



Discontinuous Constituents in Generalized Categorical Grammars

Item Type	article;article
Authors	Bach, Emmon W.
Download date	2024-08-10 06:23:52
Link to Item	https://hdl.handle.net/20.500.14394/36347

DISCONTINUOUS CONSTITUENTS IN GENERALIZED CATEGORIAL GRAMMARS

EMMON W. BACH

UNIVERSITY OF MASSACHUSETTS/AMHERST

0. Introduction. Categorical grammars were invented by K. Ajdukiewicz (1935), on the basis of ideas of Lesniewski, Husserl, and ultimately Aristotle. They have attracted the attention of linguists and philosophers since 1935, but mostly outside the main trends of generative grammar.¹ Recently renewed interest in non-transformational approaches to syntax² suggests that it might be well to take another look at categorical grammars, since they seem to have been neglected largely because they had been shown to be equivalent to context-free phrase structure grammars in weak generative capacity and it was believed that such grammars were incapable of describing natural languages in a natural way. It is my purpose here to sketch a theory of grammar which represents a generalization of categorical grammars of the sort invented by Ajdukiewicz and further elaborated especially by Bar-Hillel (1953) and Lambek (1961).

Impetus toward this theory has come also from a quite different quarter: Hale, Jeanne, and Platero (1977) and others have suggested that not all languages are phrase-structure languages. It is my hope to develop a theory in which phrase-structure grammars are seen as a special case of a more general type of grammar, one which allows only a very small and restricted set of transformation-like operations. In the present paper, I will concentrate on the problem of "discontinuous constituents."³

EMMON W. BACH

I take it as a necessary part of our job to provide a theory of syntax and semantics (and phonology) for natural languages. It's worth noting that the primary motivation behind Ajdukiewicz's construction, insofar as I can see it, was to give a theory with a clear and consistent relation between syntax and semantics.

In Section 1, I will review briefly the basic ideas of categorial grammar; in Section 2, I will discuss the problem of discontinuous constituents; in Section 3 we will look at a generalization of categorial grammars, and in Section 4 I will present a sample grammar in this framework for a small part of English syntax.

1. Categorial grammars. The basic idea of categorial grammar is this: think of a phrase as composed of a sequence of words, each assigned to a category which acts either as an argument or as a function taking arguments of a certain category to make a certain resultant category. A sequence of words, each assigned to a category, is a phrase of category a, if it is possible to find a way of applying functions to arguments repeatedly in such a way as to end up with just the category a by exhausting the elements of the string. The system has obvious analogies to ordinary multiplication and my first example is couched in these terms.

We first define an infinite set of categories (CAT) as follows:

- (i) $2, 3, 5 \in \underline{\text{CAT}}$
- (ii) If $\alpha \in \underline{\text{CAT}}$, and $\beta \in \underline{\text{CAT}}$, then $(\frac{\alpha}{\beta}) \in \underline{\text{CAT}}$.
- (iii) Nothing else is in CAT

Elements of CAT may be used as indices on sets of expressions. Let's use the categories as indices for the sets of expressions consisting of the categories themselves.

$$B_A = \{A\}$$

(B_A denotes the set of basic, or lexical elements of category A.) So:

$$B_2 = \{2\}$$

$$B_{\frac{2}{3}} = \{\frac{2}{3}\} \quad \text{etc.}$$

(I will usually suppress outermost parentheses.)

We can now define the language induced by the set of categories as the union of the sets of phrases in each category A, P_A , where we can make explicit the ideas sketched above as follows: P_A

GENERALIZED CATEGORIAL GRAMMARS

(iv) If $\alpha \in B_A$, then $\alpha \in P_A$

(v) If $\alpha \in P_{\frac{A}{B}}$, $\beta \in P_B$, then $\alpha \beta$ and $\beta \alpha \in P_A$

Thus, for example, $\frac{2}{3} \in P_{\frac{2}{3}}$, $3 \in P_3$, so $(\frac{2}{3})3$ and $3(\frac{2}{3}) \in P_2$

We can specify a semantics for this "language" quite simply by letting our domain of individuals be the positive rational numbers and interpreting each basic expression as the number ordinarily denoted by it, "2" denotes two, " $\frac{2}{3}$ " denotes two-thirds, etc., and letting concatenation be interpreted as multiplication.

