Review Article

Arnim von Stechow, Konstanz

Categorial Grammar and Linguistic Theory
Reflections on :

Oehrle, R.T., Bach,E. & Wheeler, D. (eds.): ***Categorial Grammars and Natural Language Structures.*** **(Studies in Linguistics and Philosophy, vol. 32) Dordrecht 1988 : Reidel, 524 pages.***

Contents

*Acknowledgments: I wish to thank Irene Heim and Wolfgang Sternefeld for critical comments and Bruce Mayo for checking my English.

1. The attraction of categorial grammar

From time to time, Categorial Grammar (CG) experiences a powerful renaissance, as the late Yehoshua Bar-Hillel noted already at the beginning of the seventies, when he spent a sabbatical in Constance. The present volume bears witness to this fact in an impressive way. It documents the state of the art in 1985, when most of the papers underlying the actual articles were read at a conference at Tucson. No doubt, the theory has evolved since then, and the influence of this kind of grammatical research is likely to grow because of the appealing mathematical features of the formalism, which make CG a good candidate for applications in computer linguistics, a branch of linguistics with a notorious rate of growth.

The reasons for this success seem straightforward: CG is very attractive for anyone interested in semantic interpretation. The classical theory, *i.e.*, Ajdukiewicz (1935), can be regarded as the formalization of exactly one semantic principle of composition, namely *functional application*. Lambek (1958) generalized the approach by allowing derived principles like *functional composition*, *Geachian* and *Montagovian type lifting* and others. We shall talk about these in a moment. Semantically, the principles have in common that they abstract over at most one occurrence of a variable. Thus, no genuine variable binding can be represented by the formalism of CG alone. In order to obtain the full expressive power of λ-languages, *i.e.*, the entire variable binding, we have to add some appropriate combinators (or open variables plus lamda abstraction, as in Cresswell's (1973)
λ-categorial languages). Then anything seems to go, as we will see.

But even Lambek grammars without combinators are very powerful. They allow simulation of most kinds of movement rules assumed in Generative Grammar - more or less in the same way as this is done in the so-called GPSG-framework (*Generalized Phrase Structure Grammar;* see Gazdar et al. 1985).

Thus, extended categorial grammars - as versions of the Lambek calculus are sometimes called - seem to combine all the nice features theoretical linguists always wanted to incorporate into a formal system. In particular, it appears that categorial grammar makes multiple level theories like the so called GB framework (*Government and Binding;* see Chomsky 1981, 1982) obsolete, because a strict surface analysis of phenomena that traditionally motivated transformations seems possible.

Let me make clear from the beginning that I do not believe that any of these great expectations will be fulfilled in the future. In the following it will turn out that it is very doubtful whether CG is a genuine alternative to transformational grammar. My personal view is that CG can best be regarded as the formalization of a certain class of semantic principles of composition. These principles can perhaps provide the semantics for some kind of movement rules. Furthermore, they may be used to analyse argument projection in syntax and morphology. Thus, categorial principles are likely to belong to a modul of grammar only, say theta theory.

However, stronger claims as to the power of categorial theory are very likely to be wrong. As it stands, the categorial formalism isn't even able to express the usual constraints investigated in generative grammar of the GB-kind. It has to be enriched by all sort of metaconventions which mimic what we know from generative theory. Whether you call the result of such extensions still a categorial grammar or not seems to be largely a matter of taste.

2. Some principles of Categorial grammar

Before I start with the discussion of the articles, I will introduce the most important principles of CG.

The classical model of CG is formulated in Ajdukiewicz (1935). In this version of the theory, each complex constituent A breaks into a *functor* A/B and an *argument* B. If the functor denotes the function **f** and the argument has the semantic value **b**, then the value of A is **f(b)**, i.e. **f** applied to **b**. The principles that combine a functor with an argument are known as the rules of *functional application*:

(1) *Functional application* (**FA**)
A/B + B --> A
B + B\A --> A

As Ajdukiewicz (1935) has shown, we can model quite interesting fragments of natural language in this extremely simple and elegant framework. Most syntacticians haven't seriously considered the possibility that CGs are a suitable tool for modeling natural languages, because they felt that syntax wasn't as simple as that. In the sixties, Bar-Hillel, Shamir & Gaifman (1960) showed that CGs are equivalent to a very special type of context-free grammars, and since the latter were believed to be inadequate for the analysis of natural language, this belief carried over to CG as well.

In the last few years, the view that context-free grammars are inadequate has been challenged, notably by GPSG-grammarians (see, e.g., Gazdar et al. 1985). In particular, it was shown that long distance dependencies like *Wh*-movement could be simulated by means of context-free rules. The renaissance of context-free grammars carried over in a natural way to CGs. One of the reasons certainly is that Gazdar's slash-categories – the crucial device for simulating movement – are categorial principles of gap inheritance (or gap projection), which were formulated in an explicit way in Lambek (1958), an article that had been ignored for two decades, but the formalism of which has been rediscovered recently and is now known as the Lambek Calculus. The Lambek Calculus is quite powerful. Among other things, it allows for the derivation of Montagovian type raising and the so called Geachian rule, which will be discussed subsequently. The two rules allow the simulation of Gazdar's slash categories. (Given the historical priority of Lambek (1985), one ought to say that Gazdar's mechanism simulates certain principles of the Lambek Calculus.) It follows that the Lambek Calculus can analyze long distance dependency more or less in the same way as this is done in the so-called GPSG-framework.

The most important principles accounting for gap projection are **M** and **G**; see (2) and (3) below.

(2) *Montagovian type raising* (**M**)

a. A --> B/(A\B),

b. A --> (B/A)\B

The interpretation of the rules is this: If an expression of category A has the semantic value **a** , then the "lifted" expression of category B/(A\B) or (B/A)\B has the value; **λff(a)** where **f** is of (semantic) category (A\B) or (B/A) respectively.

The third prominent principle is known as Geach's law.

(3) *Geach's law* (**G**)

a. A/B --> (A/C)/(B/C)

b. B\A --> (C\B)\(C\A)

The interpretation is this: If an expression of category A/B denotes the function **f**, then the "lifted" expression of type (A/C)/(B/C) denotes the function **λgλxf(g(x))**, where **g** is of (semantic) category B/C and **x** is of (semantic) category C. Similarly for (3b).

To be sure, Geach's law was not formulated in this way in Geach (1972). Geach stated a weaker principle, which, under certain assumptions, is equivalent to functional composition, which will be discussed next. Thus, the term "Geach's law" is a misnomer in several respects, but it is established in the literature and I therefore will keep it.

It is useful to remind the reader that **M** and **G** are easily derivable in the Lambek calculus (see p. 299 ff. of the volume under review) and, in fact, both have been derived as theorems in Lambek(1958).

It should be noted that the principles of functional composition are a direct consequence of **G**. Suppose we are given the sequence of categories A/B,B/C. We apply **G** to the first functor and derive (A/C)/(B/C), B/C. This reduces to A/C by **FA**. Since the principles of functional composition play a prominent role in the theory, I will list them explicitly:

(4) *Functional composition* (**FC**)
a. A/B + B/C --> A/C
b. A\B + B\C --> A\C

The semantics is the following: If an expression of category A/B denotes a function **f** and an expression of category B/C denotes a function **g** then the complex expression formed by the rule denotes the function **λxf(g(x))**. Analogously for (4b).

These are the most important principles for an understanding of the articles and the current literature in general.

3. Survey of contents and notations in the book

The articles of the book are ordered alphabetically according to the names of the authors. For convenience, I will group them into several classes, however. There are two *introductory articles* (Bach and Casadio, see section 4.1 of this review), three articles about *formal properties of the categorial formalism* (van Benthem, Buszkowski, and Lambek, see section 4.2). The main group of the articles is devoted to an elaboration of a *linguistic theory of CG* (section 4.3). The articles by Dowty, Huck, Keenan & Timberlake, and Steedman are mainly concerned with *syntax* (section 4.3.1). Two articles are about categorial *morphology* (Moortgat, and Hoeksema & Janda : section 4.3.2). One article (Wheeler) deals with categorial *phonology* (section 4.3.3). Two articles compare and combine categorial grammars with other theories (Chierchia and Pollard, cf. section 4.3.4).

Two articles are not discussed in this review: Richard T. Oehrle's "Multi-Dimensional Compositional Functions as a Basis for Grammatical Analysis" and Susan Steele's "A Typology of Functors and Categories". The former article provides a very general theory of compositionality. It is highly abstract and presumably belongs to a larger project. Furthermore, it doesn't discuss concrete examples nor is it concerned with issues that are relevant for CG proper. As for Steele's article, I have to confess that for one reason or another I haven't quite understood it. In any case, it doesn't seem to be concerned with questions central for the theory of categorial grammar.

The organization of this review is as follows. For each article, I will focus on one or two points that are central for the theory of categorial grammar. These points will be discussed and they are not touched upon anymore in our review of the later articles, even when they are central for the later authors as well. In this way we can hope to collect most of the theoretical issues which are important for an overall assessment of categorial theory. The discussion is summarized at the end.

