CATEGORIAL GRAMMAR

MARK STEEDMAN
University of Edinburgh

ABSTRACT

Categorial Grammar comprises a family of lexicalized theories of grammar characterized by very tight coupling of syntactic derivation and semantic composition, having their origin in the work of Frege. Some versions of CG have extremely restricted expressive power, corresponding to the smallest known natural family of formal languages that properly includes the context-free. Nevertheless, they are also strongly adequate to the capture of a wide range of cross-linguistically attested non-context-free constructions. For these reasons, categorial grammars have been quite widely applied, not only to linguistic analysis of challenging phenomena such as coordination and unbounded dependency, but to computational linguistics and psycholinguistic modeling.

1. INTRODUCTION. Categorial Grammar (CG) is a "strictly" lexicalized theory of natural language grammar, in which the linear order of constituents and their interpretation in the sentences of a language are entirely defined by the lexical entries for the words that compose them, while a language-independent universal set of rules projects the lexicon onto the strings and corresponding meanings of the language. Many of the key features of Categorial Grammar have over the years been assimilated by other theoretical syntactic frameworks. In particular, there are recent signs of convergence from the Minimalist Program within the transformational generative tradition (Chomsky 1995; Berwick and Epstein 1995; Cormack and Smith 2005; Boeckx 2008:250).

Categorial grammars are widely used in various slightly different forms discussed below by linguists interested in the relation between semantics and syntactic derivation. Among them are computational linguists who for reasons of efficiency in practical applications wish to keep that coupling as simple and direct as possible. Categorial grammars have been applied to the syntactic and semantic analysis of a wide variety of constructions, including those involving unbounded dependencies, in a wide variety of languages (e.g. Moortgat 1988b; Steele 1990; Whitelock 1991; Morrill and Solias 1993; Hoffman 1995; Nishida 1996; Kang 1995, 2002; Bozşahin 1998, 2002; Komagata 1999; Baldridge 1998, 2002; Trechsel 2000; Cha and Lee 2000; Park and Cho 2000; Çakıcı 2005, 2009; Ruangrajitpakorn et al. 2009; Bittner 2011, 2014; Kubota 2010; Lee and Tonhauser 2010; Bekki 2010; Tse and Curran 2010).

Categorial grammar is generally regarded as having its origin in Frege's remarkable 1879 *Begriffsschrift*, which proposed and formalized the language that we now know as first-order predicate logic (FOPL) as a Leibnizian calculus in terms of the combination of functions and arguments, thereby laying the foundations of all modern logics and programming languages, and opening up the possibility that natural language grammar could be thought of in the same way. This possibility was investigated in its syntactic and computational aspect for small fragments of natural language by Ajdukiewicz (1935) (who provided the basis for the modern notations), Bar-Hillel (1953) and Bar-Hillel et al. (1964) (who gave categorial grammar its name), and Lambek (1958) (who initiated the type-logical interpretation of CG).

It was soon recognized that these original categorial grammars were context-free (Lyons 1968), and therefore unlikely to be adequately expressive for natural languages (Chomsky 1957), because of the existence of unbounded or otherwise "long range" syntactic and semantic dependencies between elements such as those italicized in the following examples:[1]

(1) a. These are *the songs* they say that the Syrens *sang*.
b. The Syrens *sang* and say that they wrote *these songs*.
c. *Some Syren* said that she had written *each song*. $(\exists\forall/\forall\exists)$
d. *Every Syren* thinks that the sailors heard *her*.

Frege's challenge was taken up in categorial terms by Geach (1970) (initiating the combinatory generalization of CG), and Montague (1970b) (initiating direct compositionality, both discussed below) In particular, Montague (1973) influentially developed the first substantial categorial fragment of English combining syntactic analysis (using a version of Ajdukiewicz' notation) with semantic composition in the tradition of Frege (using Church's λ-calculus as a "glue language" to formalize the compositional process).

In the latter paper, Montague used a non-monotonic operation expressed in terms of structural change to accomodate long range dependencies involved in quantifier scope alternation and pronoun-binding, illustrated in (1c,d). However, in 1970b he had laid out a more ambitious program, according to which the relation between syntax and semantics in all natural languages would be strictly homomorphic, like the syntax and semantics in the model theory for a mathematical, logical, or programming language, in the spirit of Frege's original program.

For example, the standard model theory for the language of first-order predicate logic (FOPL) has a small context-free set of syntactic rules, recursively defining the structure of negated, conjunctive, quantified, etc. clauses in terms of operators $\neg$, $\wedge$, $\forall x$ etc. and their arguments. The semantic component then consists of a set of rules paired one-to-one with the syntactic rules, compositionally defining truth of an expression of that syntactic type solely in terms of truth of the arguments of the operator in question (see Robinson 1974).

Two observations are in order when seeking to generalize such Fregean systems as FOPL to human language. One is that the mechanism whereby an operator such as $\forall x$ "binds" a variable x in a term of the form $\forall x[P]$ is not usually considered part of the syntax of the logic. If it is treated syntactically, as has on occasion been proposed for programming languages (Aho 1968), then the syntax is in general no longer context-free.[2]

The second obervation is that the syntactic structures of FOPL can be thought of in two distinct ways. One is as the *syntax of the logic* itself, and the other is as a *derivational structure* describing a process by which an interpretations has been constructed. The most obvious context-free derivational structures are isomorphic to the logical syntax, such as those which apply its rules directly to the analysis of the string, either bottom-up or top-down. However, even for a context-free grammar, derivation structure may be determined by a different "covering" syntax, such as a "normal form"

[1] These constructions in English were shown by Gazdar (1981) to be coverable with only context-free resources in Generalized Phrase Structure Grammar (GPSG), whose "slash" notation for capturing such dependencies is derived from but not equivalent to the categorial notation developed below. However, Huybregts (1984) and Shieber (1985) proved Chomsky's widely accepted conjecture that in general such dependencies require greater than CF expressive power.

[2] This observation might be relevant to the analysis of "bound variable" pronouns like that in (1d).

grammar. (Such covering grammars are sometimes used for compiling programming languages, for reasons such as memory efficiency.) Such covering derivations are irrelevant to interpretation, and do not count as a representational level of the language itself. In considering different notions of structure involved in theories of natural language grammar, it is important to be clear whether one is talking about logical syntax or derivational structure.

Recent work in categorial grammar has built on the Fregeo-Montagovian foundation in two distinct directions, neither of which is entirely true to its origins. One group of researchers has made its main priority capturing the semantics of diverse constructions in natural languages using standard logics, often replacing Montague's structurally non-monotone "quantifying in" operation by more obviously compositional rules or memory storage devices. Its members have tended to either remain agnostic as to the syntactic operations involved or assume some linguistically-endorsed syntactic theory such as transformational grammar or GPSG (e.g. Partee 1975; Cooper 1983; Szabolcsi 1997; Jacobson 1999; Heim and Kratzer 1998), sometimes using extended notions of scope within otherwise standard logics (e.g. Kamp and Reyle 1993; Groenendijk and Stokhof 1991; Ciardelli and Roelofsen 2011), or tolerating a certain increase in complexity in the form of otherwise syntactically or semantically unmotivated surface-compositional syntactic operators or type-changing rules on the syntactic side (e.g. Bach 1979; Dowty 1982; Hoeksema and Janda 1988; Jacobson 1992; Hendriks 1993; Barker 2002) and the Lambek tradition (e.g. Lambek 1958, 2001; van Benthem 1983, 1986; Moortgat 1988a; Oehrle 1988; Morrill 1994; Carpenter 1995; Bernardi 2002; Casadio 2001; Moot 2002; Grefenstette et al. 2011).

Other post-Montagovian approaches have sought to reduce syntactic complexity, at the expense of expelling some apparently semantic phenomena from the logical language entirely, particularly quantifier scope alternation and pronominal binding, relegating them to offline specification of scopally underspecified logical forms (e.g. Kempson and Cormack 1981; Reyle 1993; Poesio 1995; Koller and Thater 2006; Pollard 1984), or extragrammatical discourse reference (e.g. Webber 1978; Bosch 1983)

One reason for this diversity and divergence within the broad church of categorial grammar is that the long-range and/or unbounded dependencies exemplified in (1) above, which provide the central challenge for any theory of grammar and for the Fregean approach in particular, fall into three distinct groups. Relativization, topicalization, and right node-raising are clearly unbounded and clearly syntactic, being subject to strong island constraints, such as the "fixed subject constraint", as in (2a).

(2) a. #This is *the Syren* they wonder whether *sang* a song.
b. *Some Syren* claimed that *each song* was the best. $(\exists\forall/\#\forall\exists)$
c. *Every Syren* claimed that *some song* was the best. $(\forall\exists/\exists\forall)$
d. *Every Syren* thinks that *her* song is the best.

On the other hand, the binding of pronouns and other nominals as dependents of quantifiers is equally clearly completely insensitive to islands, as in (2d), while quantifier scope inversion is a mixed phenomenon, with the universals *every and each* apparently unable to invert scope out of embedded subject positions, as in (2b), while the existentials can do so (2c).

The linguistic literature in general is conflicted on the precise details of what species of dependency and scope is allowed where. However, there is general agreement that while syntactic long-range dependencies are mostly nested, and the occasions when crossing dependencies are allowed are very narrowly specified syntactically, intrasentential binding of pronouns and dependent existentials is essentially free within the scope of the operator. For example, crossing and nesting binding dependencies in the following seem equally good:

(3) Every sailor$_i$ knows that every Syren$_j$ thinks she$_j$/he$_i$ saw him$_i$ /her$_j$.

It follows that those researchers whose primary concern is with pronoun binding in semantics tend to define their Fregean theory of grammar in terms of different sets of combinatory operators from those researchers whose primary concern is syntactic dependency. Thus, not all categorial theories discussed below are commensurable.

2. PURE CATEGORIAL GRAMMARS. In all varieties of Categorial Grammar, elements like verbs are associated with a syntactic "category" which identifies them as Fregean *functions*, and specifies the type and directionality of their arguments and the type of their result. We here use the "result leftmost" notation in which a rightward-combining functor over a domain β into a range α are written α/β, while the corresponding leftward-combining functor is written $\alpha\backslash\beta$.[3]

α and β may themselves be function categories. For example, a transitive verb is a function from (object) NPs into predicates—that is, into functions from (subject) NPs into S:

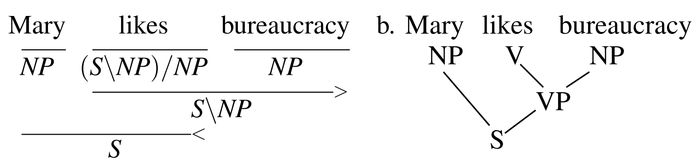
(4) likes := $(S\backslash NP)/NP$

All varieties of categorial grammar also include the following rules for combining forward- and backward–looking functions with their arguments:

(5) *Forward Application:* ($>$)
$X/Y \quad Y \quad \Rightarrow \quad X$

(6) *Backward Application:* ($<$)
$Y \quad X\backslash Y \quad \Rightarrow \quad X$

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

(7) a.


It is important to note that such tree-structures are simply a representation of the process of derivation. They do not necessarily constitute a level of representation in the formal grammar.

CG categories can be regarded as encoding the semantic type of their translation, and this translation can be made explicit in the following expanded notation, which associates a logical form with the entire syntactic category, via the colon operator, which is assumed to have lower precedence than the categorial slash operators. (Agreement

[3] There is an alternative "result on top" notation due to Lambek (1958), according to which the latter category is written $\beta\backslash\alpha$. Lambek's notation has advantages of readability in the context-free case, because all application is adjacent cancellation. However, this advantage does not hold for trans–context-free theories which include non-Lambek operators such as crossed composition. For such grammars, and for any analysis in which the semantics has to be kept track of, the Lambek notation is confusing, because it does not assign a consistent left-right position to the result α vs. the argument β.

features are also included in the syntactic category, represented as subscripts, much as in Bach 1983. The feature 3*s* is "underspecified" for gender and can combine with the more specified *3sm* by a standard unification mechanism that we will pass over here—cf. Shieber 1986.)[4]

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

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

(9) *Forward Application:* $(>)$
$$X/Y : f \quad Y : a \quad \Rightarrow \quad X : fa$$

(10) *Backward Application:* $(<)$
$$Y : a \quad X\backslash Y : f \quad \Rightarrow \quad X : fa$$

They yield derivations like the following:

(11)
$$\begin{array}{ccc}
\text{Mary} & \text{likes} & \text{bureaucracy} \\
\hline NP_{3sm} : mary' & (S\backslash NP_{3s})/NP : likes' & NP : bureaucracy' \\
 & \multicolumn{2}{c}{\overline{S\backslash NP_{3s} : likes'bureaucracy'}\;>} \\
\multicolumn{3}{c}{\overline{S : likes'bureaucracy'mary}\;<}
\end{array}$$

The derivation yields an S with a compositional interpretation, equivalent under a convention of left associativity to $(likes'bureaucracy')mary'$.

Coordination can be included in CG via the following category, allowing constituents of like type to conjoin to yield a single constituent of the same type:[5]

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

Since X is a variable over any type, and all three Xs must unify with the *same* type, it allows derivations like the following.