Let's now consider a linguistically more interesting example. Let $\underline{2}$ stand for the category of sentences (with no basic members). Members of other categories might be:

$$B_3 = \{\text{John, Mary, Bill}\}$$

$$B_{\frac{2}{3}} = \{\text{walks, runs, talks}\}$$

$$B_{\frac{(\frac{2}{3})}{3}} = \{\text{sees, kisses, loves}\}$$

$$B_{\frac{(\frac{2}{3})}{(\frac{2}{3})}} = \{\text{slowly, enjoyably}\}$$

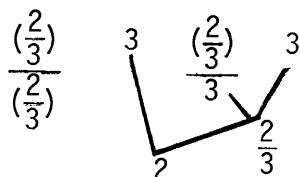
Then the language induced by the set of categories and lists of basic expressions includes these sorts of phrases:

P_2 : John walks, Mary slowly walks, walks Bill, sees John Mary,...

$P_{\frac{2}{3}}$: sees Mary, walks, Mary kisses enjoyably,...

Some expressions are excluded:

*slowly John sees Mary



EMMON W. BACH

This follows from the fact that functions must be immediately contiguous to their arguments.

The kind of language we have with a system of categories of the above type is what I've called an "M-language" (M for "mobile"): all and only permutations of siblings are permitted. (Every M-language is obviously context-free but not all CF languages are M-languages), e.g. $\{x \mid x = a^n b^n\}$ is not, as Stanley Peters pointed out to me, cf. Bach, 1975).

We can introduce ordering restrictions into our system by adopting a notation introduced by Bar-Hillel (1953) and Lambek (1961). Let $\underline{a/b}$ ($b \setminus a$) index the set of expressions "looking for" arguments of category b to their "right" ("left") to make expressions of category a . To approximate English we can replace the above categories by these:

$$B_3 \setminus 2, B_{(3 \setminus 2)/3}, B_{\frac{\binom{2}{3}}{\binom{2}{3}}}$$

The system now generates these well-formed members of $P_2(=S)$:

John enjoyably kisses Mary.
Mary walks slowly.

The system is obviously still context-free (now completely equivalent to CF grammars in weak generative capacity). [Exercise 1: make a set of categories for the language $\{x \mid x = a^n b^n\}$.] We can generate quite a lot of English with such a system. But we are still unable to do discontinuous constituents.

2. Discontinuous constituents. If all natural languages could be described by plain vanilla two-way categorial grammars of the sort we've looked at, we'd be home free. But natural languages notoriously display, or appear to display, discontinuous constructions. Some examples from English are these:

1. (a) John looked up the number.
(b) John looked the number up.
2. (a) I persuaded to leave all the men in the room.
(b) I persuaded him to leave.
3. (a) A package from India is here.
(b) A package is here from India.

In each of the (a) cases there is a contiguous string of elements which is arguably a constituent, while in the (b) cases the strings

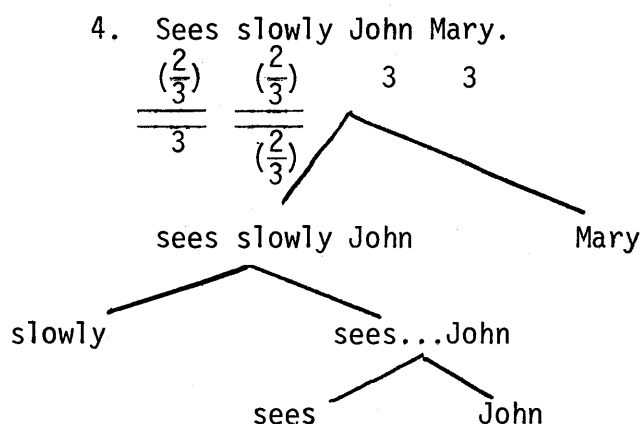
GENERALIZED CATEGORIAL GRAMMARS

are broken up. Examples like the above constituted one of the primary motivations for transformations in the first place.

A more extreme case is provided by languages which allow a lot of freedom in ordering. Here linguists have typically set up grammars which specify a fixed base order and then appealed to operations of scrambling (but see Hale, Jeanne, and Platero, 1977; Lapointe, 1980). Hale (cf. above reference and Hale, 1979) has claimed that Walbiri allows for permutations of all words around a fixed, second-place element in each clause. And there are many examples of free constituent order even in relatively fixed-order languages like English.