Before I comment on the different articles, a remark is due about the notation. The classical notation for right-application used by Bar-Hillel is A/B + B --> A, whereas B\A denotes a functor that has its B-argument on the left side. Some people (*e.g.* Dowty) find this notation counter-intuitive. He uses A/B in the same way as above,, but a functor with its B-argument on the left side is denoted by A\B. Moortgaat writes B\A for classical B\A and B/A for classical A/B. (Personally I find Moortgat's notation the hardest to read.) Other notational variants are found in the volume as well. For instance, Huck represents the classical functor B\A as $A/_{<}\underline{B}$.

In the following I will use the classical notation throughout. As to the numbering of the examples, I have indicated the original numbers in square brackets.

4. Discussion

4. 1 Introductory articles

4.1.1. Emmon Bach has been the godfather of extended CG for many years. In his "Categorial Grammar as Theories of Language" he tries to say, among other things, what CG *is*. As far as I can see, Bach carefully avoids committing himself to restrictions in syntax. The general format of the syntax rules given in section 1 is a special version (the operations have only finitely many arguments) of the rules of Montague´s *Universal Grammar*. It seems to me that virtually any of the grammars on the market can be brought into that format. On page 24, however, Bach commits himself to an important semantic constraint: "In extended Montague grammar and categorial grammar we must commit ourselves on the semantic import of our syntactic categories". In the theoretical context of Montague grammar this can only mean that to each syntactic category there is a corresponding meaning category that contains entities of the same kind, *i.e.* names denote individuals, intransitive verbs denote functions from individuals to truth-values, and so on. But this parallelism has its well-known problems. To give an example: since adverbial prepositional phrases and prepositional objects are not in the same semantic category they must belong to different syntactic categories. Cases like this can be multiplied, and they are unsatisfactory, linguistically speaking.

Be that as it may, I conclude from Bach's thoughtful introduction that it is not at all easy to say what categorial grammar is.

4.1.2. Claudia Casadio's "Semantic Categories and the Development of Categorial Grammar" gives a nice survey of the historical development of CG. My personal impression from this representation is that, although seminal ideas may be contained in the work of Husserl and Le´sniewski, true CG seems to begin only with Ajdukiewicz, since Husserl and Le´sniewski don't seem to have developed an explicit formal theory.

(Perhaps, they have. But this fact doesn't emerge from Casadio's discussion.)

I would like to add two remarks to Casadio's exposition. On page 111 she claims that Geach(1972) was the first to introduce the category s/(s/n) for quantified nominals like *every man*. In footnote 29, she says: "Ajdukiewicz tried to develop an analysis of quantifiers, but he confined himself to first level functors suggesting the category s/s together with contextual restrictions to distinguish such operators from other sentential functors as negation."

This is not correct. Ajdukiewicz(1935) explicitly rejects the possibility that the prefix in a sentence of the form (∀x)Px can have the category s/s because, if this were so, its truth-value would depend on the truth-value of Px only. He rather suggests that it has the form ∀((^x)Px) where (^x) is the abstraction operator and ∀ has the category s/(s/n). (See Ajdukiewicz (1935, p.26); to be sure, Adjukiewicz writes Π for ∀.) This is exactly the analysis proposed later by Montague, Lewis, Geach and others. Thus, the explicit introduction of the nominal goes back to Ajdukiewicz, and the idea is present in the work of Frege, of course.

The next remark concerns Casadio's claim that the decision procedure for a Lambek grammar is simpler than that for a restricted grammar of the Ajdukiewicz type (cf. p. 108 f.). It is quite obvious that the opposite is true. A Lambek-grammar contains all the combinatory possibilities of a classical grammar together with many new possible kinds of combinations stemming from additional derivation rules, which lead to such principles as **M** and G. We have to check these additional possibilities when we decide on the well-formedness of a string.

4.2. Mathematical theory of CG

4.2.1. The article "Generative Power of Categorial Grammars" by Wojciech Buszkowski is entirely technical. I am sure that the work has a high standard, but I don't feel competent to assess the importance of the results. Therefore, I have to leave the evaluation to the formal logicians.

4.2.2. The importance of J. Lambek for the theory of CG is commonly recognized in our days. His article "Categorial and Categorical Grammars" repeats some of the results of Lambek(1958) and relates the calculus with the mathematical theory of topoi. Some of the basics contained in the article is reported in the following section.

4.2.3. Johan van Benthem's "The Lambek Calculus" reports on some results of his logical investigations into the Lambek calculus or, more precisely, a special version thereof. I will not comment on this, but I will take the occasion to give the reader a rough idea what the calculus is like. I want to do this for two reasons.

Firstly, the notation may be rather unfamiliar for the reader used to classical CG only. Hence, some comments may be helpful here. Secondly, van Benthem gives a semantics for the calculus that enables us to derive the interpretation of all important categorial principles, *e.g.* for the principles **M** and **G**.

The simple calculus studied on p.39 ff. has *axioms* of the form A ==> A, the rules of functional application (which I will continue to call **FA**) and abstraction rules (which I will call **AB**).

Let us consider *functional application* first. It is formulated in this way:

(5) If $A_1 \ldots A_i$ ==> B and $A_{i+1} \ldots A_n$ ==> (B | C), then $A_1 \ldots A_n$ ==> C. (**FA**)

(A | B) is an undirected category that includes directed functors of the form (A/B) and (B\A). ==> denotes the derivability relation. The semantics is the same as that of **FA** above, i.e., where the expression corresponding to B denotes **b** and that corresponding (B | C) denotes **f**, C denotes **f(b)**.

In order to see how this works, let us derive the sentence $Mary_{NP}$ $sleeps_{NP \mid S}$:

(6) Mary		Mary		sleeps		sleeps	
NP	==>	NP	(axiom)	NP\S	==>	NP\S	(axiom)
Mary sleeps	==>	Mary sleeps	(**FA**)				
NP NP\S		S					

Such a derivation encodes all the information found in a categorial tree of the usual sort. The axioms correspond to the terminal branches and an application of a derivation rule corresponds to a branching of the tree.

The most interesting rules of the calculus are the *abstraction rules* :

(7) a. If $A_1,\ldots,A_n$,B ==> C, then $A_1,\ldots,A_n$ ==> (C/B). (=**AB**)
b. If B, $A_1,\ldots,A_n$ ==> C, then $A_1,\ldots,A_n$ ==> (B\C)

Before we comment on the semantics of these rules, let us illustrate it by deriving the NP\S $sleeps_{NP\backslash S}$ $not_{S\backslash S}$. From **AB**, we know that the following relation holds:

(8) If *x* *sleeps* *not* ==> *x sleeps not*, then *sleeps* *not* ==> *sleeps* *not*
NP NP\S S\S S NP\S S\S NP\S

Here, x is an unspecified NP which may be thought as a variable. The semantics of **AB** makes sure that this variable is bound in the conclusion by λ-abstraction (hence the name of the rule). In other words, the interpretation of (7) is this: if B denotes the variable x and the expression corresponding to C – which has the form $A_1,...,A_n,B$ or $B, A_1,...,A_n$– denotes the entity c, then the expression corresponding to (C | B) has the semantic value λxc. That B can denote a variable x, is a metalinguistic stipulation. The variable is not explicit in the syntax. In the example, we have spelled it out for convenience.

Returning to (8), it should be obvious that the antecedent of the statement is derivable by means of **FA** alone. Hence the conclusion is valid, too. Since the right hand side of the antecedent denotes the open proposition **not'(sleep'(x))**, the right hand side of the conclusion denotes the function **λx.not'(sleep'(x))**.

AB is the only principle of the calculus that binds a variable. It is not possible to abstract over variables occurring at different structural positions. Thus, the Lambek calculus allows us to abstract over exactly one variable occurrence. Clearly, this is a severe restriction of the expressive power of the language. No general variable binding is possible without the introduction of some further device. I shall return to this point when I discuss Steedman's paper.

Nevertheless, the abstraction rule considerably increases the expressive power of the grammar as compared to Ajdukiewicz's classical model. Let me illustrate this by deriving one version of Montagovian type raising together with its semantics:

(9) A ==> A (A\B) ==> (A\B) (axioms)
      a   P
______________________________ **FA**
A,(A\B) ==> B
P(a)
- **AB**
A ==> ((A\B)\B) (=**M***)
a **λPP(a)**

The result of the derivation is the rule **M*** with its standard semantics.

In order to complete our first experiences with the calculus, let us derive the rule **G**, too.

(10) $C ==> C \qquad C\backslash A ==> C\backslash A$ (axioms)

$$\underline{\qquad x \qquad\qquad\qquad y}\ \mathbf{FA}$$

$$C, C\backslash A ==> A \qquad\qquad A\backslash B ==> A\backslash B \text{ (axiom)}$$

$$\underline{\qquad\qquad y(x) \qquad\qquad\qquad\qquad f}\ \mathbf{FA}$$

$$C, C\backslash A, A\backslash B ==> B$$

$$\underline{\qquad\qquad\qquad f(y(x))}\ \mathbf{AB}$$

$$C\backslash A, A\backslash B ==> C\backslash B$$

$$\underline{\qquad\qquad\qquad \lambda x f(y(x))}\ \mathbf{AB}$$

$$A\backslash B ==> (C\backslash A)\backslash(C\backslash B)$$

$$f \qquad\qquad \lambda y \lambda x f(y(x))$$

Van Benthem (1984) has shown the following result: A "one-directional" grammar whose categories are built up by means of "/" only and which contains the rule **FA** , **G** and a type raising rule **M*** of the form A --> B/(B/A) is closed under permutation, in other words, if a sequence of categories X reduces to A, then any permutation of X reduces to A as well. (The original Lambek calculus has rules of type **M** only. This system doesn't have the permutation property.)