(13)
$$\begin{array}{ccccc}
\text{I} & \text{detest} & \text{and} & \text{oppose} & \text{bureaucracy} \\
\hline NP & (S\backslash NP)/NP & (X\backslash X)/X & (S\backslash NP)/NP & NP \\
 & & \multicolumn{2}{c}{\overline{((S\backslash NP)/NP)\backslash((S\backslash NP)/NP)}\;>} & \\
 & \multicolumn{3}{c}{\overline{(S\backslash NP)/NP}\;<} & \\
 & \multicolumn{4}{c}{\overline{S\backslash NP}\;>} \\
\multicolumn{5}{c}{\overline{S}\;<}
\end{array}$$

3. COMBINATORY CATEGORIAL GRAMMAR. In order to allow coordination of contiguous strings that do not constitute constituents, CCG allows certain further operations on functions and arguments related to Curry's combinators **B**, **T**, and **S**, (Curry and Feys 1958).

3.1. FUNCTION COMPOSITION **B**. Functions may not only apply to arguments, but also *compose* with other functions, under the following rule, first proposed in modern

[4] Another notation, more in the spirit of Prolog-style unification-based formalisms like Lexical Functional Grammar (LFG) and Head-driven Phrase Structure Grammar (HPSG) associates a unifiable logical form with each primitive category, so that the same transitive verb might appear as follows (cf. Uszkoreit 1986; Karttunen 1989; Bouma and van Noord 1994; Zeevat 1988):

(i) likes := $(S : likes'y\,x\backslash NP_{3s} : x)/NP : y$

The advantage is that the predicate-argument structure is built directly by unification with X and Y in rules like (5) and (6), which need no further modification to apply (cf. Pereira and Shieber 1987). Otherwise, the choice is largely a matter of notational convenience.

[5] The semantics of this category, or rather category schema, is given by Partee and Rooth (1983), and is omitted here as a distraction. We will come to certain restrictions on the combinatory potential of this category below.

terms in Ades and Steedman 1982 but with antecedents in Geach (1970) and as a theorem of Lambek 1958:

(14) *Forward Composition:* ($> B$)
$X/Y \quad Y/Z \quad \Rightarrow \quad X/Z$

The most important single property of combinatory rules like this is that their semantics is completely determined under the following principle:

(15) *The Principle of Combinatory Transparency:* The semantic interpretation of the category resulting from a combinatory rule is uniquely determined by the interpretation of the slash in a category as a mapping between two sets.

In the above case, the category X/Y is a mapping of Y into X and the category Y/Z is that of a mapping from Z into Y. Since the two occurrences of Y identify the *same* set, the result category X/Z is that mapping from Z to X which constitutes the composition of the input functions. It follows that the only semantics that we are allowed to assign, when the rule is written in full, is as follows:

(16) *Forward Composition:* ($>\mathbf{B}$)
$X/Y:f \quad Y/Z:g \quad \Rightarrow \quad X/Z:\lambda x.f(gx)$

No other interpretation is allowed.[6]

The operation of this rightward composition rule in derivations is indicated by an underline indexed $>\mathbf{B}$ (because Curry called his composition combinator $\mathbf{B}$). Its effect can be seen in the derivation of sentences like *I detest, and will oppose, bureaucracy*, which crucially involves the composition of two verbs to yield a composite of the same category as a transitive verb (the rest of the derivation is given in the simpler notation). It is important to observe that composition also yields an appropriate interpretation for the composite verb *will oppose*, as $\lambda x.\lambda y.will'(oppose'\ x)\ y$, a category which if applied to an object *bureaucracy* and a subject *I* yields the proposition $will'(oppose'\ bureaucracy')\ me'$. The coordination will therefore yield an appropriate semantic interpretation.[7]

(17)
$$
\begin{array}{cccccc}
\text{I} & \text{detest} & \text{and} & \text{will} & \text{oppose} & \text{bureaucracy} \\
\overline{NP} & \overline{(S\backslash NP)/NP} & \overline{(X\backslash X)/X} & \overline{(S\backslash NP)/VP : will'} & \overline{VP/NP : oppose'} & \overline{NP} \\
 & & & \multicolumn{2}{c}{\underline{\hphantom{(S\backslash NP)/NP : \lambda x.\lambda y.will'(oppose'x)y}}_{>\mathbf{B}}} & \\
 & & & \multicolumn{2}{c}{(S\backslash NP)/NP : \lambda x.\lambda y.will'(oppose'x)y} & \\
 & \multicolumn{4}{c}{\underline{\hphantom{(S\backslash NP)/NP : \lambda x.\lambda y.will'(oppose'x)y \quad (S\backslash NP)/NP}}_{<\Phi>}} & \\
 & \multicolumn{4}{c}{(S\backslash NP)/NP} & \\
 & & & \multicolumn{3}{c}{\underline{\hphantom{(S\backslash NP)/NP : \lambda x.\lambda y.will'(oppose'x)y \quad NP}}_{>}} \\
 & & & \multicolumn{3}{c}{S\backslash NP} \\
\multicolumn{5}{c}{\underline{\hphantom{(S\backslash NP)/NP : \lambda x.\lambda y.will'(oppose'x)y \quad S\backslash NP}}_{<}} & \\
\multicolumn{5}{c}{S} & \\
\end{array}
$$

3.2. TYPE-RAISING $\mathbf{T}$. Combinatory grammars also include type-raising rules, originally proposed in Steedman 1985 but with antecedents in generalized quantifier theory, which turn arguments into functions over functions-over-such-arguments. These rules allow arguments to compose, and thereby take part in coordinations like *I detest, and Mary likes, bureaucracy*. For example, the following rule allows the conjuncts to form as below (again, the remainder of the derivation is given in the briefer notation):

(18) *Subject Type-raising:* ($>$T)
$NP : a \quad \Rightarrow \quad S/(S\backslash NP) : \lambda f.f\ a$

[6] This principle would follow automatically if we were using the alternative unification-based notation discussed in note 4 and the composition rule as as it is given in 14.

[7] The analysis compresses two applications into a single coordination step labeled $<\phi>$, and begs some syntactic and semantic questions about the interpretation of modals.

(19)
$$\begin{array}{cccccc}
\text{I} & \text{detest} & \text{and} & \text{Mary} & \text{likes} & \text{bureaucracy}\\
\overline{NP} & \overline{(S\backslash NP)/NP} & \overline{(X\backslash X)/X} & \overline{NP} & \overline{(S\backslash NP)/NP} & \overline{NP}\\
 & & & : mary' & : \lambda x.\lambda y.likes'xy & \\
\overline{}{>}\mathbf{T} & & & \overline{}{>}\mathbf{T} & & \\
S/(S\backslash NP) & & & S/(S\backslash NP) & & \\
 & & & : \lambda f.f\ mary' & & \\
\multicolumn{2}{c}{\overline{\hspace{8em}}{>}\mathbf{B}} & & \multicolumn{2}{c}{\overline{\hspace{10em}}{>}\mathbf{B}} & \\
\multicolumn{2}{c}{S/NP} & & \multicolumn{2}{c}{S/NP} & \\
 & & & \multicolumn{2}{c}{: \lambda x.likes'x\ mary'} & \\
\multicolumn{5}{c}{\overline{\hspace{28em}}{<}\Phi{>}} & \\
\multicolumn{5}{c}{S/NP} & \\
\multicolumn{6}{r}{\overline{\hspace{22em}}{>}}\\
\multicolumn{6}{c}{S}
\end{array}$$

Rule 18 has an "order-preserving" property. That is, it turns the NP into a *rightward* looking function over *leftward* function, and therefore preserves the linear order of subjects and predicates specified in the lexicon for the language.

Like composition, type-raising rules are required by the Principle of Combinatory Transparency (15) to be transparent to semantics. This fact ensures that the raised subject NP has an appropriate interpretation, and can compose with the verb to produce a function that can either coordinate with another nonstandard constituent of the same type or reduce with an object *bureaucracy* to yield *likes' bureaucracy' mary'* via a nonstandard left-branching alternative derivation to (11), delivering the same logical form.

The latter alternative derivational structures are sometimes misleadingly referred to as "spuriously" ambiguous (Karttunen 1989), and deprecated as exacerbating the search problem for the parser. However, any theory of grammar that covers the same range of coordination phenomena will engender the same degree of nondeterminism in derivation. We return to this question in section 7.

While other solutions to the problem of getting subjects to combine with the transitive verb can readily be imagined, the inclusion of order-preserving type-raising is essential to the account of coordination afforded by CCG, because it allows *sequences of arguments* to compose. We defer discussion of this question until section 3.6, where we will see that type-raising should be constrained as an essentially lexical operation, identifiable with the traditional notion of *case*, whether morphologically marked, as in Latin and Japanese, or positionally marked or "structural", as in English.

3.3. SUBSTITUTION **S**. For reasons that we will come to directly, the following rule of a type closely related to composition, first proposed by Szabolcsi (1983, 1992) under the name "connection", and discussed at length in Steedman 1987, is required for the analysis of the parasitic gap construction, illustrated in (21):

(20) *Backward Crossed Substitution:* ($<\mathbf{S}$)
$$Y/Z : g \quad (X\backslash Y)/Z : f \quad \Rightarrow \quad X/Y : \lambda x.(fx)(gx)$$

(21) I will [[burn]$_{VP/NP}$ [without reading]$_{(VP\backslash VP)/NP}$]$_{VP/NP}$ [any report longer than 100 pages]$_{NP}$

3.4. THE SPACE OF POSSIBLE COMBINATORY RULES IN CCG. Rule (20) is of interest for exploiting all and only the degrees of freedom that are available under the following universal syntactic projection principles (the term "Principal Functor" refers to the input functor whose range is the same as the range of the output—in the notation used above, the functor whose range is X):

(22) *The Principle of Directional Consistency*
The inputs to a combinatory rule must be directionally consistent with the principal functor—if the latter applies to the result of the subordinate fuctor to the left, it must be rightmost of the inputs, and vice versa.

(23) *The Principle of Directional Inheritance*
The argument(s) in a function that is the output of a combinatory rule must be directionally consistent with the corresponding argument(s) in the input functors—if an argument of the output bears a slash of a given directionality, all occurrences of that argument in the inputs must bear a slash of the *same* directionality.

These rules are pefectly illustrated by Szabolcsi's rule (20): although the the inputs bear different "crossing" directionality (which is allowed), the principal functor $(X\backslash Y)/Z$ is looking for the result Y of the subordinate functor Y/Z to its left, so it is rightmost (*consistency*), and the argument Z of the output X/Z bears a rightward slash in both inputs (*inheritance*).

The lexical type-raising rules are limited by these principles to the order preserving cases (30) and (31): raised categories which change word order, such as topics and the relative pronouns discussed in the next section, have to change the result type.

3.5. EXTRACTION. Since complement-taking verbs like *think*, VP/S, can in turn compose with fragments like *Mary likes*, S/NP, we correctly predict that right-node raising is unbounded, as in a, below, and also provide the basis for an analyis of the similarly unbounded character of leftward extraction, as in b:[8]

(24) a. [[I detest]$_{S/NP}$ and [you think Mary likes]$_{S/NP}$]$_{S/NP}$ [bureaucracy.]$_{NP}$
b. The bureaucracy$_{N}$ [[that]$_{(N\backslash N)/(S/NP)}$ [you think Mary likes]$_{S/NP}$]$_{N\backslash N}$

(25) a. I will [[sieze]$_{VP/NP}$ and [burn without reading]$_{VP/NP}$]$_{VP/NP}$ [any report longer than 100 pages.]$_{NP}$
b. The reports$_{N}$ [[that]$_{(N\backslash N)/(S/NP)}$ [I will burn without reading]$_{S/NP}$]$_{N\backslash N}$

3.6. COORDINATION. This apparatus has been applied to a wide variety of coordination phenomena, including "argument-cluster coordination", "backward gapping" and "verb-raising" constructions in a variety of languages by the authors listed in the introduction. The first of these is relevant to the present discussion, and is illustrated by the following analysis, from Dowty (1988, cf. Steedman 1985):[9]

(26)

$$\begin{array}{ccccc}
\text{introduce} & \text{Bill} & \text{to Sue} & \text{and} & \text{Harry} \quad \text{to George} \\
\hline (VP/PP)/NP & \overline{(VP/PP)\backslash((VP/PP)/NP)}^{<\mathbf{T}} & \overline{VP\backslash(VP/PP)}^{<\mathbf{T}} & \overline{(X\backslash X)/X} & \overline{VP\backslash((VP/PP)/NP)}^{<\mathbf{B}} \\
 & \multicolumn{2}{c}{\overline{VP\backslash((VP/PP)/NP)}^{<\mathbf{B}}} & & \\
 & \multicolumn{4}{c}{\overline{VP\backslash((VP/PP)/NP)}^{<\Phi>}} \\
\multicolumn{4}{c}{\overline{VP}^{<}} &
\end{array}$$

The important feature of this analysis is that it uses "backward" rules of type-raising $<\mathbf{T}$ and composition $<\mathbf{B}$ that are the exact mirror-image of the two "forward" versions introduced as examples 14 and 18, which similarly guarantee that the semantics of nonstandard constituents like *Bill to Sue* is such as to reduce appropriately with a

[8] See the earlier papers and Steedman (1996, 2000a) for details, including ECP effects and other extraction asymmetries, and the involvement of similar apparently non-constituent fragments in intonational phrasing.

[9] In more recent work, Dowty has disowned this analysis, on the ground that it makes an "intrinsic" use of logical form to account for binding phenomena. This issue is discussed further in Dowty 1997 and Steedman 1996, and more briefly below.

ditransitive verb like *give*. It is in fact a prediction of the theory that such a construction can exist in English, and its inclusion in the grammar requires no additional mechanism whatsoever.

