CATEGORIAL GRAMMARS AND NATURAL LANGUAGE STRUCTURES

Edited by

RICHARD T. OEHRLE
Dept. of Linguistics, University of Arizona

EMMON BACH
Dept. of Linguistics, University of Massachusetts at Amherst

and

DEIRDRE WHEELER
English Dept., University of New Hampshire

D. REIDEL PUBLISHING COMPANY

A MEMBER OF THE KLUWER ACADEMIC PUBLISHERS GROUP

DORDRECHT / BOSTON / LANCASTER / TOKYO

Library of Congress Cataloging-in-Publication Data

CIP

Categorial grammars and natural language structures.

(Studies in linguistics and philosophy ; v. 32)
Bibliography: p.
Includes indexes.
1. Categorial grammar. I. Oehrle, Richard T.
II. Bach. Emmon W., 1929- . III. Wheeler, Deirdre.
IV. Series.
P161.C39 1988 415 88-3052

ISBN 1-55608-030-1

Published by D. Reidel Publishing Company,
P.O. Box 17, 3300 AA Dordrecht, Holland.

Sold and distributed in the U.S.A. and Canada
by Kluwer Academic Publishers,
101 Philip Drive, Norwell, MA 02061, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers Group,
P.O. Box 322, 3300 AH Dordrecht, Holland.

Printed in The Netherlands

CATEGORIAL GRAMMARS
AND NATURAL LANGUAGE STRUCTURES

STUDIES IN LINGUISTICS AND PHILOSOPHY

formerly *Synthese Language Library*

VOLUME 32

EDITORIAL PREFACE

For the most part, the papers collected in this volume stem from presentations given at a conference held in Tucson over the weekend of May 31 through June 2, 1985. We wish to record our gratitude to the participants in that conference, as well as to the National Science Foundation (Grant No. BNS-8418916) and the University of Arizona SBS Research Institute for their financial support. The advice we received from Susan Steele on organizational matters proved invaluable and had many felicitous consequences for the success of the conference. We also would like to thank the staff of the Departments of Linguistics of the University of Arizona and the University of Massachusetts at Amherst for their help, as well as a number of individuals, including Lin Hall, Kathy Todd, and Jiazhen Hu, Sandra Fulmer, Maria Sandoval, Natsuko Tsujimura, Stuart Davis, Mark Lewis, Robin Schafer, Shi Zhang, Olivia Oehrle-Steele, and Paul Saka. Finally, we would like to express our gratitude to Martin Scrivener, our editor, for his patience and his encouragement.

TABLE OF CONTENTS

INTRODUCTION

The term 'categorial grammar' was introduced by Bar-Hillel (1964, page 99) as a handy way of grouping together some of his own earlier work (1953) and the work of the Polish logicians and philosophers Leśniewski (1929) and Ajdukiewicz (1935), in contrast to approaches to linguistic analysis based on phrase structure grammars. The most accessible of these earlier works was the paper of Ajdukiewicz, who, under the influence of Husserl's *Bedeutungskategorien* and the type theory that Russell had introduced to fend off foundational problems in set theory, proposed a mode of grammatical analysis in which every element of the vocabulary of a language belongs to one or more categories, and each category is either basic or defined in terms of simpler categories in a way which fixes the combinatorial properties of complex categories.

As an example, consider the language described below:

Basic Categories: s

Recursive Definition of the Full Set of Categories: The set CAT is the smallest set such that (1) if A is a basic category, then A belongs to CAT, and (2) if A and B belong to CAT, then A/B belongs to CAT.

Basic expressions ('$B(A)$' denotes the set of basic expressions of category A):

$$B(s) = \{p_i \mid i \in \mathbb{N}\}$$
$$B(s/s) = \{\sim\}$$
$$B((s/s)/s) = \{\wedge\}$$

Simultaneous recursive definition of the full set of expressions ('$P(A)$' denotes the full set of expressions of category A): if A is a category and x is a member of $B(A)$, then x is a member of $P(A)$; if x is a member of $P(A)$ and y is a member of $P(C/A)$, for some category C, then yx is a member of $P(C)$.