In order to realize what this means consider the following sentences:

(11)a. John told Mary that Ede would be late.
b. John told Ede that Mary would be late.
c. Ede told John that Mary would be late.

Since the rules **G** and **M*** don't affect the interpretation, we have it that in a grammar of the type considered by van Benthem, each of these three sentences has an analysis under which it means the same as the natural interpretation of any of the two other sentences. Furthermore, we can generate a lot of rubbish like *would told that John be Mary late Ede,* which, according to the theory, should have the same meaning as these sentences. Van Benthem is not too worried about this consequence, but most linguists probably would be.

Quite generally, extended CGs have the property that, when you add some innocent looking principle, the system suddenly explodes with respect to combinatorial power. Van Benthem has called this law of categorial nature *Moortgat's nuisance.* I will return to this point when I discuss Moortgat's article. Thus, as in any other syntactic theory the question arises as to how we can restrict the formalism in such a way that it mirrors natural grammar.

On the other hand, the calculus is too weak for modeling natural language, since it lacks a complete equivalent of variable binding. It is obvious that we need variable binding for doing semantics of natural language. I shall come back to this question when I discuss Steedman's paper.

4.3 Linguistic theory of CG

4.3.1 Syntax

4.3.1.1. The paper "Phrasal Verbs and the categories of postponement" by G.J. Huck is interesting for at least two reasons. Firstly, it discusses the close analogy between the principles of gap projection (alternatively "slash projection" or "postponement", the latter being Huck's own term) and movement. Secondly, it suggests that at least some discontinuities cannot be treated by the methods of CG, but would rather have to be treated by something like syntactic transformations. (Actually, Huck assumes crossing branches instead of movement.) If Huck is right on this, it will follow that syntactic movement does not reduce entirely to categorial principles of gap projection. I have to add, however, that the data discussed by Huck are a problem for any grammatical theory.

Consider an example. We can analyze *Max seems to sleep* by means of a "disharmonic" version of functional composition - here called FC* - in the following way:

(12) Max seems to sleep

$$\dfrac{NP \qquad \dfrac{S/S \qquad NP\backslash S}{NP\backslash S}\,FC^{*}}{S}\,FA$$

If we think of the phrase *to sleep* as an S with a subject gap t_{NP}, we can represent this in the following way:

(13) Max seems t_{NP} to sleep

```
                  \ /
NP     S/S    NP\S  FC*
       __________
        NP\S  FA
_______________
   S
```

Thus, **FC*** might be thought as a principle to project a gap in an argument to the functor governing this argument. Since this gap ultimately has to be filled by a missing argument, we obtain the same result as if we had "raised" the argument to the superordinate subject position. In other words, the particular instance S/S + NP\S --> NP\S of the rule **FC*** can be interpreted in a straightforward way as movement. The consequence is, whatever can be done by rules of this kind, can be done by movement as well.

To my mind, Huck is right in remarking that the slash mechanism of GPSG is essentially the same as the principles of "postponement" in CG, *i.e.* the principles of gap inheritance (see p. 255). The same point has been made by Dowty and Steedman in their papers. Under this perspective, GPSG may be considered as a combination of categorial and phrase structure grammar. The analogy between the categorial slash and movement doesn't come as a surprise. The slash category has been introduced into context-free grammar in order to mimic movement in a context-free framework, and if slash rules in GPSG can be interpreted as movement, **FC** and **FC*** can be interpreted as movement as well.

Is this version of categorial grammar strong enough to analyse discontinuous particle verbs as in the following sentences?

(14)a. He looked it up
b. He threw the package away

Huck assumes that *threw* has the category (TV/e)/Part, where e is the category of proper names, VP is (e\S), where TV is VP/e and Part is the category of particles. Then the VP of (14a) has the analysis

(15) threw the package away

```
TV/Part    TV\VP  FC*   Part
__________________
      VP/Part
      ___________________________  FA
           VP
```

This works fine. One of the problems for this approach is, however, that it has to assume that particle verbs have a compositional semantics, because only the gaps of semantic entities can be projected in CG. Now, this is obviously not the case for a verb like *look up*. In order to analyze (14b) by means of **FC***, one would have to assume an entry like:

(15) *look* $\mathrm{TV}/[\mathrm{up}]_{\mathrm{Part}}$

Here $[\mathrm{up}]_{\mathrm{Part}}$ is a category that is instantiated only by *up*; furthermore, it has no independent meaning. In order to make the analysis work, one would have to stipulate a dummy meaning for any idiosyncratic category of this sort. It must be assumed then that the meaning of the functor *look* does not depend on the *up*-argument. Clearly this is greatly *ad hoc*. Further complications arise when we consider the interaction of particle verbs with adverbs. Among others, Huck points out the following contrast:

(16) a. He threw the package quickly away
b. *He looked the number quickly up

It is very hard to see how (16b) could be blocked in a natural way invoking the principles of CG only, given that the two particle verbs have entirely parallel lexical entries. The problem is that we need restrictions for movement – perhaps even non-local ones – which we can't formulate in this framework. Considerations of this kind make Huck favour a movement solution. He assumes the following two transformations:

(17) a. The direct object can be placed immediately after the main verb.
b. A heavy or focused constituent can be moved to the end of a phrase.

Using these principles, Huck assumes a D-structure of the following kind for the VP in (17b) (recall that **FC*** is dismissed in favour of the two movement rules):

(18) (((look up) quickly) the number)

Applying rule (17a), we can place the direct object *the number* directly behind *look* and obtain (19):

(19) ((look the number up) quickly)

If *up* is not focused, we cannot proceed to (16b), since *up* does not count as a heavy constituent. On the other hand, if *up* is focused, we can move the particle to the end of the phrase and obtain the surface order exhibited by (16b).

In his article, Huck uses trees with crossing branches in order to simulate the movement rules mentioned. Unfortunately, the details of the formalism are not spelled out. But the arguments are largely convincing and lead to the conclusion that there is no natural way to treat phenomena like these by categorial means. This suggests that the relation between movement rules and the categorial principles of gap inheritance is less direct than has been assumed so far: gap inheritance always has a semantic parallel, *viz.* λ-abstraction. Movement can have this semantic pendant, but it need not have it. Some examples of particle movement belong to the latter case: this type of movement depends on certain formal features of the particle, for instance intonation.

It remains to say that the data considered by Huck are a challenge for any grammatical theory. Properties like "heaviness" cannot be defined by the usual structural methods. Furthermore, structures like (18) are not possible in standard accounts (for case theoretical reasons), and the movement (or branch crossing) relation assumed in (18) and (19) are not possible in GB-theory either. So the phenomena discussed here remain rather mysterious and can't be used PRO or CONTRA one or another theory.

4.3.1.2. With David Dowty's article "Type Raising, Functional Composition, and Non-Constituent Conjunction" we are right at the heart of extended CG. Dowty defends the thesis that non-constituent coordination (NCC) is the natural outcome of independently motivated processes like type raising and functional composition. Let us consider some typical examples of NCC:

```
(20) John saw  Mary    yesterday   and  Bill    today      [=(23)]
      e    TV  TV\VP   VP\VP FC    Con  TV\VP   VP\VP FC
               ______________           _____________
                  TV\VP                     TV\VP C
                  ______________________________
                               TV\VP FA
          ______________________
                   VP FA
     ____________
           t
```

e is the category of proper names ("entities"); t is the category of sentences ("truth-values"); IV is (e\t), and TV is (IV/e). Con stands for an appropriate category of conjunction. **C** applies this functor to its arguments. Con, to be sure, is an abbreviation for many categories. The restriction is that the two conjuncts are of the same category. In the categorial tree above, the "non-constituents" *Mary yesterday* and *John today* have become constituents. In order to achieve this, *John* and *Mary* have to be functors applying on the TV. Otherwise, the rule **FC** could not form the two "non-constituents".

Dowty gives no semantics for the rule **C**, but it is clear what the interpretation must be. Consider a conjunction of two VPs, for instance. The meaning of [VP and VP] has to be **λx [VP'(x) and' VP'(x)]**. Clearly, we have abstraction over more than one occurrence of a variable at this point, an option not available in the restricted form of the Lambek Calculus. Thus the expressive power of the language is increased by this rule. It should be noted that Lambek himself pointed out this extension in Lambek (1958).

The question that immediately arises is this: Why should we not analyse (20) as a special case of Gapping, i.e. why should we not assume the much simpler analysis (21)?

(21) John [$_{VP}$[$_{VP}$ saw Mary yesterday] and [$_{VP}$ SAW Bill today]]

The verb to be gapped is in capital letters. Note that we need a gapping rule (which presumably is a PF-rule) anyway for conjunctions with a subject in the second conjunct:

(22) Mary drinks beer and John DRINKS wine

There is no way to account for this sentence by categorial methods. Once we subsume (20) under gapping, most of the problems Dowty discusses vanish.

Dowty is aware of this possible objection and claims that NCCs are not cases of Gapping. He provides three arguments in support of this view: 1. Gapping involves an intonation break, whereas NCC does not. 2. Gapping belongs to a formal register, while NCC does not. 3. Gapping becomes significantly worse if there are more than two 'remnant' constituents, whereas we do not observe this effect for NCC.