The earlier papers show that no *other* non-constituent coordinations of dative-accusative NP sequences are allowed in any language with the English verb categories, given the assumptions of CCG. Thus the following are ruled out in principle, rather than by stipulation:

(27) a. *Bill to Sue and introduce Harry to George
b. *Introduce to Sue Bill and to George Harry

In English the phenomenon shows up in all constructions that can be assumed to involve multiple arguments of the same functor:[10]

(28) a. I introduced Bob to Carol, and Ted to Alice.
b. I saw Thelma yesterday, and Louise the day before.
c. Will Gilbert arrive and George leave?
d. I persuaded Warren to take a bath and Dexter to wash his socks.
e. I promised Mutt to go to the movies and Jeff to go to the play.
f. I told Shem I lived in Edinburgh and Shaun I lived in Philadelphia.
g. I bet Sammy sixpence he would win and Rosie a dollar she would lose.
h. I like Ike and you, Adlai.

A number of related well-known cross-linguistic generalizations first noted by Ross (1970) concerning the dependency of so-called "gapping" upon lexical word-order are also captured (see Dowty (1988) and Steedman (1985, 1990, 2000b)). The pattern is that in languages whose basic clause constituent order subject-verb-object (SVO), the verb or verb group that goes missing is the one in the right conjunct, and not the one in the left conjunct. The same asymmetry holds for VSO languages like Irish. However, SOV languages like Japanese show the opposite asymmetry: the missing verb is in the *left* conjunct.[11] The pattern can be summarized as follows for the three dominant constituent orders (asterisks indicate the excluded cases):[12]

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

This observation can be generalized to individual constructions within a language: just about any construction in which an element apparently goes missing preserves canonical word order in an analogous fashion: (26) above is an example of this generalization holding of a verb-initial construction in English.

Phenomena like the above immediately suggest that all complements of verbs bear type-raised categories. However, we do not want anything *else* to type-raise. In particular, we do not want raised categories to raise again, or we risk infinite regress in our rules. One way to deal with this problem is to explicitly restrict the two type-raising rules to the relevant arguments of verbs, as follows, a restriction that is a natural expression of the resemblance of type-raising to some generalized form of (nominative, accusative, etc., morphological or structural) grammatical *case*—cf. Steedman (1985, 1990).

[10] This assumption precludes a small clause analysis of the basic constructions.

[11] A number of apparent exceptions to Ross's generalization have been noted in the literature and are discussed in Steedman 2000b. Ross's constraint is there stated in terms of overall order properties of languages and constructions rather than any notion of "underlying" word order.

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

(30) *Forward Type-raising:* $(> \mathbf{T})$
$X : a \quad \Rightarrow \quad \mathrm{T}/(\mathrm{T}\backslash X) : \lambda f.fa$

(31) *Backward Type-raising:* $(< \mathbf{T})$
$X : a \quad \Rightarrow \quad \mathrm{T}\backslash(\mathrm{T}/X) : \lambda f.fa$

The other solution is to simply expand the lexicon by incorporating of the raised categories that these rules define, so that categories like NP have raised categories, and all functions into such categories, like determiners, have the category of functions into raised categories.

These two tactics are essentially equivalent, because in some cases we need both raised and unraised categories for complements. (The argument depends upon the observation that any category that is not a barrier to extraction must bear an unraised category, and any argument that can take part in argument-cluster coordination must be raised—cf. Dowty 2003). The correct solution from a linguistic point of view, inasfar as it captures the fact that some languages appear to lack certain unraised categories (notably PP and S'), is probably the lexical solution. However the restricted rule-based solution makes derivations easier to read and allows them to take up less space.

Since categories like NP can be raised over a number of different functor categories, such as predicate, transitive verb, ditransitive verb etc, and since the resulting raised categories $S\backslash(S/NP)$, $(S\backslash NP)\backslash((S\backslash NP)/PP)$, etc. of NPs, PPs, etc are quite hard to read, it is sometimes convenient to abbreviate the raised categories as a schema written $NP^{\uparrow}$, $PP^{\uparrow}$, etc.[13]

3.7. GENERALIZING COMBINATORY RULES. A number of generalized forms of the combinatory rules are allowed. We have already noted the need for crossed directionality in rules. Thus for composition we have the following rule allowed under principles (22) and (23)

(32) *Backward Crossed Composition:* $(>\mathbf{B_x})$
$Y/Z : g \quad X\backslash Y : f \quad \Rightarrow \quad X/Z : \lambda x.f(gx)$

This rule allows "heavy NP-shifted" examples like the following:

(33) I will $[[\text{go}]_{VP/PP}$ $[\text{next Sunday}]_{VP\backslash VP}]_{VP/PP}$ $[\text{to London}]_{PP}$

We also generalize all composition rules to higher valency in the subordinate input function, to a small bound equal to the highest valency specified in the lexicon, say ≤ 4, thus:[14]

(34) *Backward Crossed Second-order Composition:* $(<\mathbf{B^2_x})$
$(Y/Z)/W : g \quad X\backslash Y : f \quad \Rightarrow \quad (X/Z)/W : \lambda w \lambda z.f(gwz)$

Such rules allow examples like the following, related to (33)

(35) I shall $[[\text{present}]_{(VP/PP)/NP}$ $[\text{next Sunday}]_{VP\backslash VP}]_{VP\backslash VP}$ $[\text{a prize}]_{PP}$ $[\text{to each winner}]_{PP}$

The inclusion of crossed second order composition in the theory of grammar allows unbounded crossing dependency of the kind investigated in Dutch and Zurich German by Huybregts and Shieber, and is the specific source of greater than CF power in CCG.

The inclusion of crossed combinatory rules means that lexical items must be restricted to avoid overgeneration in fixed word-order languages like English. For exam-

[13] In computational implementations English type-raised categories are usually schematized in this way, because its word order is sufficiently rigid to allow the statistical parsing model to resolve the ambiguity locally.

[14] If there is no bound on n the expressive power of the system jumps to that of full Indexed Grammar (Srinivas 1997).

ple, while adverbial adjuncts can combine with verbs by crossed combination as above, adnominal adjuncts cannot:

(36) *An $[\text{old}]_{N/N}$ $[\text{with a limp}]_{N\backslash N}$ man

Accordingly, we follow Hepple (1990); Morrill (1994); Moortgat (1997), and more specifically Baldridge and Kruijff (2003) in lexically distinguishing the slash-type of English adnominals and their specifiers as only allowed to combine by non-crossed rules, writing them as follows:

(37) with := $N\backslash_{\diamond}N$

Similarly, the coordination category (12) must be rewritten as follows, where the * modality limits it to only combining via the application rules:

(38) and := $(X\backslash_{\star}X)/_{\star}X$

This is managed using a simple type-lattice of features on slashes in categories and rules due to Baldridge, whose details we pass over here (see Steedman and Baldridge 2011).

3.8. FREE WORD-ORDER. When dealing with the phenomenon of free order, it may be convenient for some languages to represent some group of arguments of a head as combining in any order to yield the same semantic result. Following Hoffman (1995) and Baldridge (2002), if three arguments A, B, C are completely free-order with respect to a head yielding S, we might write its category as in (a) below. If it requires all three to the left (right) in any order, we might write the category as in (b (c)). If the head requires A to the left and the other arguments in any order to the right, we might write the category as in (d):

(39) a. $S\{|A, |B, |C\}$
b. $S\{\backslash A, \backslash B, \backslash C\}$
c. $S\{/A, /B, /C\}$
d. $S\{\backslash A, /B, /C\}$

Braces $\{\ldots\}$ enclose multisets of arguments that can be found in any order in the specified direction. The question then arises of how multisets behave under combinatory rules such as composition.

Baldridge points out that, to preserve the near context-free expressive power of CCG, it is crucial to interpret the multiset notation as merely an abbreviation for exhaustive listing of all the ordered categories that would be required to support the specified orders. (For example, $S\{\backslash A, \backslash B, \backslash C\}$ abbreviates the set $\{((S\backslash A)\backslash B)\backslash C,\ ((S\backslash A)\backslash C)\backslash B,\ ((S\backslash B)\backslash A)\backslash C,\ ((S\backslash B)\backslash C)\backslash A,\ ((S\backslash C)\backslash A)\backslash B,\ ((S\backslash C)\backslash B)\backslash A\}$.)

The combinatory rules, such as the various forms of composition, must then be defined so that they preserve this interpretation, crucially involving the same limit of $n \leq 4$ on the degree of generalized composition. Again, we pass over the details here, but if this constraint is not observed, then the expressive power of the grammar explodes to that of full indexed grammars (cf. note 14).

Hoyt and Baldridge (2008) use composition as a unary rule to define certain further linguistically motivated combinatory species, notably those corresponding to Curry's $\mathbf{D}$ combinator (Wittenburg 1987). All of these rules obey the principles (22) and (23).

3.9. COMBINATORS AND THE THEORY OF GRAMMAR. It is important for any theory to keep its intrinsic degrees of freedom as low as possible with respect to the degrees of freedom in the data it seeks to explain. In the case of the theory of grammar, this means limiting the set of languages covered or generated to the smallest "natural family of languages" that includes all and only the possible human languages. That set is known to lie somewhere between the context-free and context-sensitive levels of the Chomsky Hierarchy. Since most phenomena of natural language look context-free, the lower end of that vast range is the region to aim for.

Quite simple systems of combinators, including the typed **BTS** system that underlies CCG, are, in the absence of further restriction, equivalent to the simply typed λ-calculus—that is, to unconstrained recursively enumerable systems. It is the universal restrictive "projection principles" (22) and (23), together with the restriction of generalized composition $\mathbf{B^n}$ and X in the coordination category schema $(X\backslash X)/X$ in (12) to a bound on valency tentatively set at $n \leq 4$, that crucially restricts the (weak) expressive power of CCG to the same low level as tree-adjoining grammars (TAG, Joshi 1988) and linear indexed grammars (LIG, Gazdar 1988) (Vijay-Shanker and Weir 1994). This level is the least more expressive natural family of languages that is known than context free grammars.

This "nearly context-free" class is very much less expressive than the very properly sub-context-sensitive" classes such as full indexed grammars (IG, Aho 1968, 1969), which can express "non-constant growth" languages like a^{2^n}, and full linear context-free rewriting systems (LCFRS, Weir 1988), also known as multiple context free grammars (MCFG). The latter are shown by Kanazawa and Salvati 2012 to include the total scrambling language MIX, and have been argued by Stabler 2011 to characterize Chomskian Minimalist Grammars.

A word is in order on the relation of the "nearly context-free" formalisms CCG, TAG, and LIG. While weakly equivalent, they belong to different branches of the (extended) Chomsky hierarchy. Weir showed that the natural generalization of TAG is to full LCFRS/MCFG, while the natural generalization of CCG is to full IG. LCFRS and IG do not stand in an inclusion relation, although both are very properly contained in—that is, very much less expressive than—full context sensitive grammars, which for practical purposes are essentially as expressive as recursively enumerable systems (Savitch 1987).

This low expressive power brings a proof of polynomial parsability (Vijay-Shanker and Weir 1994), the significance of which is that standard "divide and conquer" context-free parsing algorithms readily generalize to CCG and the related formalisms. This fact, together with its semantic transparency, is the key to the widespread uptake of CCG in computational linguistic applications, particularly those which require semantic interpretation and/or accuracy with long-range dependencies, such as semantic parsing, question answering, and text entailment (see below).

In considering the rival merits and attractions of some of the alternative forms of CG considered below, it is worth keeping in mind the question of whether they are comparable constrained in expressive power and complexity.

4. CATEGORIAL GRAMMARS WITH WRAP COMBINATORY RULES. Categories like (8) exhibit a pleasing transparency between the syntactic type $(S\backslash NP)/NP$ and the logical form $\lambda x \lambda y.likes'xy$: both functions want the object as first argument and the subject as last argument, reflecting the binding asymmetry below in the fact that the subject commands (is attached higher than) the object:

(40) a. John$_i$ likes himself$_i$.
b. #Himself$_i$ likes John$_i$.

(Indeed, in such cases the λ-binding in the logical form is redundant: we could just write it as *likes*$'$.)

However, similar binding asymmetries like the following are not so simple:

(41) a. I introduced John$_i$ to himself$_i$.
b. #I introduced himself$_i$ to John$_i$.

If we want to continue to capture the binding possibilities in terms of command, then we need to amke a choice. If we want to retain the simple surface compositional account of extraction and coordination proposed in sections 3.5 and 3.6 shall need a category like the following for the ditransitive, in which derivational command and logical form command are no longer homomorphic:

(42) introduced := $((S\backslash NP)/PP)/NP : \lambda n \lambda p \lambda y.introduced' pny$

This category captures the asymmetry in (41) in the fact that the first NP argument n lf-commands the second PP argument p in the logical form $introduced' pny$, despite the fact that the former commands the latter in the derivation according to the syntactic type. As a consequence, we can no longer eliminate the explicit λ-binding logical form, and if binding theory asymmetries like (41) are to be explained in terms of command, it cannot be derivational command, but must be commmand at the level of logical form.

This use of logical form should not been seen as proliferating levels of representation in the theory of grammar, as Dowty has suggested. In such versions of CG, logical form is the *only* structural level of representation: no rule of syntax or semantics is defined in terms of derivation structure.

If on the other hand we wish to continue to account for such binding asymmetries in terms of derivational command, maintaining strict isomorphism between syntactic and semantic types in ditransitive categories like (42), and eliminating any distinction between derivational structure and logical form, as Bach, Dowty, and Jacobson, among others, have argued, we need a different category for *introduced*, taking the PP as its first most oblique argument, and the NP as its second.