Imagine a language in which any order of constituents is possible, i.e. for any sentence, all "scrambles" of words (or morphemes) is also a sentence. We could represent functions that are looking for arguments "anywhere" as $\frac{a}{b}$. [Exercise 2: show that the free "scramble" of a context free $\frac{a}{b}$ language is not necessarily context-free.] Similarly, we could introduce the notation $a//b$ ($b\backslash a$) for functions looking for arguments indefinitely far to their "right" ("left").

The introduction of such categories entails that we will not be able to construct phrase-markers for all possible permutations. For example, consider a sentence that is disallowed for "contiguous English" but allowed in "scrambled English". We can analyze the sentence after the fashion of a Montague analysis tree, but the analysis can't be mapped onto a phrase-marker:



(This is only one of two possible analyses; in the other John is the subject.)

Of course, no language has such complete freedom of order; typically, certain constituents act as "boundary domains" for scrambling. (I conjecture that such bounding has the effect of keeping them all "almost context-free" (if not context-free) i.e.

EMMON W. BACH

that such grammars with bounding on recursive nodes do not generate the full class of context-sensitive languages.) We can achieve this result by stating general conditions, e.g. that S's be coherent.

3. Generalized Categorical Grammars. Let me sum up what I've suggested as a generalization of the syntactic categories of categorical grammar in the form of a recursive definition schema (for universal grammar):

(i) $\underline{a}, \underline{b} \in \text{CAT}$ ($\underline{a}, \underline{b}$ primitive categories)

(ii) If $\underline{a}, \underline{b} \in \text{CAT}$ so are

(a) $\underline{a}/\underline{b}$

(b) $\underline{b}\backslash\underline{a}$

(c) $\frac{\underline{a}}{\underline{b}}$

(d) $\underline{a}//\underline{b}$

(e) $\underline{b}\bowtie\underline{a}$

(f) $\frac{\underline{a}}{\underline{b}}$

(iii) Nothing else is in CAT.

I assume that this basic schema will be supplemented by a system of features, so that we can include e.g. $\underline{a}/\underline{b}$ to index a function that requires an argument immediately to its right with the additional stipulation that the argument have the case value 1 (e.g. accusative, more on this in Bach, ms.).

So much for syntax. The more difficult question is the semantics. For a first step, I'll assume (along the lines of Montague's PTQ) that there is a uniform mapping from categories to types and that every function argument rule (cancellation rule) carries with it automatically a rule of interpretation which says that the interpretation of the resultant category is the result of applying the interpretation of the function expression to the intension of the interpretation of the argument expression.

I've also not said very precisely how we stipulate how to analyze (or construct) sentences with these "long-distance" functions. I consider it still an open problem to try to give a general characterization of such syntactic operations. In the next section I'll show how to accomplish a syntactic/semantic analysis for several particular cases in English. The basic idea is this: we let function expressions of category $\underline{a}/\underline{b}$ (etc.) act syntactically as expressions of category \underline{a} , and supply a free variable of the type of intensions of the type corresponding to \underline{b} in the translation. This variable remains free until we find an expression of category \underline{b} , then we bind the variable with a lambda and let it loose.

GENERALIZED CATEGORIAL GRAMMARS

4. Some English Constructions. I now want to show how a generalized categorial grammar can be used to capture certain facts about English. I will exhibit a grammar which includes persuade and promise and so-called heavy NP-shifts.

The primitive categories (I ignore further analysis) are S (sentence), T (term-phrase, NP), CN (common noun), IV (tenseless VP's), IV (to VP's). The basic categories we need are these, with some typical members.

$$\begin{aligned}
 B_{CN} &= \{\text{woman, fish, elephant, ...}\} \\
 B_T &= \{\text{Bill, Mary, John, ...}\} \\
 B_{T/CN} &= \{\text{every, the, a, ...}\} \\
 B_{(T \setminus S)} &= \{\text{walks, runs, ...}\} & B_{IV} &= \{\text{walk, run, ...}\} \\
 B_{(T \setminus S)/T} &= \{\text{hits, sees, loves, ...}\} & B_{IV/T} &= \{\text{hit, see, love, ...}\} \\
 B_{(T \setminus S)/\overline{IV}} &= \{\text{tries, ...}\} & B_{IV/\overline{IV}} &= \{\text{try, ...}\} \\
 B_{((T \setminus S)/T)/\overline{IV}} &= \{\text{persuades, ...}\} & B_{(IV/T)/\overline{IV}} &= \{\text{persuade, ...}\} \\
 B_{((T \setminus S)/\overline{IV})/T} &= \{\text{promises, ...}\} & B_{(IV/\overline{IV})/T} &= \{\text{promise, ...}\} \\
 B_{(T \setminus S) \setminus (T \setminus S)} &= B_{IV \setminus IV} = \{\text{slowly, today, ...}\} \\
 B_{\overline{IV}/IV} &= \{\text{to}\}
 \end{aligned}$$