I am not convinced by these arguments, for we might say just as well that Gapping in VP-conjunctions has exactly the properties Dowty assumes for NCC, whereas Gapping in S-conjunctions goes together with the properties which Dowty reserves for Gapping exclusively. Let us put this criticism aside, however, and let us assume that NCC is a genuine linguistic problem.

Applying the same method we can combine an indirect object and a direct object to form a constituent:

(23) John gave Mary a book and Sue a record [=(27)]

```
(23) John gave   Mary       a book      and  Sue        a record    [=(27)]
          TTV   TTV\TV   TV\VP FC            TTV\TV   TV\VP FC
                     TTV\VP                        TTV\VP C
                                TTV\VP FA
                     VP
```

TTV is (TV/e). In this particular case, *Mary* has to belong to the category TTV\TV. As the examples show, a noun like *Mary* belongs to many different categories: e, TV\VP, TTV\TV and others. We may think of these categories as the outcome of a generalized version of the type raising rule **M**. Remember that **M** raises the category e to the category t/VP. This is a functor converting a one-place predicate into a zero-place one. Generalizing this procedure, we can raise e to a functor that makes an (n-1)-place relation out of an n-place one. This raising operation is implicitly present in Montague's PTQ-analysis of object transparent verbs like *to find, to lose* etc. (see PTQ, p.263, (4)). Semantically, the operation amounts to quantifying into VP. Dowty believes that this kind of type raising is independently motivated. Among other things, we need the raised types for combining quantifier phrases with a verb (cf. p. 162 ff., for pertinent discussion).

Still more complicated cases are NCC's which include partial subconstituents as in the following examples:

(24)a. John went to Chicago on Monday and New York on Tuesday [= (35)]
b. Susan gave Mary a book on Monday and a record on Tuesday [=(62)]

Dowty analyzes (24a) as

```
(25)  John went to  Chicago on Monday  and  Detroit on Tuesday  [=(41)]
        e   VP/e       (VP/e)\VP              (VP/e)\VP
                                  (VP/e)\VP
```

From here on, we can use **FA** to finish the derivation. The analysis of the subconstituents proceeds as follows:

(26) went to Chicago on Monday

$$\frac{\underline{VP/(VP\backslash VP)\quad (VP\backslash VP)/e}\ \mathbf{FC}}{VP/e}\qquad \frac{\underline{(VP/e)\backslash VP\quad VP\backslash VP}\ \mathbf{FC}}{(VP/e)\backslash VP}$$

The climax as to complication is reached with the following sentence:

(27) Bill gave and Max sold a book to Mary and a record to Sue [=(63)]

In order to understand Dowty's analysis, we have to advance stepwise. Dowty assumes that a ditransitive verb has the category TTV=(((VP/PP)/e)\VP). From example (23) we know already that a sequence consisting of a direct object (DO) plus an indirect object (IO) has the category TTV\VP. Now, *Bill gave* is a sentence with missing DO+IO-constituent.Therefore, it must have the category t/(TTV\VP). The conjunction of *Bill gave and Max sold* has the same category, of course. This reflection motivates the following analysis:

(28) <u>Bill gave and Max sold</u> <u>a book to Mary and a record to Sue</u>

$$\frac{\underline{t/(TTV\backslash VP)\qquad\qquad TTV\backslash VP}\ \mathbf{FA}}{t}$$

In his article, Dowty says that the categorial analysis is much simpler than a (somewhat enriched) GPSG-analysis would be. I find claims of this sort unconvincing, because it follows from the discussion of Huck's paper that there is a straightforward respelling of the categorial principles of gap-projection into GPSG or GB-theory. Let us first consider the last example. On page 186, Dowty himself says what a GPSG-analysis would be, namely something like the following:

(29)Bill gave $t_{VP/V}$ and Max sold $t_{VP/V}$ t_V a book to Mary and t_V a record to Sue

$$\frac{\underline{S/(VP/V)\qquad\qquad VP/V}}{S}$$

Dowty calls this "a horror", but this characterization should apply to his own analysis as well, because the GPSG-structure is virtually the same as his categorial structure. (The reader should convince himself of this by going step by step through the motivation for Dowty's complicated categories given above.) To be sure, we have to enrich the slash categories of GPSG-grammar so that they may dominate slash categories.

Dowty remarks on this on page 187: "It would be unfortunate to have to complicate the theory in general and the grammar of English in particular in such a way for this one type of sentences to be produced." If Dowty's analyses are correct, then there can't be any doubt that GPSG has to be strengthened along these lines. As it stands, GPSG is not rich enough. But it doesn't follow from this that CG is simpler than GPSG.

And, of course, we can reformulate Dowty's analysis in terms of movement. First we adjoin the verbs to their respective VPs and then we extrapose the headless VP-conjunct "across the board", obtaining the following structure:

(30) $[_S [_S$ Bill $[_{VP}$ gave$_i$ VP$_j$] and Max $[_{VP}$ sold$_i$ VP$_j$]]
$[_{VP_j} [_{VP}$ V$_i$ a book to Mary] and $[_{VP}$ V$_i$ a record to Sue]]]

Note that I do not claim that this structure is licensed by current GB-principles. I merely want to point out that Dowty's analysis has an obvious equivalent in terms of movement.

The simulation of example (24a) by means of movement rules is a bit more complicated. We first have to incorporate the preposition *to* into *went* and then we have to adjoin the result across the board to the VP-conjunction:

(31) John $[_{VP} [_V$ went to$_j$ $]_i$ $[_{VP} [_{VP}$ t_i [t_j Chicago] on Monday]
and $[_{VP}$ t_i [t_j Detroit] on Tuesday]]]

In order to reformulate this in a GPSG-framework, we have to create an equivalent for incorporation. It is straightforward to do this; therefore, I leave it out.

This comparison relativizes Dowty's claim that NCS's like the ones cited above cannot be regarded as left node raising. If we incorporate the relevant preposition before we apply left node raising, then the structure (31) can be generated. On the contrary, Dowty's analyses can all be said to be instances of left node raising. This is clear for (20) and (23), because in both cases the VP-conjunct has the structure $[_{VP}$ V$_i$ $[_{VP}$ t_i ...]] & $[_{VP}$ V$_i$ $[_{VP}$ t_i ...]]. As we have observed, (25) has essentially the same structure. To be sure, (30) is a bit more complicated; it arises as a combination of left node raising plus right node raising under Dowty's analysis.

What do we have to make of all this? To my mind, the essential question is whether a movement analysis of the kind advocated by Dowty is correct for these phenomena. I have expressed my doubts already. It seems to me that a treatment by means of Gapping (and other rules of ellipsis) is more adequate for these phenomena. Among other things, a weakness of Dowty's treatment is that it does not seem to generalize for adjacency pairs of utterances, where we find the same configurations as in the examples discussed:

(32) Did John give a book to Mary? No, a record to Sue.

I think this is clear case of ellipsis. Thus, we need the relevant principles for ellipsis anyway. Dowty, on the other hand, has to claim that this is an entirely different construction.

4.3.1.2. E.L. Keenan & A. Timberlake's paper "Natural Language Motivation for Extending Categorial Grammar" takes up a particular case of systematic ambiguity of functors. The problem seems to be this. The copula subcategorizes NP and AP (*John is a fool/John is fat*) and converts them into VP. Hence, it should belong to both VP/NP and VP/AP. Since one word cannot belong to more than one category, it follows that we have two copulae. However, intuitively speaking, we only have one. In order to make precise this intuition, Keenan & Timberlake say that the copula belongs to the *2-tuple category* <VP,VP>/<AP,NP>. The syntactic rules combining this functor with an argument are these:

(33) a. <VP,VP>/<AP,NP> + AP ==> VP
b. <VP,VP>/<AP,NP> + NP ==> VP

The first rule receives the same interpretation as the rule VP/AP + AP ==> VP, and the second rule is interpreted as if it were VP/NP + NP ==> VP. The notion is generalized to the concept of *n-tuple category*

(34) $<C_1,...,C_n>/<D_1,...,D_n>$.

The pair C_i/D_i is called the *i-th coordinate* of the category. The semantic interpretation of these categories may best be thought as a sequence of functions $<f_1,...,f_n>$, where f_i is the value of the i-th coordinate. The two authors codify this information a bit differently (see, p.267), but it amounts to precisely this, as far as I can see.

Thus, the formalism seems little more than a complex notation for several rules, similar to the brackets used in earlier generative grammar for a more condensed notation of context-free rules, though this interpretation is disputed by the authors. They justify the approach by the so called *n-Tuple Universal* to which I will turn at the end of my discussion.

Let us consider some applications of the theory first. One example is the interaction between the applicative and the passive morpheme in Kinyarwanda. The applicative **-ir-** is a predicate functor which introduces a new argument - the beneficient - and is interpreted in this way:

(35) **ir** $(p^n)(x_n)(x_{n-1})...(x_1) = (p^n)(x_n)(x_{n-1})...(x_1)$ & $\text{ben}(p^n,x_{n+1})$ [=(14)]
where $\text{ben}(p^n,x_{n+1})$ means "x_{n+1} is the beneficient of the action p^n", and p^n is an n-place predicate.