Richard T. Oehrle et al. (eds.), Categorial Grammars and Natural Language Structures, 1—16.

While the language defined by these definitions is rich enough to develop a formulation of the propositional calculus, and makes use of the 'Polish' notation introduced by Łukasiewicz, what is most important from a grammatical standpoint is that expressions belonging to complex categories of the form *A/C*, where *A* and *C* are categories that are either simple or derived, may be identified with functions which map expressions of category *C* into the set of expressions of category *A*. In fact, it is the systematic use of functions to characterize the compositional properties of grammatical expressions which is most symptomatic of the modern versions of 'categorial grammars' whose investigation is the purpose of this volume. And in this respect, categorial grammars hark back directly to ideas first introduced by Frege in his *Begriffsschrift* (cf. especially §9, pp. 21—23 of the translation in van Heijenoort, 1971).

I

The role that functions play in categorial grammars confers on them a number of properties that make them a focus of interest to linguists, philosophers, logicians, and mathematicians. One way to see this is to consider some of the ways of extending the simple grammar given above. First of all, while it is possible to construct categorial grammars which are purely syntactical systems, it is extremely useful to consider functions and arguments which have non-syntactic, as well as syntactic, properties. The crucial aspect of the enormous impact of Montague's work, for example, is perhaps the deep relation between syntactic properties and semantic properties in the systems he constructed. Other recent work by Bach, Schmerling, and Wheeler has emphasized the applicability of the basic perspective of categorial grammar to the analysis of phonological properties and the phonological composition of expressions. Extended in this way, then, categorial grammars offer a way of studying the composition of grammatical expressions across a variety of phenomenally-accessible domains.

Second, it is possible to have a richer inventory of basic types. While the grammar above has only one basic type, the system discussed by Ajdukiewicz consists of two types, Bar-Hillel's formulation of bidirectional category systems allows any finite number of basic categories, Lewis (1972) suggests a system with three basic categories (corresponding to sentence, name, and common noun), and the intensional logic of Montague's PTQ (Paper 8 in Thomason, 1974a) is based on the three types *s*, *e*, and *t*.

A third way in which categorial grammars can be enriched is to allow the composition of expressions in ways not directly definable in terms of concatenation. This possibility is explicitly recognized in the writings of Curry (1961) ("A functor is *any* kind of linguistic device which operates on one or more phrases (the argument(s)) to form another phrase" [1961, p. 62]). Lewis (1972), following a suggestion of Lyons (1966) (though with possibly different motives), proposes a categorially-based transformational grammar — that is, a transformational grammar whose base component is a categorial grammar rather than a phrase structure grammar. The syntactic rules in Montague's fragments make it abundantly clear that he felt under no compulsion to restrict syntactic operations to concatenative ones — in fact, he apparently felt no need at all to explicate the notion of possible syntactic operation. One goal of more recent work, typified by an important series of papers by Bach (1979, 1981, 1984), has been to study the properties of categorial systems with basic operations which are not all characterizable in terms of the concatenation of their operands.

Since categorial grammars are based on the algebraic notions *function* and *argument*, there is another way in which extended categorial grammars arise: namely, by exploiting certain natural relations among functions and their arguments. The exploration of these ideas goes back at least to the discovery of the basic principles of combinatory logic by Schönfinkel and Curry in the 1920s. But the first grammatical system in which they play a fundamental role is the associative syntactic calculus of Lambek (1958: cf. also 1961). This system allows operations corresponding to functional composition, type-lifting, and a variety of other rules of type-changing. Rules of functional composition and higher-order types for conjoined terms (as well as for expressions corresponding to quantifiers) can be found in Geach (1972), and a higher order type for all NP's is one of the most characteristic traits of Montague's PTQ. Levin (1976, 1982) also seeks to enrich simple categorial grammars of the type found in the work of Ajdukiewicz and Bar-Hillel with combinatorial operations not restricted to a single type, but raises the problem of which of the various possible operations are desirable in different contexts.