(43) introduced := $((S\backslash NP)/_{\mathscr{RW}}NP)/PP : \lambda p \lambda n \lambda y.introduced'' pny$

We must also introduce a further combinatory family of “wrapping” rules that “infix” the second string argument as first argument, as in (44), marking the second argument of the syntactic type of English polytransitive verbs for combination by such a rule, as in (43):[15]

(44) *Right Wrap:* $(>\mathscr{RW})$

$$(X/_{\mathscr{RW}}Y)/Z : f \quad Y : a \quad Z : b \quad \Rightarrow \quad X : fba$$

Bach (1979), Jacobson (1992), and Dowty (1997) have strongly advocated the inclusion of wrap rules and categories in CG, both empirically on the basis of a wide variety of constructions in a number of languages, and on grounds of theoretical parsimony, based on the fact that all logical forms in categories and rules like (43) and (44) are eliminable, and derivation structure is the only level of representation in the theory.

There can be no doubt of the empirical strength of the generalization that the linear order of arguments in verb-initial constructions is typically the inverse of that required for the command theory of binding. However, the argument from parsimony is less conclusive. Dowty’s claim seems to be that the naive category (42) makes an intrinsic use of logical form to switch the order of application to the arguments. However, this operation is entirely local to the lexicon, and therefore seems entirely equivalent to the same information implicit in the syntactic category inclusion of wrap categories like (43) in place of (42), together with wrap rules like (44) engenders considerable complication in the syntactic theory. In particular, the simple and elegant account of

[15] Such rules correspond to Curry and Feys 1958 combinator **C**. There are actually a number of ways that wrap might be written as a combinatory rule, and it is not always clear from the literature which is assumed. I follow the categorial notation of Bach (1979) and Jacobson (1992).

constituent cluster coordination as a necessary corollary prediction in the **BST** system including type-raising and composition alone, exemplified earlier by (26), is no longer available in a combinatory grammar including rule (44) (see Dowty 1997 for a very clear exposition of the difficulty).

Dowty has proposed a number of solutions to this problem, including an "ID/LP" version of categorial grammar (Dowty 1996), following Zwicky (1986), in which combination (Immediate Domination ID) is discontinuous, and is filtered by constraints on order (Linear Precedence, LP). Dowty (1997) shows how a multimodal categorial grammar of a kind discussed in the next section can be extended with uninterpreted structural rules of permutation triggered by typed string concatenation operators, (The latter are reminiscent of the product operator used by Pickering and Barry (1993) as an alternative to composition and type-raising for making argument clusters like "Bill to Sue" into constituents.) Dowty's grammar captures the full range of constructions addressed in standard CCG, including crossed dependency and argument cluster coordination. However, the theoretical difficulties are considerable, and the details must be passed over here. It remains unclear whether structural rules and product operators can be made to force Ross' generalization (29) in the straightforward way that CCG restricted to **BTS** combinatory rules does.

5. LAMBEK GRAMMARS. Despite a superficially similar categorial slash notation, Lambek Grammars constitute a quite different approach to the extension of the pure Categorial Grammar of Ajdukiewicz (1935) and Bar-Hillel (1953), building on the view of Categorial Grammar *as a logic* initiated by Lambek (1958, 1961). This approach treats the categorial slash as a form of (linear) logical implication, for which the antecedent-canceling application rules perform the role of modus ponens.[16]

Lambek's contribution was to add further logical axioms of associativity, product-formation, and division, from which harmonic composition and order-preserving type raising emerge as theorems (although crossing composition and the **S** combinator do not). The resulting calculus was widely assumed to be weakly equivalent to context-free grammar, although the involvement of an axiom schema in its definition meant that the actual proof of the equivalence was not forthcoming until the work of Pentus (1993). (In fact, the original Lambek calculus supports essentially the same analysis of unbounded dependency as context-free GPSG, but with the advantage of semantic isomorphism.) Paradoxically, despite CF equivalence, the original context-free calculus lacks a polynomial-time recognition algorithm, because the mapping to equivalent CF grammars that would afford such an algorithm is not itself polynomial, a result that was also widely anticipated but hard to prove, until Pentus (2003) proved that as well.

5.1. PREGROUP GRAMMARS. Partly in response, Lambek (2001) has recently proposed Pregroup Grammar as a simpler context-free base for type-based grammars. Like LG, pregroup grammars (PG) are context-free (Buszkowski 2001), but they have a polynomial conversion to CFG (Buszkowski and Moroz 2008). They inherit the associative property of LG of combining the English subject and transitive verb as a single operation, rather than tediously requiring two successive operations of type raising and composition. While Pregroup Grammar thereby obscures the relation between type

[16] The attraction of viewing grammars as logics rather than combinatory algebras or calculi seems to be that they then support a model theory that can be used as a basis for proofs of soundness and completeness of the syntax. It should be noticed that such a logic and model theory is distinct from the standard logic implicit in the applicative semantics for the categorial grammar itself or the corresponding set of standard context-free productions.

raising and grammatical case, the associativity operator obeys the Principles of Adjacency, Consistency, and Inheritance defined above, and is definable in terms of the combinators **B** and **T**. It could therefore be consistently incorporated in CCG grammar or parser, if so desired, without losing the advantages of the latter's mildly trans-context-free expressive power.

Pregroup grammars are so called because of the involvement of free pregroups as a mathematical structure in their definition. Kobele and Kracht (2005) show that one natural generalization of pregroup grammars defined in terms of *all* pregroups and allowing the empty string symbol generates all recursively enumerable languages—that is, is equivalent to the universal Turing machine. Tupled Pregroup Grammars, a more constrained generalization of context-free Pregroup Grammars by Stabler (2004a,b), have been shown to be weakly equivalent to Set-Local MC-TAG, the Multiple Context-Free Grammars (MCFGs) of Seki et al. (1991), and the Minimalist Grammars of Stabler and Keenan (2003). Such calculi are, as noted earlier, much more expressive than CCG and TAG, being equivalent to full LCFRS. Their learnability is investigated by Béchet et al. (2007).

Other generalizations of the original Lambek calculi include Abstract Categorial Grammar (de Groote 2001; de Groote and Pogodalla 2004), Lambda Grammar (Muskens 2007), Convergent Grammar (CVG) (de Groote et al., forthcoming), and the Lambek-Grishin calculus or Symmetric Categorial Grammar (Moortgat 2007, 2009). The expressive power of these systems seems likely to be also that of full LCFRS.

5.2. CATEGORIAL TYPE LOGIC. In a separate development from the original Lambek calculi, Moortgat (1988a) and van Benthem (1988) importantly showed that simply adding further axioms such as permutation or crossing composition to the Lambek calculus causes it (unlike CCG) to collapse into permutation completeness. Instead, they and Oehrle (1988) proposed to extend the Lambek calculus using "structural rules" and typed slashes of the kind originated by Hepple (1990) and discussed in section 3.7, to control associativity and permutativity. Specific proposals of this kind include Hybrid Logical Grammar (Hepple 1990, 1995), Type-Logical Grammars (Morrill 1994, 2011; Carpenter 1997), Term-Labeled Categorial Type Systems (Oehrle 1994), Type-Theoretical Grammars (Ranta 1994), Multimodal Categorial Grammar (Moortgat 1997, Carpenter 1995, Moot and Puite 2002; Moot and Retoré 2012, cf. Baldridge 2002).

Carpenter and Baldridge showed that the Type-Logical Categorial Grammars were potentially very expressive indeed. TLG thus provides the general framework for comparing categorial systems of all kinds, including CCG (Baldridge and Kruijff 2003).

6. SEMANTICS IN CATEGORIAL GRAMMARS. Although the primary appeal of categorial grammars since Montague derives from the aforementioned close relation between categorial derivation and semantic interpretation, there is a similar diversity in theories of the exact nature of the mapping from categorial syntax and semantics to that concerning the exact nature of the syntactic operations themselves. The main focus of disagreement is on the question of whether a representational level of logical form distinct from syntactic derivational structure is involved or not.[17]

Montague himself was somewhat divided on this question. In Montague 1970a,b he

[17] Montague's stern use of the word "proper" in his 1973 title may have reflected the fact that his treatment of quantified terms like *Every Syren* and *a song* assigned them the type of proper nouns like *John* and *she* under the Description Theory of of Russell and Frege. The implication that other treatments were somehow *im*proper may have been a donnish pun. This possibility is not always appreciated by those who proliferate titles in semantics of the form "The Proper Treatment of X".

argued that there was no logical reason why natural languages, any more than formal logical languages, should have more than one representational level. However, in 1973, he included operations such as "quantifying in" that made apparently essential use of logical form via operations with no surface syntactic reflex. The literature since Montague has similarly followed two diverging paths, which I will distinguish as "sacred" and "profane".[18]

The sacred path usually adhered to by Partee, Dowty, Jacobson, Hendriks, and Szabolci follows the Montague of 1970 in assuming a standard logic such as first-order (modal) predicate logic (FOPL) as the language of thought or logical form, and seeking to eliminate 1973-style intrinsic use of logical form by elaborating the notion of syntactic derivation via various additional combinators, such as wrapping, type-lowering, and specialized binding combinators. The official name for this sacred approach is "Direct Surface Compositionality": it seeks to eliminate logical form and make derivation the sole structural level of representation.[19]

Because phenomena like the dependency of bound variable pronouns in (1d) tend to be much less restricted than strictly syntactic dependencies like relativisation, the sacred approach has in practice shown itself quite willing to abandon the search for low expressive power in surface syntax characteristic of GPSG and CCG. There is a natural affinity between the sacred approach to semantics and the more expressive forms of Lambek and Type-Logical grammars, although Carpenter 1997 combines Morrill's type logical syntax with an essentially profane semantics.

The profane approach follows the opposite strategy. If the chosen representation for logical form appears to require non-monotonic structure-changing operations such as quantifier-raising and quantifying-in (or extraneous equivalent type-changing derivational operations), then there must just be something wrong in the choice of logical representation. The logical language itself should be changed, perhaps by eliminating quantifiers and introducing discourse referents (Kamp and Reyle 1993) and referring expressions (Fodor and Sag 1982), or Skolem terms (Kratzer 1998; Steedman 1999, 2012; Schlenker 2006) in their place. If pronoun binding obeys none of the same generalizations as syntactic derivation, do it non-derivationally (Steedman 2012). It is surface derivation that should be eliminated as a representational level. In fact there may be very many derivation structures yielding the same logical form. There is a close affinity between the profane approach to semantics and computational linguistics.

In the end, the difference between these two approaches may not be very imporant, since they agree on the principle that there should be only one level of structural representation, and differ only on its relation to surface derivation. Nevertheless, they have in practice led to radically different semantic theories. We will consider them briefly in turn.

6.1. THE SACRED APPROACH: DIRECT SURFACE COMPOSITIONALITY. The sacred approach to semantics differs from the profane in making the following assumptions:

1. Surface derivational structure is the *only* representational level in the theory of grammar corresponding to logical form and supporting a model theoretic semantics.

[18] Of course, I am exaggerating the differences for mnemonic reasons. Most of us combine both traits.

[19] If all one wants to do with a logic is prove truth in a model, then structural representation itself can technically be eliminated entirely, in favor of direct computation over models. But if you want to do more general inference, then in practice you need *some* structural representation.

2. No other representation of logical form is necessary to the theory of grammar, and any use of logical formulae to represent meanings distinct from derivation structures is a mere notational convenience.

Dowty (1979), Partee and Rooth (1983), Szabolcsi (1989), Chierchia (1988), Hendriks (1993), Jacobson (1999), Barker (2002), Jäger (2005), and colleagues follow Montague in seeking to extend categorial grammar to the problem of operator semantics, including pronoun-binding by quantifiers, exemplified in the following:

(45) a. Every sailor$_i$ believes every Syren$_j$ knows he$_i$ heard her$_j$
b. Every sailor$_i$ believes every Syren$_j$ knows she$_j$ saw him$_i$

As noted in the introduction, such examples show that bound variable pronouns are free to nest or cross dependencies with quantificational binders. Such dependencies are also immune to the island boundaries that block relativization, of which the fixed subject condition illustrated in (46).

(46) a. Every sailor$_i$ believes that he$_i$ won.
b. #A Syren who(m) every sailor$_i$ believes that won$_i$

Following Szabolcsi (1989), these authors seek to bring pronoun binding within the same system of combinatory projection from the lexicon as syntactic dependency. Jacobson (1999; 2007:203-4) assigns pronouns the category of a nominal syntactic and semantic identity function, with which the verb can compose. However, instead of writing something like $NP|NP : \lambda x.x$ as the category for "him" she writes $NP^{\{NP\}} : \lambda x.x$.[20]

Constituents of the type of verbs can be subject to a unary form of composition or "division", which she calls the Geach Rule.[21]

For example, intransitive "won" $S\backslash NP$ can acquire a further category $S^{\{NP\}}\backslash NP^{\{NP\}} : \lambda i\lambda y.won'y$. Such Geached types can combine with a pronoun (or any NP $NP^{\{NP\}}$ containing a pronoun, such as "his mother" or "a man he knows") by function composition, so that "he won" yields the category $S^{\{NP\}} : \lambda y.won'y$ rather than the standard category $S : won'him'$.[22]

Constituents of the same type as verbs can also undergo a unary combinatory rule which Jacobson calls **z**. For example, "believes" $(S\backslash NP)/S : \lambda s\lambda y.believes'sy$ can become $(S\backslash NP)/S^{\{NP\}} : \lambda p\lambda y.believes'(py)y$, which on application to "he won", $S^{\{NP\}} : \lambda y.won'y$, yields "believes he won", $S\backslash NP : \lambda y.believes'(won'y)y$. This predicate can combine as the argument of the standard generalized quantifier category for "Every sailor", $S/(S\backslash NP) : \lambda p.\forall y[sailor'y \Rightarrow believes'(py)y]$, to yield:

(47) Every sailor believes he won $S : \forall y[sailor'y \Rightarrow believes'(won'y)y]$

However, the freedom of multiple pronouns to either nest or intercalate bindings to scoping quantifiers, as in (45), coupled with the fact that binding may be into and out of strong islands such as English subject position, as in (46a), means that Jacobson's binding categories $X^{\{Y\cdots\}}$ have to form a separate parallel combinatory categorial system. It is not easy to see who to combine these two combinatory systems in a single grammar without exploding the category type system. Jacobson 1999:105,n.19 in fact proposes to give up the CCG account of extraction entirely, and to revert to something like the GPSG account (although it is known to be incomplete—see Gazdar 1988).