We need the following Rule Schemata (from Universal Grammar). The rules may be read as follows: $A := B C$: If $\alpha \in B$, $\beta \in C$, then $\alpha \beta \in A$ and if α translates as α' (i.e. 1') and β as β' (2'), then $\alpha \beta$ translates as (0'=) whatever follows = underneath the rule. I ignore completely, for simplicity, conditions on feature values for constituents, except for two; VST ("variable store") is a function which tells us what free variables have been put into the translation and remain to be bound at appropriate points, while $[-A]$ is a feature telling us that there is a "missing" phrase of category A (comparable to Gazdar's slash categories, Gazdar, forthcoming). Note further that the rules are schemata, and they (as well as the list of categories⁴) must be further specified for particular languages.

$$0. P_A := B_A$$

$$0' = 1'$$

$$1a. P_A := \left\{ \begin{array}{l} P_{A/B} \\ P_{A//B} \end{array} \right\} P_B$$

$$0' = 1' (^2')$$

$$1b. P_A := P_B \left\{ \begin{array}{l} P_{B \setminus A} \\ P_{B // A} \end{array} \right\}$$

$$0' = 2' (^1')$$

GENERALIZED CATEGORIAL GRAMMARS

Finally, let's consider phrases like give Bill a fish (not in our fragment). I've argued (Bach, 1980) that give (in this usage) belongs to the category $B_{(IV/T)//T}$. The standard derivation follows exactly that of persuade Bill to go.

$$\underline{\text{give Bill}} \in P_{IV/T} \quad (\text{by R2})$$

and is translated as $\lambda P_3[\text{give}'(P_3)(\wedge\text{Bill}')]]$, so by 1a we have

$$\begin{aligned} \text{give Bill a book} &\in P_{IV} \\ \lambda P_3[\text{give}'(P_3)(\wedge\text{Bill}')] &(\wedge\text{a-fish}') \\ &\equiv \text{give}'(\wedge\text{a-fish}')(\wedge\text{Bill}') \end{aligned}$$

Note that the inner argument of give, in this usage, is the thing given and that Bill is the direct object of the TVP give a fish.

This same standard derivation will assign to the phrase give a fish Bill an interpretation that reduces to this: give'(\wedge\text{Bill}')(\wedge\text{a fish}'), which is of course syntactically OK and means the same as give Bill to a fish.

What remains to be explained is why the option of R3 and R4 is not available to yield "heavy NP-shift" interpretations of phrases like give a fish the elephant (that is heavy). Note that R1a allows this derivation

$$\begin{aligned} \text{give} &\in P_{(IV/T)//T} & \text{give}' \\ \text{give a fish} &\in P_{IV/T} & \text{give}'(\wedge\text{a fish}) \\ \text{give a fish elephant} &\in P_{IV} & \text{give}'(\wedge\text{a fish}')(\wedge\text{the}'(\wedge\text{elephant}')) \end{aligned}$$

(Compare above: this is identical with the result of the standard derivation of give Bill a fish.) We don't want to block the option of letting double slash categories take immediately following arguments ("anywhere to the right" must include "immediately following"). I propose that what is crucial here is the identity of inner and outer arguments, so we can amend the rule schema 1a (second option) as follows):

$$P_A := P_{A//B} P_B \text{ unless } A = (X/B) \text{ for some } X^5$$

EMMON W. BACH

Footnotes:

¹Among those who have taken up categorial grammars are Lyons (1966), Lewis (1972), Geach (1972), Montague (1974: Paper 8: henceforth PTQ).

²To be mentioned here are Harman (1963), Brame (1978), Bresnan (1978), Schachter (1978), Hudson (1976), K. Ross (1980), Lapointe (1980), Gazdar (forthcoming). I am building in this paper on ideas of K. Ross (1978), R. Saenz, S. Lapointe (1980), and M. Flynn (1980).

