committed to the view that everything that *can* be heavy shifted over must be an argument or an adjunct.[51]

[50] As in the case of the backward crossed substitution rule 43, we must restrict the backward crossed composition rule to permit only categories of the form $S_x\backslash NP$ (equivalently, VP_x) to unify with the variable Y, in order to exclude the analogous derivation of

(i) a. *a [*curious*]$_{N/N}$ [*by Rembrandt*]$_{N\backslash N}$ [*engraving*]$_N$.

[51] It follows that infinitival VPs in sentences like the following must be arguments of verbs like *consider*.

(i) a. I consider this novel to be poorly written.
b. I consider to be poorly written the novel about the man in the grey flannel suit.

That is to say that a "small clause" analysis of these phenomena at the level of syntax

Heavy NP shift interacts very interestingly with the account of preposition stranding outlined above. Identical compositions to those involved in the Heavy-shifted sentences 113 and 114 are crucial in the derivation of relative clauses such as the following:

(116) a. an engraving which I will buy today and sell tomorrow.
b. an engraving which I will show to him and give to you.

The combinatory grammar correctly prevents both rightward and leftward extraction out of the shifted-over modifier in such examples, whether or not it is subcategorised for, to exclude sentences like the following, which violate Kuno's 1973 Clause Non-final Incomplete Constituent Constraint (CNICC).

(117) a. * A woman who(m) I will $[\text{give}]_{(VP/PP)/NP}$ $[\text{to}]_{PP/NP_{+wh}}$ [an engraving by $\text{Rembrandt}]_{NP}$
b. * I will $[\text{give}]_{(VP/PP)/NP}$ $[\text{to}]_{PP/NP_{+wh}}$ [an engraving by $\text{Rembrandt}]_{NP}$ this very interesting woman.

It is important to notice that this restriction does not depend upon the restriction on the argument of the stranding prepositional category PP/NP_{+wh}. No language could violate Kuno's constraint.

However, it is only because of the restriction to NP_{+wh} that heavy NP shift over a non-subcategorised modifier is blocked, since in this case leftward movement is allowed, an asymmetry noted by Ross 1967

(118) a. Which island did Keats travel to with Chapman?
b. *Keats travelled to with Chapman the beautiful Isle of Capri

However, there is more to this asymmetry than meets the eye, because (as Ross also noticed) the related right node raising is fine:

(119) Keats travelled to, and Chapman returned from, the beautiful Isle of Capri.

To allow this example, we also need the French style stranding category PP/NP_{-wh}. However, this example requires focal stress on the stranded

appears to be incompatible with the present theory. See Pollard and Sag forthcoming, Ch.3 for arguments against the small clause analysis.

prepositions (an observation that also holds of the earlier French examples). Morever, similar stress on the supposedly impossible example 118b, coupled with coordination, seems to improve it considerably:

(120) Keats travelled TO YESTERDAY, and returned FROM this AFTERNOON, the beautiful Isle of Capri.

The unacceptability of 118b therefore seems to be discourse-related. We can capture the data by assuming that English has not only the category PP/NP_{+wh}, but the category PP/NP_{-wh}, the latter bearing the restriction that it must be stressed – see Steedman 1991a for some discussion of the integration of prosodic information into CCG.

As in the case of subject extraction, such asymmetries are rare in comparison with the overwhelming tendency for constraints on rightward and leftward movement to be parallel. It therefore seems appropriate to handle them via minor features on argument terms.

§3 Parasitic Gaps

The Principles of Directional Consistency and Directional Inheritance leave two degrees of freedom to Universal Grammar to specify the directionality of the slashes in each instance of the substitution rule in example 53. As in the case of functional composition, two of these are "forward" instances, with the principal functor on the left, and two are "backward" instances. One of each is slash-crossing, while the other is non-slash-crossing. The theory must predict that all four rule-types are potentially implicated in the languages of the world. The present section examines this consequence of the theory.[52] The present paper revisits many of the examples in Steedman 1987. Where the present version of the theory leaves the earlier analysis unchanged, the reader will be referred there.

¶3.1 Parasitic Gaps in Adjuncts

The backward crossing substitution rule, which is repeated below, was introduced to account for parasitic gaps in adjuncts, such as the following famous

[52] See Steedman 1990 for a similar examination of the four possible rules of functional composition.

relative clause, parallel to the earlier example 44:[53]

(121) Articles which I will $[\text{file}]_{VP/NP}$ $[\text{without reading}]_{(VP\backslash VP)/NP}$

(122) BACKWARD CROSSED SUBSTITUTION (<Sx)
$Y/Z \quad (X\backslash Y)/Z \quad \Rightarrow_{\mathbf{S}} \quad X/Z$
where $Y = S_x\backslash NP$

As usual, VP is an abbreviation for the predicate category, and the binding proceeds as before. The reader may care to satisfy themselves that "stacked" parasitic gaps, as in the following multiply ambiguous examples, are accepted as well:

(123) a. Articles which I will file without reading in order to evaluate.
b. Articles which I will file without reading in order to evaluate before burning

Examples like the following also require the rule 43, and provide further evidence for the mechanism for extracting non-peripheral arguments exemplified in the earlier heavy NP shift examples 113 and 114, and for the existence in the grammar of constituents like *throw in the trash*$_{VP/NP}$, derived by composing the (raised) PP with the verb by the backward crossing composition rule:

(124) a. Articles which I will $[\text{throw in the trash}]_{VP/NP}$ $[\text{without reading}]_{(VP\backslash VP)/NP}$
b. $[\text{I will}]_{S/VP}$ $[\text{reject without reading}]_{VP/NP}$ $[\text{any article longer than sixty pages}]_{NP}$

The theory captures the fact that extraction obeys exactly the same subjacency and ECP-related constraints within the adjunct as it does everywhere else, despite the surprising claim to the contrary by von Stechow 1990, p.458, as the following examples reveal:[54]

[53]The restriction on the variable Y in the rule is discussed in the earlier note 50.

[54]Von Stechow gives no examples to support the claim that the combinatory theory "doesn't restrict the occurrences of a parasitic gap enough" in respect of extraction site, but he appears to be thinking of sentences like these, which are adapted from Chomsky 1986b, p.57-58.

(125) a. Articles which [I will]$_{S/VP}$ [file]$_{VP/NP}$ [without believing that you will read]$_{(VP\backslash VP)/NP}$
b. *Articles which [I will]$_{S/VP}$ [file]$_{VP/NP}$ *[without believing that will please you]$_{(VP\backslash VP)/NP}$
c. Articles which [I will]$_{S/VP}$ [file]$_{VP/NP}$ [without believing will please you]$_{(VP\backslash VP)/NP_{+wh}}$
d. *Articles which [I will]$_{S/VP}$ *[file without reading the name of the person who wrote]$_{(VP\backslash VP)/NP}$

No invocation of empty operators and complex chain formation of the kind proposed in Chomsky 1986b is required to explain this result.