[20] Jacobson does not usually include set-delimiting braces $\{\dots\}$ in her notation for categories including bindable pronouns, but sentences like (45) show that in general these superscripts are composable (multi)sets.

[21] The unary Geach rule is implicitly schematized as unary $\mathbf{B}^n$ along lines exemplified for (34).

[22] Jacobson η-reduces the redundant abstraction in terms like $\lambda y.won'y$ to e.g. won', but in the absence of explicit types (as in $won'_{\langle e,t\rangle}$) I let it stand as more intelligible.

Similar problems attend attempts to treat quantifier scope via surface syntactic derivation. There is a temptation to think that the "object wide scope" reading of the following scope-ambiguous sentence (a) arises from a derivation in which a type-raised object $S\backslash(S/NP) : \lambda p.\exists x[woman\, x \wedge p\, x]$ derivationally commands the nonstandard constituent $S/NP : \lambda y.\forall z[man\ z \Rightarrow loves\ yz]$ to yield $S : \exists x[woman\ x \wedge \forall z[man\ z \Rightarrow loves\ xz]]$ (eg. Bernardi 2002:22).

(48) a. Every man loves a woman.
b. Every man loves and every boy fears a woman.

However, (48b) has *only* the object c-command derivation. Yet it undoubtedly has a narrow scope reading involving possibly different women. If scope is to be handled by derivational command we therefore have to follow Hendriks (1993) in introducing otherwise syntactically unmotivated type-*lowering* operations, with attendant restrictions to ensure that they do not then raise again to yield unattested mixed-scope readings, such as the one in which men love possibly different narrow scope women, and boys all fear the same wide scope woman.

Solutions to all of these problems have been proposed by Hendriks, Jacobson, and Jäger, and in the continuation-based combinatory theory of Barker (2001) and Shan and Barker (2006), but it is not yet clear whether they can be overcome without compromising the purely syntactic advantages of CG with respect to, for example, extraction. If not, there is some temptation to consider binding and scope as distinctively anaphoric properties of logical form that are orthogonal to syntactic derivation, as is standard in logic proper and in programming language theory.

6.2. THE PROFANE APPROACH: NATURAL SEMANTICS. The profane approach to semantics haas been called "natural semantics," in homage to Lakoff's 1970 proposal for a natural logic underlying the Generative Semantics approach to the theory of grammar in the early seventies, of which Partee (1970) was an early categorial exponent. The proposal was influentially taken up by Sánchez Valencia (1991, 1995) and Dowty (1994) within categorial frameworks, and extended elsewhere by MacCartney and Manning (2007). Natural semantics departs from the sacred approach, and from other proposals within generative semantics of that period, in three important respects.

1. Logical form is the *only* representational level in the theory of grammar supporting a model-theoretic semantics.

2. Surface-syntactic derivation is *not* a level of representation in the theory of grammar, and does not require or support a model theory distinct from that of logical form. It is merely a description of the computation by which a language processor builds (or expresses) a logical form from (or as) a string of a particular language, and is entirely redundant with respect to interpretation or realization of meaning.

3. The language of natural logical form should not be expected to be anything like traditional logics such as FOPL, invented by logicans and mathematicians for very different purposes. The sole source of information we have as to the nature of this "hidden" language of thought is linguistic form, under the strong assumption of syntactic/semantic homomorphism shared by all categorial grammarians.

The profane natural approach to semantics therefore questions the core assumption of Lambek and Type Logical approaches that surface syntax is itself a *logic*. By the same token, natural semantics questions the core assumption of Direct Surface Compositionality concerning the redundancy of logical form. It is derivational structure that is semantically redundant, not logical form. If keeping derivation simple requires lexical logical forms to wrap derivational arguments, as in (42), then let them do so (Carpenter 1997:437). If the attested possibilities for quantifier scope alternation do not seem to be compatible with any simple account of derivation, then replace generalized quantifiers with devices that simplify derivations, such as referring expressions or Skolem terms (Steedman 2012).

There have been recent signs of a rapprochement between these views. Jacobson (2002:60) points out that the use of WRAP rules in some combinatory versions of CG demands structural information that is exactly equivalent to that in the profane λ-binding category (42) and Dowty's 1997 concatenation modalities. Dowty (1996, drafted around 1991;2007) has drawn a distinction following Curry (1961) between a level of "tectogrammatics", defining the direct compositional interpretation of the equivalent of logical form, and one of "phenogrammatics", equivalent to surface derivation. Dowty 2007:58-60 regards the responsibiity for defining the ways phenogrammatical syntax can "encode" tectogrammatical structure as a question for psycholinguistics, as concerning a processor which he seems to view as approximate and essentially related to what used to be called performance. These are clear and welcome signs of convergence between these extremes. Perhaps, as is often the case in human affairs, the sacred and the profane are quite close at heart.

7. COMPUTATIONAL AND PSYCHOLINGUISTIC APPLICATIONS. It is unfashionable nowadays for linguistic theories to concern themselves with performance. Moreover, most contemporary psychological and computational models of natural language processing return the compliment by remaining ostentatiously agnostic concerning linguistic theories of competence.

Nevertheless, one should never forget that linguistic competence and performance must come into existence together, as a package deal in evolutionary and developmental terms. The theory of syntactic competence should therefore ultimately be transparent to the theory of the processor. One of the attractions of categorial grammars is that they support a very direct relation between competence grammars and performance parsers.

The central problems for practical language processing by humans or by machine are twofold. First, natural language grammars are very large, involving thousands of constructions. (The lexicon derived from section 02-21 of the categorial CCGbank version of the 1M-word Penn Wall Street journal corpus (Hockenmaier and Steedman 2007) contains 1224 distinct category types, of which 417 only appear once, and is known to be incomplete.)

Second, natural grammars are hugely ambiguous. As a result, quite unremarkable sentences of the kind routinely encountered in an even moderately serious newspaper have thousands of syntactically well-formed analyses. (The reason that human beings are rarely aware that a sentence has more than a single analysis is that nearly all of the other analyses are semantically anomalous, especially when the context of discourse is taken into account.)

The past few years have shown that ambiguity of this degree can be handled practically in parsers of comparable coverage and robustness to humans, by the use of statistical models, and in particular those that approximate semantics by modeling semantically relevant head-dependency probabilities such as those between verbs and the nouns that head their (subject, object, etc.) arguments (Hindle and Rooth 1993; Magerman 1995; Collins 1997). Head-word dependencies compile into the model a powerful mixture of syntactic, semantic, and world-dependent regularities that can be amazingly effective in reducing search.[23]

Categorial grammars of the kinds discussed here were initially expected to be poorly adapted to practical parsing, because of the additional derivational ambiguity introduced by the nonstandard constituency discussed at the end of section 3.2. However, a number of algorithmic solutions minimizing redundant combinatory derivation have been discovered (König 1994; Eisner 1996; Hockenmaier and Bisk 2010).

Doran and B. Srinivas (2000),Hockenmaier and Steedman (2002), Hockenmaier (2003, 2006), Clark and Curran (2004), and Auli and Lopez (2011a,b,c) have shown that CCG can be applied to wide-coverage, robust parsing with state-of-the-art performance. Granroth-Wilding (2013) has successfully applied CCG and related statistical parsing methods to the analysis of musical harmonic progression. Birch et al. (2007) Hassan et al. (2009) and Mehay and Brew (2012) have used CCG categories and parsers as models for statistical machine translation.

Prevost (1995) applied the nonstandard surface structures of CCG to the control of prosody and intonation in synthesis of spoken English from information-structured semantics. White (2006) extended this to efficient sentence realization for CCG, while Kruijff-Korbayová et al. (2003) have applied CCG in dialog generation. Gildea and Hockenmaier (2003) and Boxwell et al. (2009, 2010) have applied CCG to and with Semantic Role Labeling.

Briscoe (2000), Buszkowski and Penn (1990) and Kanazawa (1998) discuss learnability of categorial grammars, while Villavicencio (2002, 2011), Buttery (2004, 2006), McConville (2006), Zettlemoyer and Collins (2005, 2007), Kwiatkowski et al. (2010, 2011, 2012), and Krishnamurthy and Mitchell (2012) have exploited the semantic transparency of CCG to model semantic parsing and grammar induction from pairs of strings and logical forms, as a model of child language acquisition in humans. Piantadosi et al. (2008) have used CCG to model acquisition of quantifier semantics.

Pereira (1990) applied a unification-based version of Lambek Grammar to the derivation of quantifier scope alternation. Bos and Markert (2005a,b, 2006) and Zamansky et al. (2006) have applied DRT-semantic CCG and Lambek Grammars to text entailment, while Harrington and Clark (2007, 2009) have used a CCG parser to build semantic networks for large-scale question answering, using spreading activation to limit search and update. Indeed, the main current obstacle to further progress in computational applications is the lack of labeled data for inducing bigger lexicons and models for stronger parsers, a problem to which unsupervised or semisupervised learning methods appear to offer the only realistic chance of an affordable solution. The latter methods have been applied to categorial grammars by Watkinson and Manandhar (1999); Thomforde (2013) and Boonkwan (2013).

Grefenstette et al. (2011), Grefenstette and Sadrzadeh (2011), and Kartsaklis et al. (2013) propose an application of Pregroup Grammars to compositional assembly of vector-based distributional semantic interpretations (cf. Mitchell and Lapata 2008). Lewis and Steedman (2013) propose a different form of distributional semantics for open-class words in CCG based on paraphrase clustering, and apply it to text entailment tasks.

Multimodal and Type-Logical Categorial Grammars support the notion of "proof nets" (Moortgat 1997; Moot and Puite 2002), related to an earlier idea of "count invariance", which has been applied in wide coverage parsing in the Grail system by Moot (2010), Moot and Retoré (2012). Morrill (2000, 2011) seeks to model "garden-path"

[23] The two main varieties of statistical model, the probabilistic/generative and the weighted/discriminative, are discussed by Smith and Johnson (2007)..

phenomena and other aspects of human psycholinguistic performance algorithmically using proof nets, exploiting the potential of generalized categorial grammars to deliver incremental predominantly left-branching analyses supporting semantic interpretation.

8. CONCLUSION. Categorial grammars of all kinds have attractive properties for theoretical and descriptive linguists, psycholinguists, and computational linguists, because of their strict lexicalization of all language-specific information, and the consequent simplicity of the interface that they offer between syntactic derivation and compositional semantics on the one hand, and parsing algorithms and the statistical and head dependency-based parsing models that support robust wide-coverage natural language processing on the other. The non-standard notion of surface derivational structure that they offer is particularly beneficial in the cross-linguistic analysis of coordination, extraction, and intonation structure.

FURTHER READING. Lyons 1968 includes an early and far-seeing introduction to pure categorial grammar and its potential role in theoretical linguistics. Wood 1993 provides a balanced survey of historical and early modern approaches. Buszkowski et al. 1988, Oehrle et al. 1988, and Casadio and Lambek 2008 provide useful collections of research articles, the former reprinting a number of historically significant earlier papers. Moortgat 1997, Steedman and Baldridge 2011, Lambek 2008, Morrill 2011, Moot and Retoré 2012, and Bozşahin 2012 are more specialized survey articles and monographs on some of the contemporary varieties of categorial grammar discussed above, often with comparisons across approaches. Partee 1976, Dowty 1979, and Barker and Jacobson 2007 represent (mostly) sacred approaches to semantics within categorial grammar, while Carpenter 1997 and Steedman (2012) represent the unabashedly profane.

A number of open-source computational linguistic tools for CCG applications are available at `http://groups.inf.ed.ac.uk/ccg/software.html` and via SourceForge at `http://openccg.sourceforge.net` The categorial CCGbank version of the Penn WSJ treebank is available from the Linguistic Data Consortium (Hockenmaier and Steedman 2005) and has been improved by James Curran. Hockenmaier has developed a German CCGbank (Hockenmaier 2006). The Grail type-logical parser and related resources are available at `http://www.labri.fr/perso/moot/grail3.html`.

ACNOWLEDGEMENTS I am grateful to Polly Jacobson for answering a number of questions during the preparation of this article, which was supported by EU ERC Advanced Fellowship 249520 GRAMPLUS and EU IP EC-FP7-270273 Xperience.

BIOGRAPHICAL NOTE. The author is the Professor of Cognitive Science in the School of Informatics at the University of Edinburgh. He was trained as a psychologist and a computer scientist. He is a Fellow of the British Academy, the American Association for Artificial Intelligence, and the Association for Computational Linguistics. He has also held research positions and taught at the Universities of Sussex, Warwick, Texas, and Pennsylvania, and his research covers a wide range of issues in theoretical and computational linguistics, including syntax, semantics, spoken intonation, knowledge representation, and robust wide-coverage natural language processing.