The mathematical rationale of these relations is unlikely to be familiar to those with no background in modern logic or modern algebra. But it is easy to grasp and important in many of the papers which follow. In the next section, we provide enough of an informal exposition of the nature of functions to justify the equivalences among functions appealed to above, such as functional composition and type-lifting. In the following sections, we show how these ideas are relevant to the papers contained in this volume, and how the research reported here bears on current issues in linguistic theory.

II

The terms *function* and *argument* have been used for a long time in linguistics, with various looser and tighter senses. In mathematics and logic, they have an even longer history and in general a much more precise sense. As the above discussion makes clear, categorial grammar can be seen as the result of taking the mathematical notions seriously in the analysis of language structures.

In mathematical contexts, there are a variety of ways in which the notion of a function has been explicated (cf. MacLane, 1986, pp. 126—127). Two of these stand out. In one, a function is thought of as a procedure, operation, or computer program, with a well-defined domain (or ordered set of n domains) of possible arguments, which yields a unique result when it is given an argument (or n arguments) in the domain(s). In the other, a function is thought of more statically as simply a relation between sets (more generally, n sets) such that given any argument (or sequence of arguments) in the first $(n - 1)$ set(s), there is a unique member of the second (or nth) set in the relation. It is necessary to distinguish between *total* functions, where every element in the domain(s) yields a result when the function is applied to it, and *partial* functions, where this is not the case. Easy examples from arithmetic are the functions Square (total) and Divide, a two-argument function, which is a partial function when its domain is the set of all rational numbers, since such things as 'Divide 3 by 0' are illicit. In the latter case, we say that the result is 'undefined'. A partial function from A to a set B can always be extended to a total function from A to $B+$, where $B+$ is gotten from B by adding some new element ('zilch', say) to B and letting that be the element to which the members of A for which the function is undefined correspond. (A familiar kind of example here is the extension of a partial truth function for formulas in a two-valued logic to a total function from the formulas to a set $\{0, 1, \text{zilch}\}$.)

For a long time in mathematics, the concept of a function was used but had no nice representation. This was remedied by Church in the 1930s with the introduction of the lambda operator to create names for functions. Since this notation is fundamental to understanding many of the papers here, let us review it briefly. For a function from a *domain* A to a *range* (or *codomain*) B, we write

> $\lambda x[P]$, where x ranges over elements in A and P denotes some element in B.

This expression, relative to some assignment g of values to variables, denotes that function f such that for any element a in A, $f(a)$ is the denotation of P in B relative to an assignment g' which is exactly like g, except (possibly) that $g'(x) = a$. The reason for this complicated clause is that P may contain free variables other than x. Notice, by the way, that P may lack free occurrences of x, in which case the lambda expression will be a constant function yielding the denotation of P no matter what argument we feed it.

Students of Montague grammar will be well acquainted with this definition. This notation is probably most often used for giving names to sets by writing the name of a characteristic function from some set (or Cartesian product of several sets), that is, for so-called 'set abstraction'. (A characteristic function is a function into the set of truth values.) However, the lambda notation is completely general: the codomain itself may be, for example, a set of functions. Easy examples are the arithmetical functions mentioned above, which we may write thus:

$$\text{Square} = \lambda x[x, x]$$
$$\text{Divide} = \lambda x[\lambda y[y/x]]$$

A number of the papers in this volume make use of several results about functions. We will briefly review these ideas here. (In the discussion below, we use the exponential notation 'X^Y' to stand for the set of all functions with domain Y and co-domain X.)