Written as an n-tuple category, **ir** has the following category:

(36) $< (S/A_1A_2), (S/A_1A_2A_3), (S/A_1A_2A_3A_4) / (S/A_1), (S/A_1), (S/A_1A_2),(S/A_1A_2A_3)>$ [=(12)]

A_i is an abbreviation for any type of argument category, NP, PP, S' and perhaps others.

The passive operator **PASS** has the following universal analysis:

(37) a. $(S/A_nA_1...A_{n-1}) / (S/AA_1...A_n)$ $n = 0,1,2,3$ [=(21)]
b. $\mathbf{PASS}(p^{n+1})(x_{n-1})...(x_1)(x_n) = (\exists y)(p^{n+1})(x_n)...(x_1)(y)$

In other words, **PASS** existentially generalizes the former subject and"promotes"the last argument to subject.

It is claimed that this treatment can explain the interaction between **APPL** and **PASS**. The authors say that, in Kinyarwanda, the direct object, the indirect object and the applied object (the beneficient=AO) can invariably be promoted to subject. Thus, from the form (38a), we can derive (38b) to (38d) by passivization:

(38)a. Umugore	a -ra -he -er -a	umugabo	imbwa	ibiryo [=(16b)]
woman	she-Prs-give-**APPL**-Asp	man	dog	food
SU		AO	IO	DO

b. umugabo a -ra -he -er -w -a imbwa ibiryo [=(19a)]
man he -Prs-give-**APPL**-**PASS**-Asp dog food (AO promoted)
4-->3

c. imbwa a-ra -he -er -w -a umugabo ibiryo [=(19b)]
dog it-Prs-give-**APPL**-**PASS**-Asp man food (IO promoted)
3-->2

d. ibiryo a-ra -he -er -w -a umugabo imbwa [=(19c)]
food it-Prs-give-**APPL**-**PASS**-Asp man dog (DO promoted)
2-->1

Note that **PASS** does not simply affect the arguments of the verb, but that it rather operates on a phrase. For instance, in (38c), the phrase [gives-for man] is passivized. This is a 3-place predicate. **PASS** converts this into the 2-place predicate $\lambda x \lambda y$ [Someone give for man x_{IO} y_{DO}]. This is indicated here by the notation 3 --> 2, and analogously for the other cases. **PASS** operates on a phrase though it is a bound morpheme. As the reader may convince himself, the assumption that there are morphemes which modify phrases is essential for the approach. It is not possible to analyze (38c) and (38d) by means of an operator that simply affects the argument structure. If the facts are correctly reported, then they represent a serious challenge for any kind of strong lexicalistic theory such as has been advocated, e.g., by Di Sciullo & Williams (1987). Keenan & Timberlake say nothing as to the question of how they conceive of the relation between syntax and morphology. Clearly, their account is at odds with surface oriented categorial theories of morphology like that advocated by Moortgat in the same volume.

Beside this question, there are empirical problems with their approach. Thus, Baker (1988, p.407 ff.) disputes some of the linguistic data. According to him, at least (38d) should not be possible. In addition, Baker claims that the relative order of **APPL** and **PASS** must always be as in these examples, in other words, there are no cases where a passivized verb undergoes applicativization. This follows from his theory, but it is not implied by Keenan & Timberlake's account. There is no reason, why we shouldn't find the form *give*-**PASS**-**APPL**. For instance, we could express the same meaning as does (38c) by means of the following analysis:

(38') c. dog give-**PASS**-**APPL** food man
3 -->2

If Baker is right, this is not possible, however. The relevant examples (due to Kemenyi 1980, quoted from Baker 1988), are:

(39) a. Ibaruwa i -ra -andik-**ir** **-w** -a umukoobwa n-umuhuungu
letter SP-pres-write-**APPL-PASS**-asp girl by-boy
b. *Ibaruwa i-ra-andik-w -ir -a umukoobwa n-umuhuungu
PASS-APPL

Consequently, Keenan & Timberlake would have to exclude (39b) by morphological stipulation.

Note further the following descriptive inadequacy of the universal passive rule:

(40) Otto is respected by everyone

If the passive operator always existentially generalizes the subject, then it cannot be universally quantified by *everyone* in a *by*-phrase, as (40) requires. The authors remain silent on questions like these which have traditionally worried linguists working on these phenomena.

A criticism of a similar kind applies to all other proposals made in this paper. An example is the authors' treatment of *tough*-movement. Consider their meaning rule for *difficult*, which is roughly this:

(41) $\mathbf{difficult}(p^n) = \lambda x_n \lambda x_2 ... \lambda x_{n-1}\ \mathbf{DIFFICULT}(\exists x_1 p^n(x_1,...,x_n))$

Thus, **difficult** converts an n-place predicate phrase into an n-1-place one. The semantics is analogous to that of **PASS**, the difference being that the passivized sentence is imbedded under the sentential operator **DIFFICULT**. This correctly describes the meanings of the following sentence pair:

(42) a. Flowers are difficult to give to John
b. John is difficult to give flowers to

But the approach cannot block the ungrammatical sentence (43), which means the same as (42b).

(43) *John is difficult to give flowers

((43) can be derived from "It is difficult to give John flowers".) Finally, (41) presupposes that the complement of *difficult* cannot have an overt subject, which is contrary to the facts:

(44) John is difficult for Mary to talk to

Keenan & Timberlake could react by saying that (44) requires something additional and that the blocking of cases like (43) should be achieved by syntactic constraints which are disregarded in their context of research. As far as I can see, it is far from trivial to say what these constraints are. On the other hand, the meaning rules proposed are rather trivial, I believe. So, the really tough questions are left out in this paper.

Let me conclude my comments by a remark on the mentioned *n-tuple Universal*(p.293):

(45) If $<C_1,...,C_n>/<D_1,...,D_n>$ is an *n*-tuple category in the grammar of a possible natural language, then there is a uniformly definable function *g* from Cat into Cat such that for all *i* between 1 and *n*, $C_i = g(D_i)$.

The authors' intention is that "C_i differs in form from C_i in the same way that C_j differs in form from D_j." They illustrate the principle by means of the n-tuple category for the passive, where C_i has the form $g(S/AA_1...A_n) = (S/A_nA_1...A_{n-1})$. In this particular case, g means that the subject is absorbed and the lowest argument is promoted to subject. The other n-tuple categories considered can be analysed in an analogous way. Thus, the constraint is an attempt to account for the intuition that each n-tuple category reflects a uniform syntactic process. Surely, this is not a particularly revealing account.

4.3.1.3. Another central paper for the syntactic theory of CG is Mark Steedman's "Combinators and Grammars". Steedman was among the pioneers that caused the comeback of CG. In earlier contributions, he applied FC for an analysis of long distance dependencies. Since the idea how this can be done by means of FC has been discussed in my comments to Huck's paper, it will not be repeated here. The most interesting question taken up by Steedman is variable binding. He considers constructions that involve abstraction over more than one variable, and he asks what kind of operations are possible in natural language. His answer is that the operations of natural languages reduce to certain combinators of combinatory logic. They make variable binding unnecessary, when there is no open pronoun. This point has been taken up by a number of categorial grammarians, notably Szabolcsi(1987), where it is claimed that natural languages never have bound variables in syntax.

I believe that Steedman's way of looking at variable binding obscures the issue. In the following I want to show that Steedman (and Anna Szabolcsi) should rather say that there is no *vacuous* variable binding in natural language. Their arguments amount exactly to that claim. Stated in this way, however, most linguists would agree. For instance, Chomsky (1982) has stated a similar principle.

The crucial examples considered by Steedman are constructions with parasitic gaps:

```
(46) (articles) which I    will    file    without       reading   [=(16)]
                NP    NP VP|VP  VP|NP  (VP|VP)|Cing  Cing|NP  FC
                                       ____________(VP|VP)|NP  S
                               _____________VP|NP  FC
                          _________VP|NPFC
                 _______S|NPFA
                       S
```

The categories are undirected. Cing is S|NP. Note that the analysis cannot be entirely correct as it stands because, for semantic reasons, the relative clause should not be an S but rather an S|NP. But this can be repaired one way or the other. The important new principle is the substitution rule S which projects the DO-gaps of *file* and *reading* to the phrase *file without reading* and binds the two. In other words, the semantics of S makes sure that the phrase means **λx[(without reading' (x))(file'(x))]**. Hence, the syntax and semantics of **S** can be stated as follows:

(47) *Substitution* (S) [=(18)]

(C|A)|B + A|B==> C|B

f **g** **λx[f(x)(g(x))]**

Recall that the system is undirected. If we wanted to incorporate the surface order, we have to replace this rule by a number of directed rules. In any case, this rule describes the facts correctly.

Now, Steedman notices that S is a well-known combinator in combinatory logic, whose semantics is determined by the following equation:

(48) $\mathbf{Sfg = \lambda x[f(x)(g(x))]}$ [=(37)]

Another combinator **B** performs **FC**:

(49) $\mathbf{Bfg = \lambda x f(g(x))}$ [=(33)]

Furthermore, there is the *commuting operator* **C**, which permutes two adjacent arguments of a function:

(50) $\mathbf{Cf = \lambda x \lambda y f(x)(y)}$ [=(35)]

Finally, we have a *reflexivization operation* **W**:

(51) $\mathbf{Wf = \lambda x f(x)(x)}$ [=(36)]

In combinatory logic, there are two other constant operators, namely the *identity operator* **I** and the *cancellation operator* **K**, whose meanings are given by the two next equations:

(52) $\mathbf{Ix = x}$ [=(39)]

(53) $\mathbf{Kxy = x}$ [=(40)]

The thesis advocated by Steedman can now be made precise: The combinatory rules of natural language only make use of the combinators **S**, **B**, **C** and **W** but not of **I** and **K**.