¶3.2 Parasitic Gaps in Complements

There is another kind of parasitic gap in English, in which both of the dependencies are into an argument or arguments. These constructions are somewhat less acceptable than the ones in the previous section. The following is one of the better examples:

(126) a man who(m) I persuaded every friend of to vote for

The grammar of English will allow the example if we include one more of the four types of substitution rules permitted by the Principles of Consistency and Inheritance, namely the forward, non-crossing version 53a, repeated here:

(127) Forward Substitution (>S)
$(X/Y)/Z \quad Y/Z \quad \Rightarrow_S \quad X/Z$

We will assume for the sake of argument and simplicity of presentation that *persuade* can compose into *every friend of.* [55] This implies that the quantifier has an unraised "stranding" category and the latter therefore bears

[55] This assumption is reasonable, given the relative acceptability of *?A man whom I persuaded every friend of to vote Conservative.* However, it is not actually necessary. A further analysis is also possible, in which *every friend of* and *to vote for* bear the category of functions from NP to *raised* NP and VP′ categories. Such categories can combine by one of the generalised instances of the backward crossed substitution rule whose existence is noted in footnote 23 to become a function from NP into functions from object raising verbs to VP. If the verb *persuade* is raised over this category, as the quasi-island violation feel of the sentence suggests is reasonable, then it can compose with *every friend of to vote for* to permit the sentence. Similar analyses are in principle available for the other examples in this section. Their existence is actually required if coordinate sentences like

the category NP/NP_{+wh} after application to the noun and composition with the stranding preposition category. The rule 127 then permits the following derivation:[56]

(128)
```
?a man  who(m)       I will  persuade       every friend of   to vote for
        ------------ ------  -----------    ---------------   -----------
        (N\N)/(S/NP) S/VP    (VP/VP')/NP       NP/NP+wh         VP'/NP
                             ------------------------->B
                                    (VP/VP')/NP+wh
                                    --------------------------------->S
                                              VP/NP+wh
                     --------------------------------->B
                                  S/NP+wh
        ------------------------------>
                     (N\N)
```

Persuade every friend of can therefore combine with *to vote for* by the forward rule, as shown above.[57]

The category S/NP_{+wh} of *persuade every friend of to vote for* inherits from the stranded preposition the restriction to combination with relative pronouns. We therefore correctly predict the same resistance to rightward movement that we saw in sentence 118a, at least in the absence of marked intonation:

(129) * I persuaded every acquaintance of to vote for, my very heavy friends.

A similar analysis to 128 is available for the following sort of example:

(130) A man who(m) I told every friend of that I would support

---

the following (which are related to examples noted by Morrill 1988) are grammatical:
(i) Who(m) did you persuade every friend of to vote for, and every enemy of to ignore?

[56]As usual, the symbols VP, VP', etc. are abbreviations for the predicate category. Again, the analysis differs from the one proposed in Steedman 1987.

[57]We are assuming a non-raised type for the VP' argument, to simplify the presentation. If *to vote for Chapman* is in fact $VP\backslash(VP/VP')$, as is suggested bythe possibility of *I persuaded Keats to vote for Chapman and Gilbert to vote for George*, then the above derivation goes through via the *backward* crossing substitution rule.

However, since the constituent *persuade/tell every friend of* bears the same category as *persuade/tell*, the grammar appears likely to permit all of the following much less acceptable sentences:[58]

(131) a. * (a man) whom I [persuaded]$_{(VP/VP')/NP}$ [to like]$_{VP'/NP}$
b. ?* (a man) whom I [persuaded]$_{(VP/VP')/NP}$ [to believe that I like]$_{VP'/NP}$
c. ? (a man) whom I [told]$_{(VP/S')/NP}$ [that I liked]$_{S'/NP}$

Engdahl (1983, p.24) accepts b and c, and has suggested that the difference between them and a is parallel to the constraints on possible coreference of pronouns and bound anaphors. Others have rejected them all, and have suggested that they should be excluded by a stronger "anti-*c*-command condition", forbidding one gap from *c*-commanding the other in a parasitic construction (cf. Taraldsen 1979 and discussion in Chomsky 1982, p.40-48).

The following examples also suggest that some such constraint applies. The first three are rejected by just about everyone: the fourth more borderline example is permitted by both the proposed constraints, and is tolerated by Engdahl, and by Chomsky 1982, although Chomsky 1981, Sag 1983 and Contreras 1984 reject it:

(132) a. *(a man) whom I will [show]$_{(VP/PP)/NP}$ [to]$_{PP/NP}$
b. *(a man) whom I will [show]$_{(VP/NP)/NP}$ [a picture of]$_{NP/NP}$
c. *(a man) whom I will [talk]$_{(VP/PP)/PP}$ [to]$_{PP/NP}$ [about]$_{PP/NP}$
d. ?(a man) whom I will [show a picture of]$_{(VP/PP)/NP}$ [to]$_{PP/NP}$

It is striking that the stronger of the two conditions is already imposed on the grammar by one of its fundamental principles. Condition C of the Binding Theory, 72, is precisely an anti-*c*-command condition, or equivalently an anti-*F*- (or *o*-) command condition, obeyed by all lexical and derivational categories. Examples 131 and 132a, b, and c are therefore already excluded by the Binding Theory, while 132d is allowed.[59] This claim depends crucially

[58] The annotation "?" on sentences means that there is disagreeement among the authorities concerning the grammaticality of the example. It should not be taken to imply that the present author accepts all such examples.

[59] The suggestion that the anti-*c*-command condition on parasitic gaps should be captured at the level of interpretation goes back at least to Sag 1983. The interpretation of the anti-*c*-command condition as a special case of Condition C was proposed by Chomsky 1982 (and opposed by Chomsky 1986b), and has been more recently revived by Koster 1987, p.356-368.

on the preservation of the obliqueness hierarchy in the interpretation of verbs, and the lexical "wrapping" analysis of the categories for verbs like *persuade* and *show* given in examples 26 and 8. The exclusion of c further depends upon the treatment of prepositions as categorial identity functions in Part I, example 63,b.[60] No additional stipulation is required. In particular, we continue to escape any need to complicate the notion of "government" by a distinction between "antecedent" and "head" varieties, or to draw the associated distinction between A and $\overline{A}$ positions, or to introduce notions like chain composition, or distinctions between varieties of empty category and/or operator, again contrary to the claim of von Stechow 1990.

An example will show how the Binding Theory acts in these cases. We saw earlier that *persuades* has the following category in the lexicon:[61]

(133) persuades $:= ((S : persuade'\ s\ x\ y \backslash NP : y)/(S_{to-inf} : s \backslash NP : PRO\ x)/NP : x$

The category of *to like* is the following, obtained via composition of the proposition and the bare infinitival:

(134) to like $:= (S_{to-inf} : like'\ z\ w) \backslash NP : w)/NP : z$

If the forward substitution rule were to combine them, then we would get the following:

(135) *persuades to like $:= (S : \underline{persuade'\ (like' z\ (PRO\ z))\ z\ w}) \backslash NP : w)/NP : z$

However, the variable z c-commands an instance of itself that is not bound in a PRO-term in the interpretation of the result S, (underlined), The category is therefore excluded under Condition C of the binding theory, 18. Moreover, no combinatory rule whatsoever can overcome this condition, since no combinatory rule can introduce a PRO-term or anything else into an interpretation. (A "composition" rule that could manipulate interpretations in this way would not *be* functional composition).