(1) *Currying*: An early result by Schönfinkel (1924) and, independently, by Curry, showed that it is always possible to take an n-ary function and break it down into a series of a 1-ary functions. We have already used this idea in the definition of Divide just given, that is, it is given not as a two-place function but rather as a one-place function which will yield for each argument a another function which we get by setting $x = a$ (thus, Divide (2) is $\lambda y[y/2]$, which applied to 5 yields

5/2). Conversely, any complex function of the curried sort can be 'decurried'. The algebraic basis of this result is the fact that there is an isomorphism between the two function sets:

$$S^{YXZ} \cong (S^Y)^Z$$

given by associating a function f in the first set with a function F in the second just in case $f(y, z) = [F(z)](y)$. Algebraically, then, while f and F are distinguishable, it is both convenient and possible in most contexts to identify them. (The same argument extends easily to n-ary functions with $n > 2$ — take the set Y above to be an $n - 1$ fold Cartesian product.) The linguistic consequence of this result is that theories based on the theory of functions can make use either of many-place or one-place functions (analogous to ideas about 'flat' versus 'hierarchical' structures in phrase-structure grammars).

(2) *Function composition*: Suppose we have a function f from A to B and a function g from B to C. Then there always exists a function h from A to C, where

$$h = \lambda x[g(f(x))], \text{ with } h \text{ usually written as } g \circ f$$

(Strictly, this is true only for total functions: in the case of partial functions, if we want the composite to be a total function, we have to restrict ourselves to cases where the value of f on an argument is defined and within the subdomain of B for which g is also defined.)

(3) *Type-lifting*: Suppose we have an element a in the domain of a function f. Then we can always identify a with a higher-order function g whose domain contains f in such a way that the following equality holds:

$$f(a) = g(f)$$

If a is a member of some set A and f belongs to the function set C^A, this identity characterizes an injection (one-to-one map) from A to the function set

$$C^{(C^A)}$$

Thus, while a and g are distinguishable, it is often convenient to identify them. A well-known example is Montague's treatment of noun-phrases as functions from properties to truth-values (leading to the very fruitful investigations in recent years of generalized quantifiers in natural language).

(4) *Interchange*: Suppose we have a function in the function set

$$(C^B)^A$$

This function can be identified in a natural way with a function in the function set

$$(C^A)^B$$

The requisite correspondence is easily expressed in the lambda notation:

$$\lambda x[\lambda y[f(x, y)]] \rightarrow \lambda y[\lambda x[f(x, y)]]$$

And it is easy to see that this mapping characterizes an isomorphism between the two function sets. As in the cases discussed earlier, this mapping allows us to identify two distinct functions in a way which often proves convenient. Related to this fact in semantically-interpreted directional categorial grammars is the natural identification of a functor f of type $(C/B)/A$, interpreted as f', with that functor g of type $(C/A)/B$, interpreted as g' such that for all arguments a in A and b in B, interpreted as a' and b', respectively, $[f'(a')](b') = [g'(b')](a')$. In this case, however, if f and g share a common phenogrammatical shape — say V —, under the standard interpretation of directional categories, we have two distinct forms — V*ab* and V*ba* — associated with a single interpretation. Equivalences of this type which result in the permutation of arguments have obvious syntactic application.

On the other hand, here is a straightforward example from semantics: Montague defines a *proposition* as a function from indices (possible worlds cross times) to truth-values, and a *property* as a function from indices to sets, or rather characteristic functions of sets (see above). So if we let s, e, t be the types for indices, individuals, and truth-values, respectively, we have the types for propositions as $\langle s, t \rangle$, and for properties as $\langle s, \langle e, t \rangle\rangle$. Now given Interchange, we can just as well let properties be functions from individuals to propositions, that is, of type $\langle e, \langle s, t \rangle\rangle$, which will be completely equivalent. This alternative characterization is found in Chierchia (1984), as well as in earlier work by Cresswell (1972) and von Stechow (1974).