REFERENCES

ADES, ANTHONY and STEEDMAN, MARK. 1982. On the order of words. Linguistics and Philosophy 4 517–558.

AHO, ALFRED. 1968. Indexed grammars—an extension of context-free grammars. Communications of the

Association for Computing Machinery 15 647–671.

——. 1969. Nested stack automata. Communications of the Association for Computing Machinery 16 383–406.

AJDUKIEWICZ, KAZIMIERZ. 1935. Die syntaktische Konnexität. Storrs McCall (ed.), Polish Logic 1920–1939, 207–231. Oxford: Oxford University Press. Trans. from *Studia Philosophica*, 1, 1-27.

AULI, MICHAEL and LOPEZ, ADAM. 2011a. A comparison of loopy belief propagation and dual decomposition for integrated CCG supertagging and parsing. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, 470–480. Portland, OR: ACL.

——. 2011b. Efficient CCG parsing: A* versus adaptive supertagging. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, 1577–1585. Portland, OR: ACL.

——. 2011c. Training a log-linear parser with loss functions via softmax-margin. Proceedings of the Conference on Empirical Methods in Natural Language Processing, 333–343. Edinburgh: ACL.

BACH, EMMON. 1979. Control in Montague Grammar. Linguistic Inquiry 10 513–531.

——. 1983. Generalized categorial grammars and the English auxiliary. Frank Heny and Barry Richards (eds.), Linguistic Categories: Auxiliaries and Related Puzzles, II, 101–120. Dordrecht: Reidel.

BALDRIDGE, JASON. 1998. Local Scrambling and Syntactic Asymmetries in Tagalog. Master's thesis, University of Pennsylvania.

——. 2002. Lexically Specified Derivational Control in Combinatory Categorial Grammar. Ph.D. thesis, University of Edinburgh.

BALDRIDGE, JASON and KRUIJFF, GEERT-JAN. 2003. Multi-Modal Combinatory Categorial Grammar. Proceedings of 11th Annual Meeting of the European Association for Computational Linguistics, 211–218. Budapest.

BAR-HILLEL, YEHOSHUA. 1953. A quasi-arithmetical notation for syntactic description. Language 29 47–58.

BAR-HILLEL, YEHOSHUA, GAIFMAN, CHAIM, and SHAMIR, ELIYAHU. 1964. On categorial and phrase structure grammars. Yehoshua Bar-Hillel (ed.), Language and Information, 99–115. Reading, MA: Addison-Wesley.

BARKER, CHRIS. 2001. Integrity: A syntactic constraint on quantifier scoping. Proceedings of the 20th West Coast Conference on Formal Linguistics, 101–114. Somerville, MA: Cascadilla.

——. 2002. Continuations and the nature of quantification. Natural Language Semantics 10 211–242.

BARKER, CHRIS and JACOBSON, PAULINE (eds.). 2007. Direct Compositionality. Oxford University Press.

BÉCHET, DENIS, FORET, ANNIE, and TELLIER, ISABELLE. 2007. Learnability of pregroup grammars. Studia Logica 87 225–252.

BEKKI, DAISUKE. 2010. Nihongo Bunpoo no Keesiki Riron: Katsuyoo taikee, Toogo koozoo, Imi goosee [Formal theory of Japanese grammar: The system of conjugation, syntactic structure, and semantic composition]. Nihongo Kenkyuu Soosyo [Japanese Frontier Series] 24. Tokyo: Kurosio Publishers.

BERNARDI, RAFFAELLA. 2002. Reasoning with Polarity in Categorial Type-Logic. Ph.D. thesis, Universiteit Utrecht.

BERWICK, ROBERT and EPSTEIN, SAMUEL. 1995. Computational minimalism: The convergence of 'Minimalist' syntax and Categorial Grammar. A. Nijholt, G. Scollo, and R. Steetkamp (eds.), Algebraic Methods in Language Processing 1995: Proceedings of the Twente Workshop on Language Technology 10, jointly held with the First Algebraic Methodology and Software Technology (AMAST) Workshop on Language Processing. Enschede, The Netherlands: Faculty of Computer Science, Universiteit Twente.

BIRCH, ALEXANDRA, OSBORNE, MILES, and KOEHN, PHILIPP. 2007. CCG supertags in factored translation models. Proceedings of the 2nd Workshop on Statistical Machine Translation, 9–16. held in conjunction with ACL, Prague: ACL.

BITTNER, MARIA. 2011. Time and modality without tenses or modals. Renate Musan and Monika Rathert (eds.), Tense Across Language, 147–188. Tübingen: Niemeyer.

——. 2014. Temporality: Universals and Variation. Malden, MA: Wiley-Blackwell.

BOECKX, CEDRIC. 2008. Bare Syntax. Oxford: Oxford University Press.

BOONKWAN, PRACHYA. 2013. Scalable Semi-Supervised Grammar Induction using Cross-Linguistically Parameterized Syntactic Prototypes. Ph.D. thesis, Edinburgh.

BOS, JOHAN and MARKERT, KATJA. 2005a. Combining shallow and deep NLP methods for recognizing textual entailment. Proceedings of the First PASCAL Challenge Workshop on Recognizing Textual Entailment, 65–68. `http://www.pascal-network.org/Challenges/RTE/`: Pascal.

——. 2005b. Recognising textual entailment with logical inference. Proceedings of the 2005 Conference on Empirical Methods in Natural Language Processing (EMNLP 2005), 628–635.

——. 2006. When logical inference helps determining textual entailment (and when it doesn't). Proceedings of the Second PASCAL Challenge Workshop on Recognizing Textual Entailment. Pascal.

BOSCH, PETER. 1983. Agreement and Anaphora. New York: Academic Press.

BOUMA, GOSSE and VAN NOORD, GERTJAN. 1994. Constraint-based categorial grammar. Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics. Las Cruces, NM: ACL.

BOXWELL, STEPHEN, MEHAY, DENNIS, and BREW, CHRIS. 2009. Brutus: A semantic role labeling system incorporating CCG, CFG, and dependency features. Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP, 37–45. Suntec, Singapore: ACL.

——. 2010. What a parser can learn from a semantic role labeler and vice versa. Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, 736–744. ACL.

BOZŞAHIN, CEM. 1998. Deriving predicate-argument structure for a free word order language. Proceedings of International Conference on Computational Linguistics, 167–173. Montreal: ACL.

——. 2002. The combinatory morphemic lexicon. Computational Linguistics 28 145–186.

——. 2012. Combinatory Linguistics. Berlin: de Gruyter.

BRISCOE, TED. 2000. Grammatical acquisition: Inductive bias and coevolution of language and the language acquisition device. Language 76 245–296.

BUSZKOWSKI, WOJCIECH. 2001. Lambek grammars based on pregroups. Proceedings of Logical Aspects of Computational Linguistics 95–109.

BUSZKOWSKI, WOJCIECH, MARCISZEWSKI, WITOLD, and VAN BENTHEM, JOHAN. 1988. Categorial Grammar. Amsterdam: John Benjamins.

BUSZKOWSKI, WOJCIECH and MOROZ, KATARZYNA. 2008. Pregroup grammars and context-free grammars. Claudia Casadio and Joachim Lambek (eds.), Computational Algebraic Approaches to Natural Language, 1–21. Monzah: Polimetrica.

BUSZKOWSKI, WOJCIECH and PENN, GERALD. 1990. Categorial grammars determined from linguistic data by unification. Studia Logica 49 431–454.

BUTTERY, PAULA. 2004. A quantitative evaluation of naturalistic models of language acquisition; the efficiency of the triggering learning algorithm compared to a categorial grammar learner. Proceedings of the Workshop on Psycho-Computational Models of Human Language Acquisition, 3–10. Geneva: COLING.

——. 2006. Computational Models for First Language Acquisition. Ph.D. thesis, University of Cambridge.

CARPENTER, BOB. 1995. The Turing-completeness of Multimodal Categorial Grammars. Jelle Gerbrandy, Maarten Marx, Maarten de Rijke, and Yde Venema (eds.), Papers Presented to Johan van Benthem in Honor of His 50th Birthday. European Summer School in Logic, Language and Information, Utrecht, 1999. Amsterdam: ILLC, University of Amsterdam.

——. 1997. Type-Logical Semantics. Cambridge, MA: MIT Press.

CASADIO, CLAUDIA. 2001. Non-commutative linear logic in linguistics. Grammars 4 167–185.

CASADIO, CLAUDIA and LAMBEK, JOACHIM (eds.). 2008. Computational Algebraic Approaches to Natural Language. Polimetrica.

ÇAKICI, RUKET. 2005. Automatic induction of a CCG grammar for Turkish. Proceedings of the Student Workshop, 43rd Annual Meeting of the ACL, Ann Arbor MI, 73–78. ACL.

——. 2009. Parser Models for a Highly Inflected Language. Ph.D. thesis, University of Edinburgh.

CHA, JEONGWON and LEE, GEUNBAE. 2000. Structural disambiguation of morpho-syntactic categorial parsing for Korean. Proceedings of the 18th International Conference on Computational Linguistics, 1002–1006. Saarbrücken: ACL.

CHIERCHIA, GENNARO. 1988. Aspects of a categorial theory of binding. Richard Oehrle, Emmon Bach, and Deirdre Wheeler (eds.), Categorial Grammars and Natural Language Structures, 125–151. Dordrecht: Reidel.

CHOMSKY, NOAM. 1957. Syntactic Structures. The Hague: Mouton.

——. 1995. Bare phrase structure. Gert Webelhuth (ed.), Government and Binding Theory and the Minimalist Program, 383–439. Oxford: Blackwell.

CIARDELLI, IVANO and ROELOFSEN, FLORIS. 2011. Inquisitive logic. Journal of Philosophical Logic 40 55–94.

CLARK, STEPHEN and CURRAN, JAMES R. 2004. Parsing the WSJ using CCG and log-linear models. Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, 104–111. Barcelona, Spain: ACL.

COLLINS, MICHAEL. 1997. Three generative lexicalized models for statistical parsing. Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics, 16–23. Madrid: ACL.

COOPER, ROBIN. 1983. Quantification and Syntactic Theory. Dordrecht: Reidel.

CORMACK, ANNABEL and SMITH, NEIL. 2005. What is coordination? Lingua 115 395–418.

CURRY, HASKELL. 1961. Some logical aspects of grammatical structure. Roman Jakobson (ed.), Structure of Language and its Mathematical Aspects, *Proceedings of the Symposium in Applied Mathematics*, vol. 12, 56–68. Providence, RI: American Mathematical Society.

CURRY, HASKELL B. and FEYS, ROBERT. 1958. Combinatory Logic: Vol. I. Amsterdam: North Holland.

DAVIDSON, DONALD and HARMAN, GILBERT (eds.). 1972. Semantics of Natural Language. Dordrecht: Reidel.

DE GROOTE, PHILIPPE. 2001. Towards Abstract Categorial Grammars. Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics, 148–155. Toulouse: ACL.

DE GROOTE, PHILIPPE and POGODALLA, SYLVAIN. 2004. On the expressive power of abstract categorial grammars: Representing context-free formalisms. Journal of Logic, Language and Information 13 421–438.

DE GROOTE, PHILIPPE, POGODALLA, SYLVAIN, and POLLARD, CARL. Forthcoming. About parallel and syntactocentric formalisms: What the encoding of Convergent Grammar into Abstract Categorial Grammar tells us. Fundamenta Informaticae .

DORAN, CHRISTY and B. SRINIVAS. 2000. A wide coverage CCG parser. Anne Abeille and Owen Rambow (eds.), Proceedings of the 3rd TAG+ workshop, Jussieu, March 1994, 405–426. Stanford, CA: CSLI Publications.

DOWTY, DAVID. 1979. Word Meaning in Montague Grammar. Dordrecht: Reidel, 1st edn.

——. 1982. Grammatical relations and Montague Grammar. Pauline Jacobson and Geoffrey K. Pullum (eds.), The Nature of Syntactic Representation, 79–130. Dordrecht: Reidel.

——. 1988. Type-raising, functional composition, and nonconstituent coordination. Richard Oehrle, Emmon Bach, and Deirdre Wheeler (eds.), Categorial Grammars and Natural Language Structures, 153–198. Dordrecht: Reidel.

——. 1994. The role of negative polarity and concord marking in natural language reasoning. Proceedings of the 4th Conference on Semantics and Theoretical Linguistics. Rochester, NY: CLC Publications, Cornell University.

——. 1996. Towards a minimalist theory of syntactic structure. Harry Bunt and Arthur van Horck (eds.), Discontinuous Constituency, 11–62. The Hague: Mouton de Gruyter.

——. 1997. Nonconstituent Coordination, Wrapping, and Multimodal Categorial Grammars: Syntactic form as logical form. Maria Luisa Dalla Chiara (ed.), Structures and Norms in Science, 347–368. Berlin: Springer. Extended version at `http://www.ling.ohio-state.edu/~dowty/`.

——. 2003. The dual analysis of adjuncts/complements in categorial grammar. Ewald Lang, Claudia Maienborn, and Cathry Fabricius-Hansen (eds.), Modifying Adjuncts, 33–66. Berlin: Mouton de Gruyter.

——. 2007. Compositionality as an empirical problem. Chris Barker and Pauline Jacobson (eds.), Direct Compositionality, 23–101. Oxford: Oxford University Press.

EISNER, JASON. 1996. Efficient normal-form parsing for Combinatory Categorial Grammar. Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics, Santa Cruz, CA, 79–86. San Francisco: Morgan Kaufmann.