[60] While 132c is shown with both PPs subcategorised, the same claims hold if one PP is treated as an adjunct. The ill-formedness of c therefore constitutes further evidence for the earlier analysis of prepositions, independent of the theory of anaphor binding.

[61] Again, note the "wrapping" relation between the interpretation of the result and the syntactic category.

Crucially, Condition C is violated twice over by this example and example 131b, since in both cases the variable z is also bound under definition 17 to the variable in a c-commanding PRO-term $PRO\ z$. Example131 c only has a single violation of this comparitively weak condition, since there is no control. This may explain its somewhat lesser unacceptability.

By contrast, after composing into *every friend of*, the constituent *persuades every friend of* might have a category something like the following:

(136) $((S : persuade'\ s\ (every'(friend'z))\ w \backslash NP : w)/(S : s \backslash NP : (PRO\ (every'(friend'z))))/NP_{+wh} : z$

This category can combine with *to like* by forward substitution to yield the following category, because z does not c-command any instance of itself in the interpretation of S (again underlined):

(137) $((S : \underline{persuade'\ (like'\ z\ (PRO\ (every'(friend'z))))\ (every'(friend'z))\ w} \backslash NP : w)/NP_{+wh} : z$

¶3.3 Subjects and Adjunct Parasitic Gaps

Engdahl (1983, ex. 54-56) points out that subjects do not in general support parasitic dependencies, offering examples similar to the following:

(138) a. *(a man) who [painted]$_{(S\backslash NP)/NP}$ [a picture of]$_{NP/NP}$
b. *(a man) who [remembered]$_{(S\backslash NP)/VP_{ing}}$ [talking to]$_{VP_{ing}/NP}$
c. *(a man) who [remembered]$_{(S\backslash NP)/S'}$ [that John talked to]$_{S'/NP}$

Chomsky 1986b, p.55 ascribes the badness of such examples to the anti-c-command condition. However, as Koster 1987, p.346 points out, these and many other examples involving subjects are considerably worse than this comparatively weak condition would lead one to expect. I noted in Steedman 1987 that in CCG all three examples are excluded by the Principles of Consistency and Inheritance, at the level of universal grammar, without further stipulation or reference to c-/F-command, or binding theory. There is no possible combinatory rule, whether corresponding to **S** or any other combinator, that will permit 138a, b and c. All such putative rules would violate the Principle of Inheritance, by equating $/NP$ with $\backslash NP$. Thus the examples are excluded for essentially the same reason as **that-t* violations, because of incompatibility in direction of government, still without the stipulation of an autonomous Empty Category Principle.

As Engdahl pointed out, *extracted* subjects *can* take part in parasitic constructions. She gives the following examples:

(139) a. (the Caesar) whom Brutus will imply t was no good while ostensibly praising t_p
b. (a man) who everyone who meets t_p admires t
c. (a man) who you said John's criticism of t_p would make us think t was stupid

Unlike the earlier cases, examples like 139a *are* allowed by the present theory, because the subject-extracted predicate *imply was no good* can be built by directly combining the special subject-extracting category $(VP/NP_{+wh})/(S\backslash NP)$ of the verb (cf. example 103 of the previous section) with the complement predicate category $S\backslash NP$. The resulting category VP/NP_{+wh} and the adjunct function *while ostensibly praising*$_{(VP\backslash VP)/NP}$ (built by the forward composition rule) are of the appropriate form and linear order for the familiar backward instance of the substitution rule to apply, thus:[62]

(140)

```
    who(m)   Brutus will       imply      was no good while ostensibly praising
------------ ----------- ---------------- ----------- -------------------------
(N\N)/(S/NP)    S/VP     (VP/NPwh)/(S\NP)    S\NP           (VP\VP)/NP
                         ------------------------>
                                  VP/NPwh
                                  -----------------------------------<Sx
                                                VP/NPwh
                ---------------------------------------->B
                                 S/NPwh
  ------------------------------------->
                  (N\N)
```

[62]Again, this analysis differs slightly from earlier proposals, because of the new analysis of subject extraction. Besides the feature-value $+wh$, it will be recalled from the discussion of subject extraction 103 that the NP argument of the constituent *imply was no good* also bears the feature *sing*(ular), distinguished for the number of the extracted subject. The (non-subject) argument of the adjunct *while ostensibly praising* is not marked on either feature. It is therefore free to combine, and the substitution rule ensures that the argument Z of the function that it produces bears both values. (The Principle of Inheritance requires for all combinatory rules that the features on an argument in the result must bear the union of the feature-value pairs on the corresponding argument(s) in its input functions). This detail is omitted in the present derivation.

This analysis provides additional evidence that the extracting subject argument of bare complement verbs like *imply* is a *rightward* NP argument of some kind. If it were leftward, like a true subject, no analysis at all would be permitted under the present theory, because it would require a combinatory rule violating the Principle of Directional Inheritance 51. The example also forces the assumption of the present analysis that the special NP_{+wh} argument is the *second* argument of the verb. If it were the first argument, the analysis would require type-raising of the predicate *was no good*. However, the existence of the Subject Condition shows that predicates cannot type-raise, *contra* Steedman 1987.

Finally, it follows from the analysis that *rightward* movement out of this parasitic construction will be just as impossible as rightward movement of a single subject:

(141) *Brutus implied was no good while ostensibly praising the man in the Brooks Brothers shirt.

¶3.4 Subjects and Complement Parasitic Gaps

Engdahl's second variety of subject parasitics, example 139b, repeated here, is an example of a parasitic inside a subject, and is mediated by the forward substitution rule.

(142) a man who(m) everyone who meets admires.

In this case, it is the first of the two extraction sites that is normally inaccessible:

(143) a. (a man) whom everyone who likes politicians admires.
b. *(a man) whom everyone who likes admires politicians.

Example (b) is a violation of the Subject Condition of Chomsky (1970), which as we saw in example 93 is captured in the present theory. With both extractions, Engdahl's example is accepted as follows:[63]

[63] In the earlier paper I proposed an alternative analysis in terms of the backward rule and the unraised category NP/NP. Note that the analysis assumes that subjacency is

```
(144) (a man)   who(m)          everyone           who        meets    admires
              ------------ ----------------- ------------ --------- ---------
              (N\N)/(S/NP) ((S/(S\NP))/(N\N) (N\N)/(S\NP) (S\NP)/NP (S\NP)/NP
                           ------------------------------>B
                                 (S/(S\NP))/(S\NP)
                                 ----------------------------------->B
                                           (S/(S\NP))/NP
                                           ----------------------------------->S
                                                          S/NP
              ------------------------------------------------->
                                 (N\N)
```

It is predicted by the above analysis that rightward movement out of such constructions should also be allowed. Provided that the conditions for heaviness are met, this rather surprising prediction seems to be born out, for (a) below is much better than the subject condition-violating (b):