III

Even a relatively quick review of the growth of the various strands of 'categorial grammar' found in current work reveals a remarkable historical depth. The work of Ajdukiewicz perhaps represents the earliest attempt to construct a generative grammar of a fragment of a natural language. While this tradition was kept alive within philosophy, because of the possibility it affords of treating the compositional properties of form and interpretation in a common framework, it has been slower to take root in linguistics. It is probably correct to say that the lack of interest in categorial grammar among linguists derived from (1) the early results of Bar-Hillel and his colleagues concerning the equivalence of bidirectional categorial grammars (with a single operation corresponding to functional application) and context-free grammars, together with the then widely-accepted view that natural languages could literally not be described by context-free grammars, as well as from (2) the unfamiliarity of most linguists with model-theoretic approaches to interpretation — approaches which make available the possibility of detailed, explicit, and tractable explorations of the relation between syntax and interpretation.

Montague's papers, particularly PTQ and UG (Paper 7 in Thomason, 1974a), provided a framework in which semantic problems could be (and most certainly have been) profitably pursued. More importantly, perhaps, the emphasis that Montague placed upon the desirability of a homomorphic relation between the properties and composition of syntactic types and the properties and composition of corresponding semantic types exposed to a wider audience the virtues of the categorial grammars he employed. At the same time, the growing sense that transformational analyses were not always appropriate led to the exploration of various non-transformational approaches to grammatical analysis. We suspect that these two converging tendencies were not entirely independent. (For example, R. Thomason's (1974b, 1976) demonstration that coherent alternatives to purely syntactic analyses of certain central grammatical problems existed in a Montague-grammar framework had an impact that has perhaps never been adequately acknowledged.)

The research presented in this volume, much of it interconnected, shows that work in categorial grammar in the broad sense has moved beyond the stage of sporadic rediscovery and reached a critical mass. Broadly construed, categorial grammar allows a much more thoroughgoing investigation of the foundations of grammatical composition than competing frameworks. The papers that follow make it clear that the attractions of this kind of investigation have already yielded results of interest. And these results in turn imply that there is a great deal more that is within reach.

We turn now to some brief expository remarks concerning the individual papers.

Bach's paper may be read as a general introduction to the use of categorial and near-categorial systems in formulating empirical theories about the structures of natural languages.

Van Benthem's paper explores the properties of the commutative Lambek calculus — a grammatical system which contains expressions of simple type and 1-ary functors which may combine with their arguments in either order, together with certain type-changing rules found in Lambek's (1958) paper. Because of the role that permutation plays in this system, it has interesting properties syntactically — both with respect to other categorial systems, such as the Ajdukiewicz/Bar-Hillel directional system and the directional Lambek calculi, and with respect to the Chomsky hierarchy of re-writing systems. The semantics for this system that van Benthem defines has affinities with the standard interpretation of the intuitionistic logic of implication. The intertwining of logical, mathematical, and linguistic themes in this paper reveals the way in which categorial grammars are especially suited to work along the boundaries between these disciplines.

In 1959, Gaifman obtained a proof of the equivalence of bidirectional categorial grammars and context-free phrase structure grammars (Bar-Hillel 1964, p. 103). As mentioned above, this result led to the general belief in linguistic circles that categorial grammars offered nothing beyond the context-free phrase structure grammars which were generally held to be empirically inadequate. In the intervening years, little has been done to dispel this illusion, in spite of the fact that the categorial grammars introduced by Lambek, Geach, and others go beyond the syntactic devices found in the grammars involved in Gaifman's proof. Recently, however, there has been a revival of interest in the mathematical properties of various categorial grammar, beginning (fittingly enough) with work by Polish logicians. Buszkowski has been at the forefront of this revival, and in his paper here, he reviews these recent results concerning the generative capacity of different types of categorial grammars, particularly the bi-directional grammars of Bar-Hillel (following Ajdukiewicz) and the product-free, associative syntactic calculus of Lambek (1958), and other related systems. These