FODOR, JANET DEAN and SAG, IVAN. 1982. Referential and quantificational indefinites. Linguistics and Philosophy 5 355–398.

FREGE, GOTTLOB. 1879. Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Halle: Louis Nebert. Translated with commentary and corrections as "*Begriffsschrift*, a formula language, modeled upon that of arithmetic, for pure thought," in van Heijenoort 1967:1-82.

GAZDAR, GERALD. 1981. Unbounded dependencies and coordinate structure. Linguistic Inquiry 12 155–184.

——. 1988. Applicability of indexed grammars to natural languages. Uwe Reyle and Christian Rohrer (eds.), Natural Language Parsing and Linguistic Theories, 69–94. Dordrecht: Reidel.

GEACH, PETER. 1970. A program for syntax. Synthèse 22 3–17. Reprinted as Davidson and Harman 1972:483–497.

GILDEA, DAN and HOCKENMAIER, JULIA. 2003. Identifying semantic roles using combinatory categorial grammar. Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing, 57–64. Sapporo, Japan.

GRANROTH-WILDING, MARK. 2013. Harmonic Analysis of Music using Combinatory Categorial Grammar. Ph.D. thesis, University of Edinburgh.

GREFENSTETTE, EDWARD and SADRZADEH, MEHRNOOSH. 2011. Experimental support for a categorical compositional distributional model of meaning. Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, 1394–1404. Edinburgh: ACL.

GREFENSTETTE, EDWARD, SADRZADEH, MEHRNOOSH, CLARK, STEPHEN, COEKE, BOB, and PULMAN, STEPHEN. 2011. Concrete sentence spaces for compositional distributional models of meaning. Proceedings of the 9th International Conference on Computational Semantics, 125–134. Oxford: IWCS.

GROENENDIJK, JEROEN and STOKHOF, MARTIN. 1991. Dynamic predicate logic. Linguistics and Philosophy 14 39–100.

HARRINGTON, BRIAN and CLARK, STEPHEN. 2007. Asknet: Automated semantic knowledge network. Proceedings of the 22nd National Conference on Artificial Intelligence (AAAI'07), 889–894. AAAI Press.

——. 2009. ASKNet: Creating and evaluating large scale integrated semantic networks. International Journal of Semantic Computing 2 343–364.

HASSAN, HANY, SIMA'AN, KHALIL, and WAY, ANDY. 2009. A syntactified direct translation model with linear-time decoding. Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, 1182–1191. Singapore: ACL.

HEIM, IRENE and KRATZER, ANGELIKA. 1998. Semantics in Generative Grammar. Oxford: Blackwell.

HENDRIKS, HERMAN. 1993. Studied Flexibility: Categories and Types in Syntax and Semantics. Ph.D. thesis, Universiteit van Amsterdam.

HEPPLE, MARK. 1990. The Grammar and Processing of Order and Dependency: A Categorial Approach. Ph.D. thesis, University of Edinburgh.

——. 1995. Hybrid categorial logics. Logic Journal of IGPL 3 343–355.

HINDLE, DONALD and ROOTH, MATS. 1993. Structural ambiguity and lexical relations. Computational Linguistics 19 103–120.

HOCKENMAIER, JULIA. 2003. Parsing with generative models of predicate-argument structure. Proceedings of the 41st Meeting of the Association for Computational Linguistics, Sapporo, 359–366. San Francisco: Morgan-Kaufmann.

——. 2006. Creating a CCGbank and a wide-coverage CCG lexicon for German. Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics, 505–512. Sydney: ACL.

HOCKENMAIER, JULIA and BISK, YONATAN. 2010. Normal-form parsing for Combinatory Categorial Grammars with generalized composition and type-raising. Proceedings of the 23nd International Conference on Computational Linguistics, 465–473. Beijing.

HOCKENMAIER, JULIA and STEEDMAN, MARK. 2002. Generative models for statistical parsing with Combinatory Categorial Grammar. Proceedings of the 40th Meeting of the Association for Computational Linguistics, 335–342. Philadelphia.

——. 2005. CCGbank. LDC2005T13. ISBN 1-58563-340-2.

——. 2007. CCGbank: a corpus of CCG derivations and dependency structures extracted from the Penn Treebank. Computational Linguistics 33 355–396.

HOEKSEMA, JACK and JANDA, R. 1988. Implications of process morphology for categorial grammar. Richard Oehrle, Emmon Bach, and Deirdre Wheeler (eds.), Categorial Grammars and Natural Language Structures, 199–248. Dordrecht: Reidel.

HOFFMAN, BERYL. 1995. Computational Analysis of the Syntax and Interpretation of "Free" Word-Order in Turkish. Ph.D. thesis, University of Pennsylvania. Publ. as IRCS Report 95–17. Philadelphia: University of Pennsylvania.

HOYT, FREDERICK and BALDRIDGE, JASON. 2008. A logical basis for the D combinator and normal form in CCG. Proceedings of Association for Computational Linguistics-08: HLT, 326–334. Columbus, OH: ACL.

HUYBREGTS, RINY. 1984. The weak inadequacy of context-free phrase-structure grammars. Ger de Haan, Mieke Trommelen, and Wim Zonneveld (eds.), Van Periferie naar Kern, 81–99. Dordrecht: Foris.

JACOBSON, PAULINE. 1992. The lexical entailment theory of control and the tough construction. Ivan Sag and Anna Szabolcsi (eds.), Lexical Matters, 269–300. Stanford, CA: CSLI Publications.

——. 1999. Towards a variable-free semantics. Linguistics and Philosophy 22 117–184.

——. 2002. The (dis)organization of the grammar. Linguistics and Philosophy 25 601–626.

——. 2007. Direct compositionality and variable-free semantics: The case of "Principle B" effects. Chris Barker and Pauline Jacobson (eds.), Direct Compositionality, 191–236. Oxford: Oxford University Press.

JÄGER, GERHARD. 2005. Anaphora and Type-Logical Grammar. Dordrecht: Springer.

JOSHI, ARAVIND. 1988. Tree-adjoining grammars. David Dowty, Lauri Karttunen, and Arnold Zwicky (eds.), Natural Language Parsing, 206–250. Cambridge: Cambridge University Press.

KAMP, HANS and REYLE, UWE. 1993. From Discourse to Logic. Dordrecht: Kluwer.

KANAZAWA, MAKOTO. 1998. Learnable Classes of Categorial Grammars. Stanford, CA: CSLI/folli.

KANAZAWA, MAKOTO and SALVATI, SYLVAIN. 2012. MIX is not a Tree-Adjoining Language. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 666–674. Jeju Island, Korea: ACL.

KANG, BEOM-MO. 1995. On the treatment of complex predicates in Categorial Grammar. Linguistics and Philosophy 18 61–81.

——. 2002. Categories and meanings of Korean floating quantifiers—with some reference to Japanese. Journal of East Asian Linguistics 11 485–534.

KARTSAKLIS, DIMITRI, SADRZADEH, MEHRNOOSH, and PULMAN, STEPHEN. 2013. Separating disambiguation from composition in distributional semantics. Proceedings of the Conference on Natural Language Learning, 114–123. Sofia: ACL.

KARTTUNEN, LAURI. 1989. Radical lexicalism. Mark Baltin and Anthony Kroch (eds.), Alternative Conceptions of Phrase Structure, 43–65. Chicago: University of Chicago Press.

KEMPSON, RUTH and CORMACK, ANNABEL. 1981. Ambiguity and quantification. Linguistics and Philosophy 4 259–309.

KOBELE, GREGORY and KRACHT, MARCUS. 2005. On pregroups, freedom, and (virtual) conceptual necessity. Proceedings of the 28th Annual Penn Linguistics Colloquium, University of Pennsylvania Working Papers in Linguistics. Philadelphia.

KOLLER, ALEXANDER and THATER, STEFAN. 2006. An improved redundancy elimination algorithm for underspecified descriptions. Proceedings of the International Conference on Computational Linguistics/Association for Computational Linguistics, 409–416. Sydney: Coling/ACL.

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

KÖNIG, ESTHER. 1994. A hypothetical reasoning algorithm for linguistic analysis. Journal of Logic and Computation 4 1–19.

KRATZER, ANGELIKA. 1998. Scope or pseudo-scope: Are there wide-scope indefinites? Susan Rothstein (ed.), Events in Grammar, 163–196. Dordrecht: Kluwer.

KRISHNAMURTHY, JAYANT and MITCHELL, TOM. 2012. Weakly supervised training of semantic parsers. Proceedings of Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, 754–765. Jeju Island, Korea: ACL.

KRUIJFF-KORBAYOVÁ, IVANA, ERICSSON, STINA, RODRÍGUEZ, KEPA JOSEBA, and KARAGRJOSOVA, ELENA. 2003. Producing contextually appropriate intonation in an information-state based dialogue system. Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 227–234. Budapest, Hungary: ACL.

KUBOTA, YUSUKE. 2010. (In)flexibility of Constituency in Japanese in Multi-Modal Categorial Grammar with Structured Phonology. Ph.D. thesis, Ohio State University.

KWIATKOWSKI, TOM, GOLDWATER, SHARON, ZETTLEMOYER, LUKE, and STEEDMAN, MARK. 2012. A probabilistic model of syntactic and semantic acquisition from child-directed utterances and their meanings. Proceedings of the 13th Conference of the European Chapter of the ACL (EACL 2012), 234–244. Avignon: ACL.

KWIATKOWSKI, TOM, ZETTLEMOYER, LUKE, GOLDWATER, SHARON, and STEEDMAN, MARK. 2010. Inducing probabilistic CCG grammars from logical form with higher-order unification. Proceedings of the Conference on Empirical Methods in Natural Language Processing, 1223–1233. Cambridge, MA: ACL.

——. 2011. Lexical generalization in CCG grammar induction for semantic parsing. Proceedings of the Conference on Empirical Methods in Natural Language Processing, 1512–1523. Edinburgh: ACL.

LAKOFF, GEORGE. 1970. Linguistics and natural logic. Synthése 22 151–271. Reprinted in Davidson and Harman 1972:545–665.

LAMBEK, JOACHIM. 1958. The mathematics of sentence structure. American Mathematical Monthly 65 154–170.

——. 1961. On the calculus of syntactic types. Roman Jakobson (ed.), Structure of Language and Its Mathematical Aspects, *Proceedings of the Symposium in Applied Mathematics*, vol. 12, 166–178. Providence, RI: American Mathematical Society.

——. 2001. Type grammars as pregroups. Grammars 4 21–39.

——. 2008. From Word to Sentence. Milan: Polimetrica.

LEE, JUNGMEE and TONHAUSER, JUDITH. 2010. Temporal interpretation without tense: Korean and Japanese coordination constructions. Journal of Semantics 27 307–341.

LEWIS, MICHAEL and STEEDMAN, MARK. 2013. Combined distributional and logical semantics. Transactions of the Association for Computational Linguistics 1 179–192.

LYONS, JOHN. 1968. Introduction to Theoretical Linguistics. Cambridge: Cambridge University Press.

MACCARTNEY, BILL and MANNING, CHRISTOPHER D. 2007. Natural logic for textual inference. Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing, 193–200. Prague: ACL.

MAGERMAN, DAVID. 1995. Statistical decision tree models for parsing. Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics, 276–283. Cambridge, MA: ACL.

MCCONVILLE, MARK. 2006. Inheritance and the CCG lexicon. Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics, 1–8. ACL.

MEHAY, DENIS and BREW, CHRIS. 2012. Ccg syntactic reordering models for phrase-based machine translation. Proceedings of the Seventh Workshop on Statistical Machine Translation, XXX–XXX. Montréal: ACL.

MITCHELL, JEFF and LAPATA, MIRELLA. 2008. Vector-based models of semantic composition. Proceedings of the Annual Meeting of the Association for Computational Linguistics, 236–244. Columbus, OH: ACL.

MONTAGUE, RICHARD. 1970a. English as a formal language. Bruno Visentini (ed.), Linguaggi nella Società e nella Technica, 189–224. Milan: Edizioni di Communità. Reprinted as Thomason 1974:188-221.

——. 1970b. Universal grammar. Theoria 36 373–398. Reprinted as Thomason 1974:222-246.

——. 1973. The proper treatment of quantification in ordinary English. Jaakko Hintikka, J. M. E. Moravcsik, and Patrick Suppes (eds.), Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics, 221–242. Dordrecht: Reidel. Reprinted as Thomason 1974:247-279.

MOORTGAT, MICHAEL. 1988a. Categorial Investigations. Ph.D. thesis, Universiteit van Amsterdam. Published by Foris, Dordrecht, 1989.

——. 1988b. Mixed composition and discontinuous dependencies. Richard Oehrle, Emmon Bach, and Deirdre Wheeler (eds.), Categorial Grammars and Natural Language Structures, 319–348. Dordrecht: Reidel.

——. 1997. Categorial type logics. Johan van Benthem and Alice ter Meulen (eds.), Handbook of Logic and Language, 93–177. Amsterdam: North Holland.

——. 2007. Symmetries in natural language syntax and semantics: The Lambek-Grishin calculus. Proceedings of Workshop on Logic, Language, Information, and Computation (WoLLIC), Lexture Notes in Computer Science 4576, 264–284. Berlin: Springer-Verlag.

——. 2009. Symmetric categorial grammar. Journal of Philosophical Logic 38 681–710.

MOOT, RICHARD. 2002. Proof Nets for Linguistic Analysis. Ph.D. thesis, University of Utrecht.