(145) a. Everyone who meets admires, and everyone who sees remembers, that fascinating woman who chairs the Parking Space Committee
b. *Everyone who meets admires sincerity, and everyone who sees remembers the Alamo, that fascinating woman who chairs the Parking Space Committee

Because extracted embedded subjects correspond to rightward arguments the theory predicts that embedded subjects should be able to take part in parasitic dependencies into complements, and they can, when the embedded subject gap is in the complement:

(146) a man who(m) [everyone who meets]$_{(S/(S\backslash NP))/NP}$ [thinks is no good]$_{(S\backslash NP)/NP_{+wh}}$

Engdahl's third type of subject parasitic gap, 139c, *(a man) who you said John's criticism of would make us think was stupid,* which has a subject gap as the second member of the pair, is closely related to the last example, for the finite verb phrase *would make us think was stupid* bears the category $(S\backslash NP)/NP_{+wh}$. However, if an extracted embedded subject is the *first* gap,

overcome by type-raising *everyone* over its adjunct. The sentence does indeed have the feel of a subjacency violation. As we saw in connection with examples like 110, *?a book that I will walk without reading*, the possibility of such exceptions is sensitive to content in mysterious ways, so it is not suprising that some examples related to the above, such as *?a boy who the brother of admires*, are even worse.

then it is correctly predicted not to take part in parasitic constructions of this kind:

(147) * a man who(m) you must [know]$_{(VP/NP_{+wh})/(S\backslash NP)}$ [thinks Harry likes]$_{(S\backslash NP)/NP}$

The reason is once again that the subject-extracting category does not match any of the substitution rules.[64]

Finally, as in the case of the backward substitution rule, but again unlike ordinary parasitic gaps, embedded subject parasitics cannot rightward move, because of the $+wh$ feature that the lexicon imposes upon extractable subjects:[65]

(148) *[Everyone who meets thinks is no good,]$_{S/NP_{+wh}}$
and [everyone who listens to realises is utterly mad,]$_{S/NP_{+wh}}$
[that dreadful little man who chairs the Parking Space Committee]$_{NP_{-wh}}$

¶3.5 PARASITIC DEPENDENCIES IN DUTCH

Because Dutch is a predominantly verb-final language, it posesses a number of constructions discussed by Bennis 1986 and Koster 1987 which exhibit a pattern of parasitic gapping that is virtually the mirror image of the pattern in English. The dependencies in question are mediated by the rule $>$ **Sx**, 53b, and thus constitute a further confirmation that the degrees of freedom that the theory allows are indeed reflected in the languages of the world.[66]

Dutch has two kinds of relative pronoun. The relative pronoun *die/dat* is essentially like English *who(m)/that.* The relative pronoun *waar* is much more restricted, and can only be used to relativise the complement *er* of a very limited class of postpositions such as *op, voor*, and *over* which can take such complements. Many of these postpositions also act as normal prepositions as well, as illustrated in the following examples:

[64]Again, this example is much too bad for its exclusion to be merely attributed to the violation of Condition C.

[65]Certain further possibilities for subject parasitic gaps discussed by Chomsky 1982, 1986b, whose analysis is unchanged under the present theory, are discussed in Steedman 1987, p.432-433.

[66]The following data are taken from van Riemsdijk 1978 and Bennis (1986, esp. p.44)

(149) a. Jan heeft op Marie/*Marie op gerekend
John has on Mary counted
"John counted on Mary."
b. Jan heeft *op er/er op gerekend
John has there on counted
"John counted thereon."

Dutch, like most languages but unlike English, does not allow preposition stranding. However, it does allow this small class of *post*positions to strand. We therefore see the following pattern:

(150) a. *de arts die hij op *t* heeft gerekend
*the doctor that he on has counted
"the doctor that he counted on"
b. de arts waar hij *t* op heeft gerekend
the doctor where he on has counted
"the doctor whereon he counted"

Both kinds of relative in Dutch can give rise to parasitic gaps. The following example from Bennis is the Dutch mirror image to the familiar *file without reading* examples in English (traces are included in the transliteration, purely as an aid to the non-Dutch reader):

(151) Welke boeken heb je [zonder te lezen]$_{(VP/VP)\backslash NP}$ [weggezet]$_{VP\backslash NP}$
Which books have you without *t* reading *t* away-put?
"Which books did you put away without reading?"

That is, the parasitic gap lies in a (preverbal) tenseless adjunct. The rule involved is the following mirror image of the English <Sx rule:

(152) DUTCH FORWARD CROSSED SUBSTITUTION
$(X/Y)\backslash Z \quad Y\backslash Z \quad \Rightarrow_{\mathsf{S}} \quad X\backslash Z \quad (>\mathsf{Sx})$
where $Y = S_x\backslash NP$

The same rule allows parasitic *waar* clauses also. Let us distinguish the argument *er* of the first type, the postpositions, as an NP_{er}. Then the following example from Bennis is allowed, again with the parasitic gap in the adjunct:

(153) Waar heb je [na twee jaar over nagedacht te hebben]$_{(VP/VP)\backslash NP_{er}}$ [een oplossing voor gevonden?]$_{VP\backslash NP_{er}}$
What have you after two years having thought t_{er} about a solution t_{er} to found?
"What have you found a solution to after two years having thought about?"

A surprising but correct prediction immediately follows from this analysis. We have noted at a number of points that in general in English, parasitic dependencies may occur in rightward movement constructions, as well as leftward movement. This possibility arises in English because it is a VO language, and because composites like *file without reading* have the category VP/NP, like any transitive verb, and like any transitive verb, can combine to the right with an object, as in example 44. However, Dutch is an OV language. Composites like *zonder te lezen weggezetten* have the category $VP\backslash NP$, like all Dutch transitive verbs. It follows that this construction cannot take part in rightward-moved constructions. In fact, as the theory correctly predicts, rightward movement is extremely rare in Dutch (cf. Neijt 1979). Instead we correctly predict that these parasitic composites, like any Dutch transitive verb, should happily combine with an object immediately to their *left*, as in the following example.

(154) Jan heeft deze boeken [zonder te lezen]$_{(VP/VP)\backslash NP}$ [weggezet]$_{VP\backslash NP}$
Jan has these books without t reading away-put
"Jan put away without reading these (very heavy) books"

The existence of these sentences, which appear to involve a parasitic gap without a real gap, is noted by Bennis 1986, pp.54-70, who deals with them at length within a GB framework, and who also notes the resemblance between this form of "raising" and English Right Node Raising. The examples present some difficulties for his approach, requiring the assumption that the sentence has undergone an otherwise unmotivated process of "scrambling" from an underlying subject-adjunct-object-verb order (p.59). This has the effect of widening the definition of antecedent government and $\overline{A}$-position, in order to bring such sentences under the generalisation that a parasitic gap must be governed by an antecedent in $\overline{A}$-position (p.63), whilst continuing to exclude the following:

(155) *Jan heeft [zonder te lezen]$_{(VP/VP)\backslash NP}$ [deze boeken weggezet]$_{VP}$
Jan has without t reading these books away-put
"*Jan put away these books without reading"

This sentence is automatically disallowed under the present theory, without further stipulation, for the same reason as its English gloss. This result is once again achieved without invoking an $A/\overline{A}$ distinction over the positions of arguments, a distinction which plays no part in CCG.