Steedman asks the question why natural languages use combinators instead of open variables. His answer is: "...it looks as though natural languages are trying to do without bound variables".

Is this really so? I don't think so. The only difference between a combinatory system of this sort and say a GB-approach is that we perform the binding operation on different levels, in the syntax and in the metalanguage, respectively. It is obvious that we can reformulate most of the categorial combination rules by means of combinators. We have seen that for **FC** and *Substitution* already. And conjunction can be defined, too. It is the operator **S'** defined by the following equation:

(54) $\mathbf{S'\ c\ a\ b\ x = B(BS)B\ c\ a\ b\ x}$ [=(41)]

Since Steedman claims that this kind of grammar is computationally very simple and intuitive, I would like to convey a bit of his enthusiasm by evaluating the special case of VP-conjunction. It runs like this:

(55) $\mathbf{B(BS)B\ and'\ VP'\ VP'\ x}$

$\mathbf{=\lambda y[(BS)(B(y))]\ and'\ VP'\ VP'\ x}$	Def. **B**
$\mathbf{= (BS)(B(and'))\ VP'\ VP'\ x}$	λ-conversion
$\mathbf{= \lambda y[S(B(and')(y))]\ VP'\ VP'\ x}$	Def. **B**
$\mathbf{= S(B(and')(VP'))]\ VP'\ x}$	λ-conversion
$\mathbf{= S(\lambda y[and'(VP'(y)])\ VP'\ x}$	Def. **B**
$\mathbf{= \lambda x[\lambda y[and'(VP'(y)](x))\ VP'(x)]}$	Def. **S**
$\mathbf{= \lambda x[and'(VP'(x))\ VP'(x)]}$	λ-conversion

This works beautifully, and I have no fair objection. (I doubt that I would do it this way, but this is not a legitimate objection.)

What is more important is that, as far as I can see, all genuine variable binding can be simulated by means of the operators of the first group, *viz.* by means of **S**,**B**,**C** and **W**: The operation **S** allows binding two structurally independent positions; the permutation operation **C** allows us to place any argument we like at the last position of the function; **B** connects functions; and **W** does the binding within one function. That's all we need in order to bind variables. This has been realized a long time ago by Schönfinkel(1924) and by Quine(1960).

The operations **I** and **K** are needed only, if we allow for vacuous variable binding and formation of constant functions via abstraction. To give an example: The abstract **λx[y]** has to be represented as **Kxy** in combinatory logic and the abstract **λx[x]** is represented as **I**xx. The former is vacuous binding and the latter is a constant function. If we don't allow these cases in the syntax, then we do not need these operations.

As Steedman has pointed out, these cases don't seem to occur in natural language. Thus, the correct claim should be that there is no vacuous binding in natural language (and no formation of constant functions *via* abstraction). But it doesn't follow from this that there is no variable binding in natural language. Quite the opposite is true. Another way of looking at combinatory logic is that the combinators are made precisely for achieving variable binding. We need no more and no less combinators than we need for doing variable binding: The combinators are variable binders.

There is an obvious objection to this criticism: No variables, no variable binding. This objection is valid in letter, but not in spirit. We may take the operators as primitives, of course. Then the objection is valid. By I think this runs entirely against our semantic intuitions. Try to teach the semantics of combinators without recurring to variable abstraction in the metalanguage. It think that would be a hopeless endeavor. On the other hand, we quickly understand the mechanism of variable binding. We understand it, because it is deeply rooted in our semantic intuitions. Frege was the hero who made these intuitions precise. In personal communication, Ede Zimmermann has pointed out to me that it is hardly a historical accident that Schönfinkel invented the reductionistic program of combinatory logic after the invention of bound variables by Frege. Only then could he check his intuitions at each step by recurring to equivalent expressions with bound variables. (It might be interesting to recall at this point that Quine (1960) didn't call his paper "Variable binding explained away" but rather "Variables explained away". If one could dispense with bound variables once and for all then Quine's principle "To be is to be the value of a bound variable" could be eliminated as well.)

To be sure, Steedman admits that there are bound variables in syntax, for instance, in the case of overt bound pronouns: "It would be perverse to argue that these were *not* realizations of bound variables as it is to argue for 'invisible' bound variables or empty categories in the constructions considered here." (p. 437). In recent papers, Anna Szabolcsi argues precisely for the first kind of perversion (*vide, e.g.*, Szabolsci 1987). It should be clear that I am at least partly a pervert of the second kind. Theories in the GB-style that assume empty categories, which are interpreted as bound variables, seem to me much simpler than this kind of approach, both technically and conceptually. Why shouldn't we have a semantically transparent representation somewhere in the syntax, if we need it anyway?

I should add at the end that the kind of criticism put forward against Keenan & Timberlake's paper also applies to Steedman's article: The theory doesn't restrict the occurrences of parasitic gap enough. Chomsky (1986) has given arguments that a parasitic gap can't be embedded to deeply. He accounts for this fact by movement of an empty operator and complex chain formation. The subtle issues involved here are not touched in the paper. Categorial principles or the principles of combinatorial logic can't answer them.

4.3.2 Morphology

4.3.2.1. Michael Moortgat's article "Mixed composition and discontinuous dependencies" makes use of the same techniques as Huck's article, especially of the disharmonic version of functional composition, *viz.* the rule **FC***. As far as I know, Moortgat has been among the first persons who have applied these techniques to morphology. Thus, his work might claim historical priority with respect to the work of many others. The article is mainly concerned with bracketing paradoxes in morphology. These are solved by means of functional composition. Furthermore, Moortgat claims that the verb cluster in Dutch and German should be analyzed by the very same methods.

The first claim is that functional composition can explain *morphological restructuring*. Consider the example *Spielerin* "female person who plays" from German.

```
(56)            ______N______FA
          ___N___FA
           V   V\N      N\N
         spiel  er        in
           V   V\N_____N\N FC
          _________V\N FA
```

As the figure shows, there are two ways of analyzing the word, the "normal" way that only uses **FA** and the"fancy" way that first forms the complex suffix *-erin* by means of **FC** and applies it to the stem. Semantically, both derivations amount to the same.

Examples like these show that we *can* restructure a word by forming a complex suffix. Now, Moortgat claims that sometimes we *have* to restructure a word in this way. His example is the German word *Gebärerin* "female person who gives birth". Moortgat says that, for biological reasons, there is no word *Gebärer* "male person who gives birth". According to Aronoff's (1976) dictum, words have to be derived from existing words. Therefore, *Gebärerin* must have the structure *gebär*$_{V}$ + *erin*$_{V\backslash N}$, where the the complex suffix is formed by means of **FC** in the way indicated.

I find this particular example unconvincing because the German sentence *Zeus ist der Gebärer von Athene* (Jove is the male person who gave birth to Athena) is well-formed and true. On the other hand, the point Moortgat wants to make is clear. It is supported by further examples such as:

(57) Schamlos+igkeit vs. *schamlos+ig, Ruhelos+igkeit vs. *ruhelos+ig, Leblos+igkeit vs. leblos+ig, Kinderlos+igkeit vs. *kinderlos+ig

The suffix *-ig* is an adjectivizer, whereas *-heit* is a nominalizer. Since the *A+ig* forms don't exist and since, by Aronoff's dictum, words are derived from existing words, the suffix has to be *-igkeit* which is of category A\N. Comparison with Dutch, which still has some *ig*-adjectives, shows that German originally had a suffix *-ig* of category A\A but formed the A\N *-igkeit* by functionally composing *ig+keit*.

I have several worries about this. The first concerns Moortgat's claim that words like *Schamlosigkeit* "impudence" are no longer derivationally transparent. My intuitions point in the opposite direction. It is an idiosyncrasy of the suffix *-keit* that it only embeds a particular class of derived adjectives. Compare the following examples:

(58) a. $[[[$ Schamlos $]_A$ + ig $]_A$+keit $]_N$ "impudence"
b. $[[[$ Kost $]_N$ + bar $]_A$ + keit $]_N$ "preciosity"
c. $[[[$ Köst $]_N$ + lich $]_A$+keit $]_N$ "delight"
d. $[[[$ Furcht $]_N$ + sam $]_A$ + keit $]_N$ "fear"

Besides the idiosyncrasy mentioned, the morphological pattern is entirely transparent: *-keit* invariably nominalizes a derived adjective. Here, Moortgat can't argue that *-barkeit*, *-lichkeit*, *-samkeit* are derivationally nontransparent suffixes, because the intermediate forms *kost+bar*, *köst+lich*, *furcht+sam* are German words. Hence, for the derivation of these examples *-keit* must have the category A\N. This again raises the question how the application of *-keit* to nonderived adjectives is prevented in CG. It seems to me that this can't be done in a *non-ad hoc* way.