——. 2010. Automated extraction of type-logical supertags from the spoken Dutch corpus. Srinivas Bangalore and Aravind Joshi (eds.), Complexity of Lexical Descriptions and its Relevance to Natural Language Processing: A Supertagging Approach. Cambridge, MA: MIT Press.

MOOT, RICHARD and PUITE, QUINTIJN. 2002. Proof nets for the multimodal Lambek calculus. Studia Logica 71 415–442.

MOOT, RICHARD and RETORÉ, CHRISTIAN. 2012. The Logic of Categorial Grammars. No. 6850 in Lecture Notes in Computer Science. Berlin: Springer.

MORRILL, GLYN. 1994. Type-Logical Grammar. Dordrecht: Kluwer.

——. 2000. Incremental processing and acceptability. Computational Linguistics 26 319–338.

MORRILL, GLYNN. 2011. Categorial Grammar: Logical Syntax, Semantics, and Processing. Oxford: Oxford University Press.

MORRILL, GLYNN and SOLIAS, TERESA. 1993. Tuples, discontinuity, and gapping. Proceedings of 6th Conference of the European Chapter of the Association for Computational Linguistics, 287–297. Utrecht: ACL.

MUSKENS, REINHARD. 2007. Separating syntax and combinatorics in Categorial Grammar. Research on Language & Computation 5 267–285.

NISHIDA, CHIYO. 1996. Second position clitics in Old Spanish and Categorial Grammar. Aaron Halpern and Arnold Zwicky (eds.), Approaching Second: Second-Position Clitics and Related Phenomena, 33–373. Stanford, CA: CSLI Publications.

OEHRLE, RICHARD. 1988. Multidimensional compositional functions as a basis for grammatical analysis. Richard Oehrle, Emmon Bach, and Deirdre Wheeler (eds.), Categorial Grammars and Natural Language Structures, 349–390. Dordrecht: Reidel.

——. 1994. Term-labelled categorial type systems. Linguistics and Philosophy 17 633–678.

OEHRLE, RICHARD, BACH, EMMON, and WHEELER, DEIRDRE (eds.). 1988. Categorial Grammars and Natural Language Structures. Dordrecht: Reidel.

PARK, JONG and CHO, HYUNG-JOON. 2000. Informed parsing for coordination with combinatory categorial grammar. Proceedings of the 18th International Conference on Computational Linguistics, 593–599. Saarbrücken: Coling/ACL.

PARTEE, BARBARA. 1970. Negation, conjunction, and quantifiers: Syntax vs. semantics. Foundations of Language 6 153–165.

——. 1975. Montague Grammar and Transformational Grammar. Linguistic Inquiry 6 203–300.

PARTEE, BARBARA (ed.). 1976. Montague Grammar. New York: Academic Press.

PARTEE, BARBARA and ROOTH, MATS. 1983. Generalised conjunction and type ambiguity. Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow (eds.), Meaning, Use, and Interpretation of Language, 361–383. Berlin: de Gruyter.

PENTUS, MATI. 1993. Lambek grammars are context-free. Proceedings of the IEEE Symposium on Logic in Computer Science, Montreal, 429–433.

——. 2003. Lambek calculus is NP-complete. Tech. Rep. TR-2003005, Graduate Center, City University of New York, New York.

PEREIRA, FERNANDO. 1990. Categorial semantics and scoping. Computational Linguistics 16 1–10.

PEREIRA, FERNANDO and SHIEBER, STUART. 1987. Prolog and Natural Language Analysis. Stanford, CA: CSLI Publications.

PIANTADOSI, STEVEN, GOODMAN, NOAH, ELLIS, BENJAMIN, and TENENBAUM, JOSHUA. 2008. A Bayesian model of the acquisition of compositional semantics. Proceedings of the 30th Annual Meeting of the Cognitive Science Society, 1620–1625. Washington DC.

PICKERING, MARTIN and BARRY, GUY. 1993. Dependency categorial grammar and coordination. Linguistics 31 855–902.

POESIO, MASSIMO. 1995. Disambiguation as (defeasible) reasoning about underspecified representations. Papers from the Tenth Amsterdam Colloquium. Amsterdam: ILLC, Universiteit van Amsterdam.

POLLARD, CARL. 1984. Generalized Context Free Grammars, Head Grammars, and Natural Languages. Ph.D. thesis, Stanford University.

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

RANTA, AARNE. 1994. Type-Theoretical Grammar. Oxford: Oxford University Press.

REYLE, UWE. 1993. Dealing with ambiguities by underspecification. Journal of Semantics 10 123–179.

ROBINSON, ABRAHAM. 1974. Introduction to Model Theory and to the Metamathematics of Algebra. Amsterdam: North Holland, 2nd ed. edn.

ROSS, JOHN ROBERT. 1970. Gapping and the order of constituents. Manfred Bierwisch and Karl Heidolph (eds.), Progress in Linguistics, 249–259. The Hague: Mouton.

RUANGRAJITPAKORN, TANETH, TRAKULTAWEEKOON, KANOKORN, and SUPNITHI, THEPCHAI. 2009. A syntactic resource for Thai: CG treebank. Proceedings of the 7th Workshop on Asian Language Resources, 96–102. Suntec, Singapore: ACL.

SÁNCHEZ VALENCIA, VÍCTOR. 1991. Studies on Natural Logic and Categorial Grammar. Ph.D. thesis, Universiteit van Amsterdam.

——. 1995. Parsing-driven inference: Natural logic. Linguistic Analysis 25 258–285.

SAVITCH, WALTER. 1987. Context-sensitive grammar and natural language syntax. Walter Savitch, Emmon Bach, William Marsh, and Gila Safran-Naveh (eds.), The Formal Complexity of Natural Language, 358–368. Dordrecht: Reidel.

SCHLENKER, PHILIPPE. 2006. Scopal independence: A note on branching and wide scope readings of indefinites and disjunctions. Journal of Semantics 23 281–314.

SEKI, HIROYUKI, MATSUMURA, TAKASHI, FUJII, MAMORU, and KASAMI, TADAO. 1991. On multiple context-free grammars. Theoretical Computer Science 88 191–229.

SHAN, CHUNG-CHIEH and BARKER, CHRIS. 2006. Explaining crossover and superiority as left-to-right evaluation. Linguistics and Philosophy 29 91–134.

SHIEBER, STUART. 1985. Evidence against the context-freeness of natural language. Linguistics and Philosophy 8 333–343.

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

SMITH, NOAH and JOHNSON, MARK. 2007. Weighted and probabilistic context-free grammars are equally expressive. Computational Linguistics 33 477–491.

SRINIVAS, BANGALORE. 1997. Complexity of Lexical Descriptions and Its Relevance to Partial Parsing. Ph.D. thesis, University of Pennsylvania, Philadelphia. Published as IRCS Report 97-10.

STABLER, EDWARD. 2004a. Tupled pregroup grammars. Claudia Casadio and Joachim Lambek (eds.), Computational Algebraic Approaches to Morphology and Syntax, 23–52. Milan: Polimetrica.

——. 2004b. Varieties of crossing dependencies: Structure-dependence and mild context sensitivity. Cognitive Science 28 699–720.

——. 2011. Computational perspectives on minimalism. Cedric Boeckx (ed.), Oxford Handbook of Linguistic Minimalism, 617–641. Oxford University Press.

STABLER, EDWARD and KEENAN, EDWARD. 2003. Structural similarity. Theoretical Computer Science 293 345–363.

STEEDMAN, MARK. 1985. Dependency and coordination in the grammar of Dutch and English. Language 61 523–568.

——. 1987. Combinatory grammars and parasitic gaps. Natural Language and Linguistic Theory 5 403–439.

——. 1990. Gapping as constituent coordination. Linguistics and Philosophy 13 207–263.

——. 1996. Surface Structure and Interpretation. Linguistic Inquiry Monograph 30. Cambridge, MA: MIT Press.

——. 1999. Quantifier scope alternation in CCG. Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics, 301–308. College Park, MD: ACL.

——. 2000a. Information structure and the syntax-phonology interface. Linguistic Inquiry 34 649–689.

——. 2000b. The Syntactic Process. Cambridge, MA: MIT Press.

——. 2012. Taking Scope: The Natural Semantics of Quantifiers. Cambridge, MA: MIT Press.

STEEDMAN, MARK and BALDRIDGE, JASON. 2011. Combinatory categorial grammar. Robert Borsley and Kirsti Börjars (eds.), Non-Transformational Syntax: A Guide to Current Models, 181–224. Oxford: Blackwell.

STEELE, SUSAN. 1990. Agreement and Anti-Agreement: A Syntax of Luiseño. Dordrecht: Reidel.

SZABOLCSI, ANNA. 1983. Ecp in categorial grammar. Ms., Max Planck Institute, Nijmegen.

——. 1989. Bound variables in syntax: Are there any? Renate Bartsch, Johan van Benthem, and Peter van Emde Boas (eds.), Semantics and Contextual Expression, 295–318. Dordrecht: Foris.

——. 1992. On combinatory grammar and projection from the lexicon. Ivan Sag and Anna Szabolcsi (eds.), Lexical Matters, 241–268. Stanford, CA: CSLI Publications.

SZABOLCSI, ANNA (ed.). 1997. Ways of Scope-Taking. Dordrecht: Kluwer.

THOMASON, RICHMOND (ed.). 1974. Formal Philosophy: Papers of Richard Montague. New Haven, CT: Yale University Press.

THOMFORDE, EMILY. 2013. Semi-Supervised Lexical Acquisition for Wide-Coverage Parsing. Ph.D. thesis, University of Edinburgh.

TRECHSEL, FRANK. 2000. A CCG account of Tzotzil Pied Piping. Natural Language and Linguistic Theory 18 611–663.

TSE, DANIEL and CURRAN, JAMES R. 2010. Chinese CCGbank: Extracting CCG derivations from the Penn Chinese Treebank. Proceedings of the 23rd International Conference on Computational Linguistics, 1083–1091. Beijing: Coling/ACL.

USZKOREIT, HANS. 1986. Categorial unification grammars. Proceedings of the International Conference on Computational Linguistics, 187–194. Bonn: Coling/ACL.

VAN BENTHEM, JOHAN. 1983. Five easy pieces. Alice ter Meulen (ed.), Studies in Model-Theoretic Semantics, 1–17. Dordrecht: Foris.

——. 1986. Essays in Logical Semantics. Dordrecht: Reidel.

——. 1988. The semantics of variety in Categorial Grammar. Wojciech Buszkowski, Witold Marciszewski, and Johan van Benthem (eds.), Categorial Grammar, 37–55. Amsterdam: John Benjamins.

VAN HEIJENOORT, JEAN (ed.). 1967. From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931. Cambridge, MA: Harvard University Press.

VIJAY-SHANKER, K. and WEIR, DAVID. 1994. The equivalence of four extensions of context-free grammar. Mathematical Systems Theory 27 511–546.

VILLAVICENCIO, ALINE. 2002. The Acquisition of a Unification-Based Generalised Categorial Grammar. Ph.D. thesis, University of Cambridge.

——. 2011. Language acquisition with feature-based grammars. Robert Boyer and Kirsti Börjars (eds.), Non-Transformational Syntax: A Guide to Current Models, 404–442. Blackwell.

WATKINSON, STEPHEN and MANANDHAR, SURESH. 1999. Unsupervised lexical learning with categorial grammars. Proceedings of the Workshop on Unsupervised Learning in Natural Language Processing, ACL-99, College Park MA, 59–66. Association for Computational Linguistics.

WEBBER, BONNIE. 1978. A Formal Approach to Discourse Anaphora. Ph.D. thesis, Harvard University. Published by Garland, New York, 1979.

WEIR, DAVID. 1988. Characterizing Mildly Context-sensitive Grammar Formalisms. Ph.D. thesis, University of Pennsylvania, Philadelphia. Published as Technical Report CIS-88-74.

WHITE, MICHAEL. 2006. Efficient realization of coordinate structures in combinatory categorial grammar. Research on Language and Computation 4 39–75.

WHITELOCK, PETE. 1991. What sort of trees do we speak?: A computational model of the syntax-prosody interface in Tokyo Japanese. Proceedings of the Fifth Conference of the European Chapter of the Association for Computational Linguistics, 75–82. Berlin, Germany: ACL.

WITTENBURG, KENT. 1987. Predictive combinators: A method for efficient processing of combinatory grammars. Proceedings of the 25th Annual Conference of the Association for Computational Linguistics, 73–80. Stanford: ACL.

WOOD, MARY MCGEE. 1993. Categorial Grammars. London: Routledge.

ZAMANSKY, ANNA, FRANCEZ, NISSIM, and WINTER, YOAD. 2006. A “natural logic” inference system using the Lambek calculus. Journal of Logic, Language, and Information 15 273–295.

ZEEVAT, HENK. 1988. Combining Categorial Grammar and Unification. Uwe Reyle and Christian Rohrer (eds.), Natural Language Parsing and Linguistic Theories, 202–229. Dordrecht: Reidel.

ZETTLEMOYER, LUKE and COLLINS, MICHAEL. 2005. Learning to map sentences to logical form: Structured classification with Probabilistic Categorial Grammars. Proceedings of the 21st Conference on Uncertainty in AI (UAI), 658–666. Edinburgh: AAAI.

——. 2007. Online learning of relaxed CCG grammars for parsing to logical form. Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP/CoNLL), 678–687. Prague: ACL.

ZWICKY, ARNOLD. 1986. Concatenation and liberation. CLS 22: Papers from the General Session, 65–74. Chicago Linguistic Society, Chicago: University of Chicago.