The same rule also potentially permits parasitic gaps in complements. In the case of the latter kind of relative clause, Bennis claims that the following example is well-formed.[67]

(156) Dit is het artikel waar
ik [over zei]$_{(VP/S)\backslash NP_{er}}$ [dat Hendrik een reactie op moest schrijven]$_{S\backslash NP_{er}}$
This is the article which I t_{er} about said that Harry a reaction t_{er} to should write.
"*This is the article which I said of that Harry should write a reply to."

According to Bennis, examples like the following involving the other type of pronoun, are ungrammatical:

(157) * Dit is de man die ik [t vertelde]$_{(VP/S)\backslash NP}$ [dat Hendrik t zou bezoeken]$_{S\backslash NP}$
? This is the man who I told that Harry would visit.

The unacceptability of this sentence, which would otherwise be allowed by $> S_x$, is, as Bennis points out, an instance of the anti-*c*-command condition. It is therefore predicted by the present version of Condition C of the binding

[67]All indications of acceptability for the Dutch examples are from Bennis, while the English judgements remain my own. As in the case of the related English examples like *?Who did you tell that you had seen*, it is actually quite hard to get native informants to give the unequivocal judgements that these annotations suggest. However, all that matters for the present argument is that the first example be judged better than the second.

theory.[68]

§4 Conclusion

The above account necessarily remains incomplete. Many questions concerning the relation of binding and *c*-command (or equivalently, *F*- or *o*-command) have not been addressed. They include the relation of quantifier scope to argument structure, and the related phenomenon of weak "crossover". Many relevant constructions have been passed over or mentioned only in passing, including the examples that motivate "relativised minimality" (Rizzi 1990 – but cf. Szabolsci 1992b). Many pressing questions about the nature of the categorial lexicon and the associated lexical rules await further study, including the question of why the lexical specification of control relations is limited to "subjacent" complements. Perhaps the most glaring omission is the lack of any explanation whatsoever for the existence of Condition C itself. (In this, however, we are not alone).

Nevertheless, the results above suggest that all of the phenomena that have previously been described in terms of *Wh*-movement or the equivalent can be captured by the combinatory alternative, without *Wh*-traces or empty operators, and in particular without attendant conditions on traces or slash-features, such as the ECP, and without specific conditions on parasitic gaps. Relativisation can be captured using the same mechanism which at the level of interpretation associates arguments *in situ* with thematic roles, projected from the lexicon via the combinatory rules. In this respect, CCG resembles

[68]The Lost Combinator: Since English seems to include the forward substitution rule >S, example 127, as well as the backward crossing rule <Sx, example 43, one might expect Dutch, given its tendency to be the mirror image of English, to include the remaining one of the four substitution rules, <S. It appears not to. Dutch is not a perfect mirror of English, because its verbs govern sentential complements to the right. Thus as we have seen, sentences like 128 which in English need a separate rule, in Dutch are mediated by the same rule. To find a use for <S we shall undoubtedly have to turn to a "real" SOV language, like Korean, in which sentential complements are governed to the left, and relative markers are on the right of the clause. The prediction is that sooner or later we shall find one that will permit parasitic gaps on the pattern of 128, and that they also will be subject to the effects of Binding Condition C. Unfortunately, it is hard to verify this fascinating prediction. Many SOV languages (including Korean) permit null anaphora, raising the possibility that other mechanisms might mediate long range dependencies in the relevant examples.

the earliest forms of Transformational Grammar, inasfar as the combinatory rules of composition and substitution can be viewed as "double-based" or "generalised" transformations (Chomsky 1957), whose resemblance to TAG tree adjunction has been noted (cf. Chomsky 1992). The restriction to composition and substitution amounts to the claim that there only exist these two classes of syntactic transformation/adjunction in UG, and that the "proper analysis" or condition for applying these rules can be entirely defined in terms of the functional type of constituents, without regard to structure or derivation. Unlike such related theories as GPSG and HPSG, CCG achieves this result without in any way distinguishing extracting arguments of verbs from their subcategorised complements.

The core of Government-Binding theory and the other generalisations of Categorial Grammars that have been mentioned above can be seen in present terms as defining the notions "possible lexical category", and "possible argument structure". This core remains essentially unaffected by the present proposal.[69] The distinctive contribution of CCG to this consensus may lie in the claim that the phenomena that have been attributed to *Wh*-movement are closely related to phenomena that have previously been attributed to a number of devices determining Phonetic Form, sometimes referred to as "stylistic" rules, including among other things rules for "deletion under coordination". All of these can be simply captured in terms of a single combinatory system.

By adopting the combinatory alternative, CCG of course engenders a very much freer notion of surface constituency than the one implicit in the alternative theories. However, each non-standard constituent is paired with the correct interpretation. It follows that, to the extent that the alternative theories cover the same range of constructions, assigning correct interpretations, they must be carrying out the same operations in their semantics. (This is clearest in the case of relative clause dependencies: the semantics

[69]One interesting possibility here is that certain limitations on the lexicon arise from the involvement in lexical interpretations of a similar range of combinators to those active in syntax, obviating the use of variables entirely. This possibility is implicit in much work by Dowty and Jacobson (following Bach), and is discussed by Szabolcsi 1989. This work offers the enticement that it might be possible to dispense with the present structural view of interpretations and binding relations, and replace it with something more directly related to model theory.

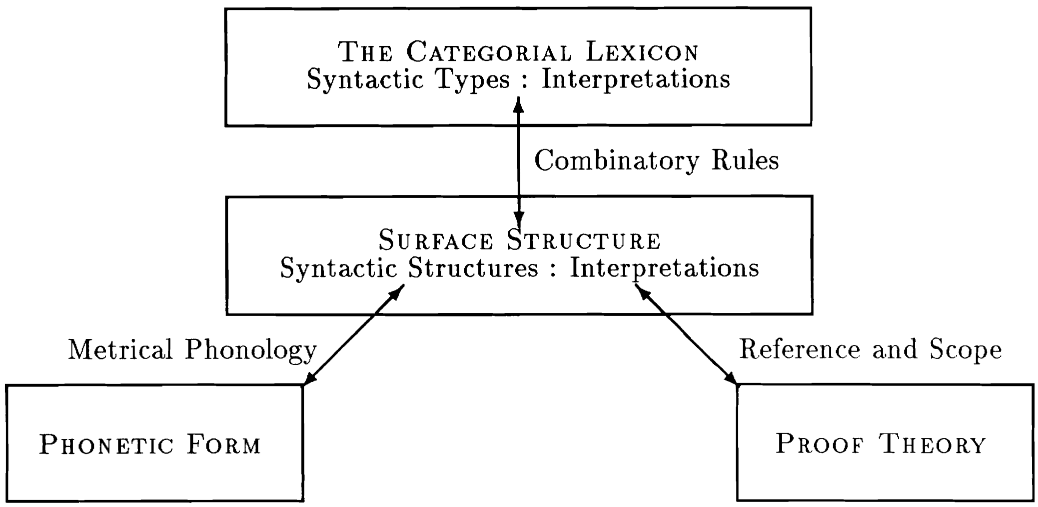

FIGURE 1.

corresponding to slash-feature passing in HPSG, the effect of (some kinds of) adjunction in TAG, and the apparatus of "Functional Uncertainty" in LFG path specifications (Kaplan and Zaenen, 1987), all bear a close resemblance to functional composition). The distinguishing characteristic of CCG is simply that it exhibits a more direct "rule-to-rule" relation between semantics and surface syntax, and thereby captures a wider variety of constructions, particularly in the domain of coordination.