³This paper is a companion piece to Bach (ms.), where I explore especially the kind of detailed dependencies found in the auxiliary system of English with a generalized categorial system. For details on the kind of system of which I am proposing a categorial variant, see K. Ross (1980), Bach and Partee (1980), Partee and Bach (1980).

⁴For an interesting exploration of the question of providing general principles determining what particular directional categories a language may pick, see M. Flynn, 1980.

⁵Answers to exercises:

Exercise 1: $B_{S/S'} = \{a\}$

$B_{S/b} = \{a\}$

$B_{S \setminus S'} = \{b\}$

$B_b = \{b\}$

Exercise 2: Let $L = \{X \mid X = (abc)^n\}$. L is CF (in fact regular). But Scramble (L) is not CF. For let $L' = \{X \mid X = a^n b^m c^k\}$ then $L' \cap L = \{X \mid X = a^n b^n c^n\}$ is not CF, but since the intersection of a CF language and a regular language is CF, L can't be CF.

REFERENCES

- Ajdukiewicz, Kazimierz (1935) Die syntaktische Konnexität. Studia Philosophica 1: 1-27. (Appears in English translation in Storrs McCall, ed., Polish Logic (Oxford: University Press, 1967)).
- Bach, Emmon (ms.) Generalized categorial grammars and the English auxiliary. Paper presented at the 4th Groningen Roundtable, 1980. (To appear in the proceedings.)
- Bach, Emmon (1975) Order in base structure. In Charles N. Li, ed., Word Order and Word Order Change (Austin and London: University of Texas Press).
- Bach, Emmon and Barbara Partee (1980) Anaphora and semantic structure. Proceedings of the Chicago Linguistic Society, Vol. 16.
- Bar-Hillel, Yehoshua (1953) A quasi-arithmetical notation of syntactic description. Language 29: 47-58.
- Brame, Michael (1978) Base-generated Syntax. Seattle: Noit Amrofer Press.
- Bresnan, Joan (1978) A realistic transformational grammar. In Morris Halle, Joan Bresnan, and George A. Miller, eds., Linguistic Theory and Psychological Reality (Cambridge, Mass: MIT Press).
- Flynn, Michael (1980) A categorial theory of the base. Unpublished Ph.D. dissertation, University of Massachusetts, Amherst.
- Gazdar, Gerald (forthcoming) Phrase structure grammar. To appear in G.K. Pullum and P. Jacobson, eds., On the Nature of Syntactic Representation.
- Geach, Peter T. (1972) A program for syntax. In D. Davidson and G. Harman, Semantics of Natural Language (Dordrecht: Reidel), pp. 483-497.
- Hale, Kenneth (ms.) On the position of Walbiri in a typology of the base. (Unpublished paper, MIT, 1979.)
- Hale, Kenneth, LaVerne Jeanne, and Paul Platero (1977) Three cases of overgeneration. In P. Culicover, T. Wasow, and A. Akmajian, eds., Formal Syntax. New York: Academic Press.
- Harman, G.H. (1963) Generative grammars without transformational rules: a defense of phrase structure. Language 39: 597-616.

EMMON W. BACH

- Hudson, Richard A. (1976) Arguments for a non-transformational grammar. Chicago: University of Chicago Press.
- Lambek, Joachim (1961) On the calculus of syntactic types. In R. Jakobson, ed., Structure of Language and its Mathematical Aspects. Providence: Proceedings of Symposia in Applied Mathematics, XII, American Mathematical Society.
- Lapointe, Steven G. (1980) A theory of grammatical agreement. Unpublished Ph.D. dissertation, University of Massachusetts, Amherst.
- Lewis, David (1972) General semantics. In D. Davidson and G. Harman, eds., Semantics of Natural Language (Dordrecht: Reidel) pp. 169-218.
- Lyons, John (1966) Towards a 'notional' theory of the 'parts of speech'. Journal of Linguistics 2: 209-236.
- Montague, Richard (1974) Formal Philosophy. Richmond Thomason, ed., New Haven.
- Partee, Barbara H. and Emmon Bach (1980) Quantification, pronouns, and VP anaphora. In J. Groenendijk, T. Janssen, and M. Stokhof, eds., Formal Methods in the Study of Language: Proceedings of the Third Amsterdam Colloquium.
- Ross, Kenneth (1980) Parsing English phrase structure. Unpublished Ph.D. dissertation, University of Massachusetts, Amherst.
- Schachter, Paul (1978) Review of Hudson, 1976. Language 54: 348-376.