The conclusion I draw from difficulties like these is that Aronoff's principle – that words are only derived from existing words – should be abandoned. It think, Di Sciullo & Williams (1987) are right when they say that the principle conflates two different notions of word, the morphological and the lexical notion. A structure is a *morphological word* if it obeys the morphological well-formedness principles. It is a *psychological word* if it is stored in the lexicon, the latter being a psychological concept. (According to Di Sciullo & Williams, the lexicon is a messy area with no particular grammatical structure. Hence, there can't be a grammatical theory of the lexicon.) Following this distinction, there is no reason why *Gebärerin* should not have the structure $[[[\text{gebär}]_V\text{+er}]_N\text{+in}]_N$, even if the intermediate $[[\text{gebär}]_V\text{+er}]_N$ did not exist. The latter would still be a morphological word, though not a lexical word.

The same holds for the examples (57): The intermediate *A+ig* forms are morphological words but not lexical words. Note that it is not true that there are no *A+ig* adjectives at all: *wahrhaft(ig)* "sincere", *leibhaft(ig)* "in person". Since the the suffix *-ig* has no meaning, these forms are felt to be redundant, or they are lexicalized. In any case, the intermediate nonlexical words *could* be German lexical words, and it seems to me that at least some of them exist in dialects.

To summarize, I see no reason why the examples (57) should not be analyzed as in (58a). The examples (58) are entirely regular as to their morphological structure, and a good theory should capture this regularity.

The most promising domain where principles of functional composition might apply is the theory of *argument inheritance* in morphology. Consider the following examples of Moortgat's:

(59) a. tevreden+heit met de soep
contentness with the soup
b. vergelijk+baar met wijn
comparable with wine

In both cases, the complement is semantically inside the scope of the affix, syntactically outside of it. This should be clear from the English paraphrases. Thus we face a bracketing paradox. The solution in CG is straightforward: FC* projects the argument of the stem to the derived word as witnessed by the analysis of (59b):

(59') vergelijk baar met wijn
TV/PP TV\A FC* PP
A/PP FA
A

Semantically, this analysis is certainly correct. But is it really as simple as one would believe at first sight? Recall that ,for the reasons given above, the morphological structure of *vergelijk+baar* should be $[[\text{ vergelijk }]_V$ baar $]_A$. Now, Moortgat's analysis doesn't quite give us that. It rather contains two operations that can be made visible in terms of movement in the following way:

(60) vergelijk t_i (t_j) bar met wijn (Moortgat)

V_{st} PP NP A_{aff} PP_i

\ | / / /

V / /

A /

AP

t_i is the PP-gap projected by means of FC*, and t_j is the gap of the DO, which is raised to subject according to the semantics of the passivizer *-bar*, whose meaning may be thought of as **λx.possibly'(∃yTV'(y,x))**. To be sure, Moortgat has to assume only the PP-gap, because the NP-gap is not encoded in the categories. I have represented it in order to be able to compare his treatment with other proposals.

Moortgat criticizes an alternative treatment due to Fabb(1984) and Pesetsky(1985). In this approach, the affix is moved in logical form to its semantically appropriate place. Thus, the analysis of (59) is something like this:

(61) [[[vergelijk$_{V_{st}}$ t_i] t_j met Bill] bar_i]

I take it that the trace t_i of the affix is ignored for the interpretation. t_j is the trace of the DO raised to subject. For the rest, the interpretation of the structure should proceed in a way similar as before.

Moortgat criticizes that this approach doesn't generalize to conjunctions, because examples like these:

(62) a. their [preparedness and willingness] to start the fight
b. John's [reluctance or inability] to accept the offer

The problem is that we have two affixes, which may even be different, as in (62b). This seems incompatible with Fabb and Pesetsky's solution. I am not sure whether this criticism is cogent, because it might be argued that these constructions are elliptic ("reluctance to accept the offer or inability to accept the offer"). If they are not, we still can move the sentential complement across the board. Thus, the LF of (62b) would be something like (63):

(63) John's [[{ [reluct t_i] t_j] ance$_i$] or [inabil t_k] t_j] ity $_k$ } [to accept the offer] $_j$]

In other words, I don't believe that Fabb and Pesetsky's approach faces descriptive difficulties that can't be overcome. The question remains, which of the two theories is the correct one: Do we move the PP over the affix or do we move the affix over the PP? The first alternative is Moortgat's, the second is Fabb and Pesetsky's. Stated that way, the answer is not at all obvious, although Moortgat's treatment seems more appealing at first sight.

Note that there is a third alternative, namely the theory of Baker(1988), according to which one could analyse (59) as (64):

(64) [$_{AP}$ [$_A$ vergelijk$_i$ bar$_A$] [$_{VP}$ t_i t_j [$_{PP}$ met wijn]]]

t_i is the trace of the verb stem incorporated into the suffix *-bar*, and t_j is the trace of the DO, which is raised to subject.

Viewed from a larger perspective, the question is not wether bracketing-paradoxes are to be treated categorially or not. The question rather is which of the three alternatives is the correct one. In a language like English, which allows for preposition stranding, a lexical treatment in the style of Moortgat is difficult to maintain in view of constructions such as the following (cf. Kayne 1984, p. 140):

(65) a. The existence of stranded prepositions is not accountable for under Moortgat's assumptions.
b. This book is readable by a 10-year old.

I can't see how (65a) can be treated in a purely lexical way. And the existence of the *by*-phrase represents a general difficulty for a lexical treatment of the passive as well (see the discussion of Keenan & Timberlake's paper).

In any case, the arguments given by Moortgat that bracketing paradoxes can't be treated at all by a syntactic approach are hardly conclusive in view of what has been said in connection with Huck's and Dowty's contributions: Any categorial principle of argument inheritance can be simulated by appropriate movement rules, perhaps across the board.

Let us consider next Moortgat's analysis of the Dutch and German verb cluster. Moortgat's argument is that the conjoinability of verb clusters shows that they are constituents. For instance, the verb cluster *will* *proberen te lezen* "wants to try to read"has the following analysis:

```
(66)  will     proberen te lezen
      VP/VP   VP/VP   NP\VP FC*
      ____________NP\VPFC*
           NP\VP
```

In order to block the ungrammatical

```
(67) *omdat  hij  het boek probeerde  op de tafel   te leggen
      because he  the book tried      on the table  to lay
                  NP       VP/VP      PP            PP\(NP\VP)FC
                           ________________ (NP\VP) FC*
                  _______________ (NP\VP)FA
                  VP
```

Moortgat has to recur to the *ad hoc*-stipulation that the verb cluster is formed by a "lexical rule". Moortgat is not explicit about what exactly he means by that. He merely says that "the lexical component is equipped with C [=FC]" (p.335). In contradistinction to German, Dutch assumes the disharmonic principle FC*. But then I see no way to block the derivation of this string in CG, given the categorization given above. Note further that there are Germanic dialects were the the construction (67) is grammatical, *viz.*, in West-Flamish and Swiss dialects. For relevant arguments that the formation of the verb cluster must be syntactic in nature, *vide* Haegeman & Riemsdijk (1986).

The last bracketing paradox considered by Moortgat is the formation of adjectives like blaw+[oog+ig] "blue eyed". The categorial analysis involves Montagovian and Geachian type raising:

```
(68)  blaw    oog              ig
      N/N     N                N\A
               |M              |G
           (N/N)\N    ((N/N)\N)\((N/N)\A) FA
      __________((N/N)\A) FA
              A
```

The type transition from N\A to the complicated type is in the lexicon, according to Moortgat. In terms of movement, this representation means the following:

(69) $[\,[\text{blaw}]_{A_i}\,[[\,t_i\ \text{oog}\,]_N\ \text{ig}\,]_A]_A$

It seems to me that we need a general theory that tells us why this should be so. I find it not very attractive to put the information that the modifier has to be moved out of the noun it modifies into the functor *-ig*. Something more general seems to be involved. I have to add that Moortgat does not discuss the question why *blawoogig* should not have the semantically transparent structure ((blaw+oog)+ig). One wonders how this follows from categorial principles, since this structure follows from the most natural categorization of the morphemes involved.

At the end of the article, Moortgat discusses the possibility whether functional composition can be confined to "the lexicon". It does not come as a surprise that he does not find this possible, because we have long distance dependencies like extraposition. And these require functional composition, the analogon of a syntactic movement rule as we know from the previous discussion.

A system with **FA, M, G, FC** and **FC*** is another instance of *Moortgat's nuisance*: It is closed under permutation. This is proved at the end of the article.(I am not sure that the proof is complete, as it stands. It seems to presuppose Lambek's sequential types (A.B), which are not assumed by Moortgat. But I might be wrong with this conjecture. Furthermore, I blindly believe the result, which is presumably caused by the disharmonic rules **FC***.) In other words, the system has to be restricted. One brake against combinatorial explosion is the lexicon where certain type transitions are explicitly listed, as in the *ig*-case. I have commented on this above. Another brake is the type driven interpretation advocated by Partee & Rooth (1983): We always take the lowest types that are needed to avoid a type clash (cf. p.322). I can't see that this principle can prevent closure under permutation: Why shouldn't we admit as many type transitions as we need for closure under permutation? To avoid this consequence, we have to know that permutation is not a goal to be striven for. But how do we know? The grammar should tell us, but CG doesn't.