The theory of grammar that is implicit in CCG can therefore be understood, as in Figure 1, in terms of the traditional "T" diagram. Under this view, both the Lexicon and the level of Surface Structure concern categories which bear a syntactic and a semantic significance, and which are either functions or arguments. The semantic interpretations or "argument structures" associated with basic categories like S, over which functions map, preserve fairly traditional relations of dominance and command. These structures conform to the obliqueness hierarchy on grammatical relations. However, the order in which categories combine syntactically with the corresponding arguments need not conform to the obliqueness hierarchy.[70]

[70]Although fuller discussion must be deferred, it is worth noting that the assumption of this freedom seems to be a forced move under the account of VSO languages proposed by

At the level of Surface Structure, the categories take part in derivations. The resultant syntactic structures do not in general conform to traditional notions of command and dominance. Instead, they capture directly the notion of constituency relevant to relativisation, coordination, and phrasal intonation. It is only in the interpretations of the categories that take part in such derivations that relations such as *c*-command are manifest. In particular, it is the interpretation of S that embodies the canonical proposition and represents its function-argument relations. At this level, the translations of quantifiers and noun properties are *in situ*, modulo the obliqueness hierarchy.

The responsibility of the combinatory rules is to "project" both components of the lexical categories onto the corresponding components of categories at the level of Surface Structure.[71] The derivations or constituent structures that they yield are considerably more diverse than traditional surface structures or S-structures. They provide the input to purely local phonological processes, such as Liason and the Rhythm Rule (Selkirk 1984), which directly map Surface Structures onto Phonetic Form proper.[72]

By contrast, the interpretation or argument structure, which is the exclusive domain of the Binding Theory, provides the input to such systems as reference, discourse binding of pronouns, and the resolution of quantifier scope. It is presumably in this process that the effects associated with "weak crossover" and "subjacency" show up. While we may find it convenient to think about this process in terms of a further structural level of Logical Form, such a representation is not in principle necessary, for the reasons discussed by Montague 1970, and in fact this level is eschewed in other versions of Categorial Grammar.[73] For this reason, the figure is non-committal on the

Dowty 1988 (cf. Steedman 1990, p.233, esp. n.30), given the fact that binding phenomena in those languages also obey the obliqueness hierarchy.

[71]See Szabolcsi 1992a for a discussion of the relation between this notion of projection and the "Projection Principle" of GB.

[72]Although the question is not pursued in detail here, I show in Steedman 1991a that the "intonational structures" postulated by Selkirk are isomorphic to surface structures under the present theory, and that the discourse "information structures" that she associates with intonational constituents are identical to the present interpretations or argument structures associated with surface constituents in the generalised sense implicated by CCG. There is therefore no need to postulate any path from PF to LF or Proof Theory other than the one shown, via surface structure.

[73]Cf. Dowty 1992, Jacobson 1990, Szabolcsi 1989, and others. The fact that the inter-

question of whether such an autonomous representation is required as an input to the Proof Theory.

Within this framework, all responsibility for long-range dependencies, whether associated with "*Wh*-movement", "scrambling", "reanalysis", "stylistic rules", "intonation structure", or "deletion under coordination", falls to the apparatus that directly projects thematic roles from the categorial lexicon to Surface Structure. This projection is mediated by language-specific instances of rules drawn from just two combinatory families, Composition and Substitution. A third combinatory rule family, Type-raising, appears to properly belong within the lexicon, rather than in syntax. The set of universally available instances of each combinator is determined by the grammatical principles of Adjacency, Consistency and Inheritance.

---

pretation or argument structure postulated here is the exclusive domain of the Binding Theory means that the present theory is entirely modular. It follows that nothing in the account of *Wh*-constructions hinges on the fact that we have chosen, in contrast to these other categorial theories, to discuss binding in structural terms. It also follows that any of the theories that embody a theory of argument structure of the kind advocated here, including LFG, HPSG, TAG, and certain versions of GB, are immediately compatible with the present theory of coordination and unbounded dependency. The pairing both at the level of the lexicon and at that of Surface Structure of syntactic structures that map directly to phonetics and interpretations the map directly to semantics is also akin to the realisation of SDs as pairs (π, λ) in Chomsky 1992, p.62.

References

Abelson, Harold, and Gerald Sussman: 1985, *Structure and Interpretation of Computer Programs*, MIT Press, Cambridge MA.

Abney, Steven: 1987, *The English Noun Phrase in its Sentential Aspect*, unpublished PhD dissertation, MIT, Cambridge MA.

Ades, Anthony E. and Mark J. Steedman: 1982, 'On the Order of Words', *Linguistics and Philosophy*, **4**, 517-558.

Ajdukiewicz, Kazimierz: 1935, 'Die syntaktische Konnexität', *Studia Philosophica*, **1**, 1-27. English translation in Storrs McCall (ed), *Polish Logic 1920-1939*, Oxford University Press, 207-231.

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

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, and Barbara Partee: 1980, 'Anaphora and Semantic Structure', in *Papers from the Parasession on Pronouns and Anaphora, Tenth Meeting of the Chicago Linguistics Society*, CLS, Chicago.

Bar-Hillel, Yehoshua: 1953, 'A Quasi-arithmetical Notation for Syntactic Description', *Language*, **29**, 47-58.

Barry, Guy: 1991, *Derivation and Structure in Categorial Grammar*, unpublished Ph.D dissertation, University of Edinburgh.

Bennis, Hans: 1986, *Gaps and Dummies*, Foris, Dordrecht.

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

Bouma, Gosse: 1987, 'A Unification-based Analysis of Unbounded Dependencies in Categorial Grammar', in J. Groenendijk, M. Stokhof, and F. Veltman, (eds.), *Proceedings of the 6th Amsterdam Colloquium*, Foris, Dordrecht.

Brame, Michael: 1976, *Conjectures and Refutations in Syntax and Semantics*, North Holland, Amsterdam.

Bresnan, Joan (ed.): 1982, *The mental Representation of Grammatical Relations*, MIT Press, Cambridge, MA.

Carpenter, Robert: 1989, *Phrase Meaning and Categorial Grammar*, unpublished Ph.D dissertation, University of Edinburgh.

Chierchia, Gennaro: 1985, 'Formal Semantics and the Grammar of Predication', *Linguistic Inquiry,* **16**, 417-443.

Chierchia, Gennaro: 1988, 'Aspects of a Categorial Theory of Binding', in Oehrle et al, 1988, 153-98.

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

Chomsky, Noam: 1970, 'Deep Structure, Surface Structure, and Semantic Interpretation', in D. Steinberg and L. Jakobovits, *Semantics*, CUP, Cambridge, 1971, 183-216.

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

Chomsky, Noam: 1982, *Concepts and Consequences of the Theory of Government and Binding*, MIT Press, Cambridge MA.

Chomsky, Noam: 1986b, *Barriers*, MIT Press, Cambridge MA.

Chomsky, Noam: 1992, 'A Minimalist Program for Linguistic Theory', MIT Working Papers in Linguistics.

Contreras, Heles: 1984, 'A Note on Parasitic Gaps', *Linguistic Inquiry,* **15**, 198-200.

Cooper, Robin: 1982, 'Binding in Wholewheat Syntax', in P. Jacobson and G.K. Pullum (eds.), *The nature of Syntactic Representation*, Reidel, Dordrecht.

Cresswell, Max: 1973, *Logics and Languages*, Methuen.

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

Dowty, David: 1982, 'Grammatical Relations and Montague Grammar', in Pauline Jacobson and Geoffrey K. Pullum (eds.), *The Nature of Syntactic Representation*, Reidel, Dordrecht. 79-130.

Dowty, David: 1988, 'Type-raising, Functional Composition, and Nonconstituent Coordination', paper to the Conference on Categorial Grammar, Tucson, AR, June 1985, in Richard T. Oehrle, E.. Bach and D. Wheeler, (eds), *Categorial Grammars and Natural Language Structures*, Reidel, Dordrecht, 153-198.

Dowty, David: 1992, '"Variable-free" Syntax, Variable-binding Syntax, the Natural Deduction Lambek Calculus, and the Crossover Constraint', in *Proceedings of the 1992 West Coast Conference on Formal Linguistics*, SLA Stanford CA, to appear.

Engdahl, Elisabet: 1983, 'Parasitic Gaps', *Linguistics and Philosophy* **6**, 5-34.

Flynn. Michael: 1983, 'A Categorial Theory of Structure Building', in Gerald Gazdar, E. Klein and G.K. Pullum, (eds.), *Order, Concord, and Constituency*, Foris, Dordrecht. 139-174.

Frank, Robert: 1991, 'Parasitic Gaps and Locality Conditions', *Proceedings of the 27th meeting of the Chicago Linguistics Society*, CLS, Chicago, 167-181.

Gazdar, Gerald: 1981, 'Unbounded Dependencies and Coordinate Structure', *Linguistic Inquiry*, **12**, 155-184.

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

Grimshaw, Jane: 1990, *Argument Structure*, MIT Press, Boston MA.

Hellan, Lars: 1988, *Anaphora in Norwegian and the Theory of Grammar*, Foris, Dordrecht.

Henderson, James: 1990, *Structure Unification Grammar*, MSE dissertation, Univ. of Pennsylvania, Tech. Report MS-CIS-90-94, University of Pennsylvania, Philadelphia.

Hepple, Mark: 1990, *The Grammar and Processing of Order and Dependency: a Categorial Aproach*, Ph.D dissertation, University of Edinburgh. Memo, Centre for Cognitive Science, University of Edinburgh.

Higgins, F.R. *The Pseudocleft Construction in English*, unpublished doctoral dissertation, MIT, Cambridge MA.

Hoeksema, Jack: 1989, *A Categorial Theory of Reanalysis Phenomena*, ms., University of Pennsylvania.

Hukari, Thomas, and Robert Levine: 1987, 'Parasitic Gaps, Slash Termination, and the *C*-command Condition', *Natural Language and Linguistic Theory* **5**, 197-222

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

Jackendoff, Ray: 1972, *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge.

Jacobson, Pauline: 1987, 'Phrase Structure, Grammatical Relations and Discontinuous Constituents', in Geoffrey Huck & Almerido Ojeda, (eds.) *Syntax and Semantics* **20**: *Discontinuous Constituency*, Academic Press, New York, 27-69.

Jacobson, Pauline: 1990, 'Raising as Function Composition', *Linguistics & Philosophy,* **13**, 423-476.

Jacobson, Pauline: 1991, 'Bach-Peters sentences in a Variable-free semantics', *Proceedings of the Eighth Amsterdam Colloquium*, to appear.

Joshi, Aravind, K. Vijay-Shanker, and David Weir: 1991, 'The Convergence of Mildly Context-sensitive Formalisms', in Peter Sells, Stuart Shieber and Tom Wasow, (eds.), *Processing of Linguistic Structure*, MIT Press, Cambridge MA. 31-81.

Joshi, Aravind, and Yves Schabes: 1992, 'Tree adjoining grammars and lexicalized grammars', in M. Nivat and M. Podelski (eds.), *Definability and Recognizability of Sets of Trees*, Elsevier, Princeton.

Kaplan, David: 1975, 'How to Russell a Frege-Church', *Journal of Philosophy,* **72**, 716-729.

Kaplan, Ron, John Maxwell, and Annie Zaenen: 1987, 'Functional Uncertainty', *CSLI Monthly*, CSLI, Stanford University.

Karttunen, Lauri: 1989, 'Radical Lexicalism', in M.R. Baltin and A.S. Kroch (eds.) *Alternative Conceptions of Phrase Structure*, Chicago University Press, Chicago.

Kayne, Richard: 1984, *Connectedness and Binary Branching*, Foris, Dordrecht.

Keenan, Ed and Bernard Comrie: 1977, 'Noun Phrase Accessibility and Universal Grammar', *Linguistic Inquiry,* **8**, 63-99.

Keenan, Ed: 1988, 'On Semantics and the Binding Theory', in John Hawkins (ed.) *Explaining Language Universals*, Basil Blackwell, Oxford, 105-144.

Keenan, Ed and Leonard Faltz: 1978, *Boolean Semantics for Natural Language*, Reidel, Dordrecht.

Koster, Jan: 1987. *Domains and Dynasties*, Foris, Dordrecht.

Kuno, Susumu: 1973, 'Constraints on Internal Clauses and Sentential Subjects', *Linguistic Inquiry,* **4**, 363-386.

Kuno, Susumo: 1987, *Functional Syntax*, Chicago University Press, Chicago.

Montague, Richard: 1970, 'Universal Grammar', *Theoria,* **36**, 373-398.

Moortgat, Michael: 1988b, *Categorial Investigations*, Ph.D dissertation, Rijksuniversiteit, Gröningen, publ. Foris, Dordrecht, 1989.

Morrill, Glyn: 1988, *Extraction and Coordination in Phrase Structure Grammar and Categorial Grammar*, unpublished Ph.D dissertation, University of Edinburgh.

Neijt, Anneke: 1979, *Gapping*, Foris, Dordrecht.

Oehrle, Richard T: 1974, 'Remarks on "the Painting of Rembrandt"', in *Proceedings of the 10th Annual Meeting of the Chicago Linguistic Society*, CLS, Chicago.

Oehrle, Richard T: 1988, 'Multidimensional Compositional Functions as a Basis for Grammatical Analysis', in Richard T. Oehrle, E. Bach and D. Wheeler, (eds), *Categorial Grammars and Natural Language Structures*, Reidel, Dordrecht. 349-390.

Oehrle, Richard T. and Zhang Shi: 1990, 'Lambek Calculus and Preposing of Embedded Subjects', *Proceedings of the 26th meeting of the Chicago Linguistics Society*, CLS, Chicago, XXX-XXX.

Partee, Barbara, and Emmon Bach: 1981, Incomplete reference.

Pereira, Fernando, and Stuart Shieber: 1987, *Prolog and Natural Language Analysis*, CSLI/Chicago University Press, Chicago.

Perlmutter, David and Paul Postal: 1977, 'Towards a Universal Characterisation of Passivization', in David Perlmutter (ed.) *Studies in Relational Grammar, I*, University of Chicago Press, Chicago, 1983.

Pesetsky, David: 1982, *Paths and Categories*, unpublished Ph.D dissertation, MIT.

Pollard, Carl and Ivan Sag: 1987, *Information-based Syntax and Semantics, vol. 1*, CSLI/Chicago University Press, Chicago.

Pollard, Carl and Ivan Sag: 1991, An Integrated Theory of Control, *Language* **67**, 63-113.

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

Pollard, Carl and Ivan Sag: 1993, *Information-based Syntax and Semantics, vol. 2*, CSLI/Chicago University Press, Chicago.

Reape, Michael: 1991, *Incomplete Reference*, unpublished PhD thesis, University of Edinburgh.

Reinhart, Tanya and Eric Reuland: 1991, 'Anaphors and Logophors: an Argument Structure Perspective', in Jan Koster & Eric Reuland (eds.), *Long Distance Anaphora*, Cambridge University Press, 283-322.

van Riemsdijk, Henk: 1978, *A Case Study in the Theory of Markedness*, Foris Dordrecht.

Rizzi, Luigi: 1982, *Issues in Italian Syntax*, Foris, Dordrecht.

Rizzi, Luigi: 1990, *Relativized Minimality*, MIT Press, Cambridge MA.

Ross, John Robert: 1967, *Constraints on Variables in Syntax*, doctoral dissertation, MIT, published as *Infinite Syntax!*. Ablex, Norton, NJ. 1986.

Safir, Ken: 1986, 'Relative Clauses in a Theory of Binding and Levels', *Linguistic Inquiry,* **17**, 663-689.

Sag, Ivan: 1983, 'On Parasitic Gaps', *Linguistics and Philosophy* **6**, 35-45.

Schönfinkel, Moses: 1924, 'On the Building Blocks of Mathematical Logic', in Jean van Heijnoort (ed.), *From Frege to Gödel*, Harvard University Press, Cambridge MA. 355-366.

Schubert, Lenhart and Francis Pelletier: 1982, 'From English to Logic: Context-free Computation of "Conventional" Logical Translations', *Computational Linguistics,* **10**, 165-176.

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

Shaumyan, Sebastian: 1977, *Applicational Grammar as a Semantic Theory of Natural Language*, Edinburgh University Press.

Shieber, Stuart: 1986, *An Introduction to Unification-based Approaches to Grammar*, CSLI/Chicago University Press, Chicago.

Smullyan, Raymond: 1985, *To Mock a Mockingbird*, Knopf, New York.

von Stechow, Arnim: 1979, 'Deutsche Worstellung und Montague Grammatik', in J.M. Meisel and M.D. Pam (eds.), *Linear Order and Generative Theory*, John Benjamins, Amsterdam.

von Stechow, Arnim: 1990, 'Categorial Grammar and Linguistic Theory', *Studies in Language,* **14**, 433-478.

Steedman, Mark J: 1985, 'Dependency and Coordination in the Grammar of Dutch and English', *Language*, **61**, 523-568.

Steedman, Mark J. 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: 1991a, 'Structure and Intonation', *Language,* **67**, 262-296.

Steedman, Mark: 1991b, 'Type-raising and Directionality on Combinatory Grammar', in *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*, Berkeley, July 1991, 71-78.

Steele, Susan: 1990, *Agreement and Anti-agreement*, Reidel, Dordrecht.

Szabolcsi, Anna: 1983, 'ECP in Categorial Grammar', ms, Max Planck Institute, Nijmegen.

Szabolcsi, Anna: 1989, 'Bound Variables in Syntax: Are there any?' in R. Bartsch, J. van Benthem, and P. van Emde Boas (eds.), *Semantics and Contextual Expression*, 295-318, Foris, Dordrecht.

Szabolcsi, Anna: 1992a, 'On Combinatory Grammar and Projection from the Lexicon' in Ivan Sag & Anna Szabolcsi, (eds.), *Lexical Matters*, CSLI, Stanford CA. 241-268.

Szabolcsi, Anna: 1992b, 'Weak Islands, Individuals, and Scope', in Barker and Dowty (eds.), *Proceedings of the Conference on Semantics and Linguistic Theory* **2**, OSU WP 40, pp. 407-436.

Szabolcsi, Anna, and Franz Zwarts: 1990. Semantic Properties of Composed Functions and the Distribution of *Wh*-phrases. Proceedings of the Seventh Amsterdam Colloquium, ed. by Stokhof and Torenvliet. 529-555. Amsterdam: ITLI.

Szabolcsi, Anna, and Franz Zwarts: 1993, 'Weak Islands and an Algebraic Semantics for Scope Taking', *Natural Language Semantics*, (to appear).

Taraldsen, Tarald: 1979, 'The Theoretical Interpretation of a Class of Marked Extractions', in A. Belletti, L. Brandi, and L. Rizzi (eds.), *Theory of Markedness in Generative Grammar*, Scuole Noarmale Superiore di Pisa, 475-516.

Uszkoreit, Hans: 1986, 'Categorial Unification Grammars', *Proceedings of the 11th International Conference on Computational Linguistics*, Bonn, August, 1986, 187-194.

Webber, Bonnie Lynn: 1978, *A Formal Approach to Discourse Anaphora*, Ph.D dissertation, Harvard, publ. Garland 1979.

Wittenburg, Kent B. 1986, *Natural Language Parsing with Combinatory Categorial Grammar in a Graph-Unification Based Formalism.* D.Phil Dissertation, U. of Texas at Austin.

Wood, M. 1988, *A Categorial Syntax for Coordinate Constructions*, Ph. D dissertation, University College, London.

Zeevat, Henk, Ewan Klein, and Jo Calder: 1987, 'An Introduction to Unification Categorial Grammar', in N. Haddock et al. (eds.), *Edinburgh Working Papers in Cognitive Science, 1: Categorial Grammar, Unification Grammar, and Parsing*, 195-222.

Zubizarreta, Maria Luisa: 1987, *Levels of Representation*, Foris, Dordrecht.

Zwarts, Franz: 1986, *Categoriale Grammatica en Algebraïsche Semantik*, Ph.D dissertation, Rijkuniversiteit, Gröningen.