4.3.2.2. Jack Hoeksema's and Richard D. Janda's "Implication of Process-Morphology for Categorial Grammar" is a less standard approach to categorial morphology. The authors relate most of the current work on morphology to Hockett's 'Item-and-Arrangement' model, according to which words are built up by the concatenations of smaller parts. They maintain that stronger algebraic operations are needed besides concatenation, *viz.* context-sensitive addition, permutation, replacement, and subtraction. Hoeksema and Janda integrate this richer approach into Hockett's 'Item-and-Process'-model. Before I discuss the theoretical implications of this view for categorial grammar, I illustrate these claims by some examples.

In Chamorro, singular agreement between the subject and the verb is expressed by infixing *-um-* after the first consonant of the verb, whereas the plural is marked by the prefix *man-*.

(70) gupu ==> g-**um**-upu
fly fly (singular)

==> **man**(g)-gupu
fly (plural)

This is an example for contextsensitive addition.

An example for morphological metathesis, *i.e.* permutation, is the derivation of the incomplete form of nouns from the complete one in Rotuman (p.227):

(71)	famóri	"people" (complete)	fä$^?$éni	"zealous" (complete)
	famóir	r-metathesis	fä$^?$ein	i-metathesis
	famör	Umlaut oi>ö	fä$^?$en	ei > e
	(incomplete form)		(incomplete form)	

An example for substitution is the formation of elative adjectives in Javanese: The elative is formed by replacing the final vowel of an adjective by *u*, if it is rounded and by *i* otherwise (p.233). Thus, we have:

(72)	gloss	primary	elative
	"difficult"	angel	angil
	"easy"	gampang	gamping
	"heavy"	abot	abut

Let us look at how Hoeksema and Janda analyze the last case. The relevant rule is the following:

(73) $< A_{prim}, A_{elat}, f_{elat} >$,

with

(74) $f_{elat}(XV[x\ round]C_0) = XV[+high,x\ back]C_0$ $(x = +,-)$.

It follows that "categorial rules" have the general format < A, B, f> , where A is the input category, B is the output category and f is the syntactic operation that brings A to B.

It seems to me that this account trivializes the theory for the following reasons. First: Any syntactic rule can be brought into this format, because this is nothing but Montague's algebraic formulation of syntactic rules, and we know that Montague´s *Universal Grammar* has no restrictions built in. In Montague's syntax we can do anything we like. Hoeksema and Janda are aware of this. On page 241 they say about constraints: "We are convinced that such limitations exist, but we do not believe that they must be direct consequences of one's descriptive framework." My reaction to this is that their formalism is indeed no more than a descriptive framework.

The second objection is that I can't see what this account has to do with categorial grammar. One of the basic tenets of CG is that an expression always breaks down into a functor and arguments. But if we look at rule (73), then we don't find a functor in the expression itself. The functor is the operation f_{elat}; however, this is not an expression but an operation. To put it differently: If this is categorial grammar, then any grammatical theory is categorial grammar and, among other things, also GB-theory. My guess is that this consequence is not what the authors had in mind.

4.3.3. Phonology

The only contribution to categorial phonology is Deirdre Wheeler's "Consequences of some categorially motivated phonological assumptions".

In Russian, we have the well-known process of *final devoicing* (**FD**): An obstruent is devoiced at the end of a syllable. There is another process of regressive assimilation called *voicing assimilation* (**VA**): A final obstruent in a C-cluster passes the feature [α voice] to all obstruents on its left, where this assimilation is not blocked by an intervening sonorant. Furthermore, **VA** ignores syllable boundaries. The following data show that **FD** feeds **VA**.

(75) a. **iz** Mcensca [smc] **FD+VA** "from Mcensk" [=(8)]
b. ot **mzdy** [dmzd] **FD+VA** "from the bribe"

However, there is a consonant, viz. /v/, which doesn't fit into this picture.

(76) a. trezv [zf] "sober" [=(11)]
b. xorugv' [gf'] "banner"

Thus, /v/ undergoes **FD**, but it does not trigger **VA**. On the other hand, /v/ undergoes **VA**, as the following data show:

(77) a. korov+ka [fk] **FD** "cow (dim.)"
b. ot vdvy [dvd] **VA** "from the widow"

The problem is how to account for this exceptional behavior of /v/. The solution of this apparent paradox advocated by Hayes (1984) is that [v] is derived from an underlying sonorant /w/. In Haye's account, sonorants undergo **FD**. Then **VA** applies. The labial sonorant [w]/[W], where the latter is the unvoiced counterpart, is strengthened to [f]/[v] by a rule called *W-Strengthening* (**WS**). The other sonorants are revoiced by a process called *Sonorant Revoicing* (**SR**). The data mentioned are deduced in the following way, where devoiced sonorants are represented by capitals:

(78)	/korow+ka/	/ot wdwy/	/iz mcenska/	/trezw/	**UR**
	koroWka	______	is mcenska	trezW	**FD**
	_______	od wdwy	is Mcenska	_____	**VA**
	korofka	od vdvy	is Mcenska	trezf	**WS**
	______	______	is mcenska	_____	**SR**

Wheeler objects that this analysis violates two general principles of CG:

(79) *Compositionality*: The interpretation of the whole is a function of the interpretation of the parts.

(80) *Invariance*: Once a feature or set of features has been specified in the phonetic interpretation, it may not subsequently be changed.

Transformational processes like **FA** and **VA** are supposed to violate compositionality, and devoicing of sonorants with subsequent revoicing clearly is a violation of the invariance principle.

Let me first comment on the compositionality requirement. It is difficult to understand what Wheeler exactly means by this. Like any transformation, Hayes' rules can be reconstructed as functions that map arguments of a particular kind to values of some other kind. I cannot see why operations are not compositional (*vide* the discussion of Hoeksema & Janda's article for a related point). If one compares Wheeler's rules discussed subsequently with those of Hayes one detects that they have virtually the same structure. It seems then that the compositionality requirement in (79) is too vague to determine a choice between the different accounts and I will disregard it in what follows. I will rather concentrate on the invariance principle.

Let us consider now Wheeler's categorial treatment of the Russian data. She assumes that a syllable of the form CVC has the categorial structure

```
(81)        S
           / \
       S/N     N
        |    /   \
        |  N      N\N
        C  V       C
```

N is the nucleus category, S stands for syllable, N\N is a part of the coda and S/N is a part of the onset. Onset and coda can be made longer by the categories (S/N)/(S/N) and (N\N)\(N\N) respectively.

The phonological interpretation of the rules is the following:

```
(82)   a. N + N\N ==> N     (offset rule)        [=(14i)]
          a    b      c
```
PI(c) is PI(a)+PI(b), with the value for [voice] left uninterpreted for PI(b).

```
       b. S/N + N  ==> S     (onset rule)         [=(14ii)]
           a    b      c
```
PI(c) is PI(a)+PI(b)

PI means "phonetic interpretation". I have certain problems with the interpretation of these rules to which I will return in a moment.

An essential assumption of this system is the

(83) *Universal Marking Convention* (**UCM**): The unmarked value of obstruents is [-voice] and that of sonorants is [+voice].

If I understand Wheeler correctly, she assumes that there are two kinds of underlying obstruents: On the one hand we have the unvoiced obstruents, which have the feature [-voice] and which are represented by small letters. On the other hand, there are the unmarked obstruents which do not contain a [± voice]-feature. These are represented by capital letters.

Before I comment on the underlying idea, let me illustrate how this works by deriving the *p~b*-alternation in *klub* [p]/*kluba* [b] "klub nom./gen.".

```
(84)          [kluB]_S ==> [klup]  by UMC     [=(15)]
               /   \
              /     [uB]_N
             /      /  \
      /kl/_S/N  /u/_N  /B/
```

(I should note that Wheeler represents unmarked obstruents sometimes by small letters in the terminal string. I take it that this is not quite in the spirit of her account.) In this derivation, nothing happens. Therefore, UMC can apply at the end and can add the feature [-voice] to the feature matrix of /B/, yielding the correct output [p]. The invariance principle is not violated.

Before I return to the derivation of [kluba], let me comment on a problem concerning Wheeler's onset rule.

```
(85)          [viSK]_S ==> [visk]  by UMC     [=(16)]
                /  \
               /    [iSK]_N
              /      / \
         /v/_S/N /i/_N /SK/_N\N
```

I take it that the onset rule (82b) does not affect the features of the input. If this interpretation is correct, it follows that Wheeler has to assume a third series of underlying consonants, *viz.* voiced consonants. Clearly, this consequence is unwelcome for reasons of economy. One would rather have expected that the onset rule voices unmarked obstruents, whereas it doesn't affect unmarked sonorants. Furthermore, my interpretation of the rules makes the qualification for the offset rule redundant: If there is no voice-feature, the rule cannot affect it. But, perhaps, the theory is intended in the way discussed here, or at least it could be revised along these lines. If this were so, [kluba] could be derived from /kLu+BA/. The revised onset rule would voice /B/ to /b/, whereas /L/ would remain unaffected and voiced by **UCM**.

In order to account for the **VA**-data, Wheeler assumes that syllables can be phonological words (W) and that a phonological word may modify another one:

(86) W/W + W ==> W [=(17b)]
a b c
PI(c) is **VA***(PI(a)+PI(b)),

where **VA*** iteratively applies from right to left and is defined as follows:

(87) $[+\text{cons}] \to [+\text{voice}] / \_ \begin{bmatrix} +\text{cons} \\ +\text{voice} \end{bmatrix}$ [=(18)]

Note that, in contradistinction to Haye's rule **VA**, Wheeler's rule **VA*** is a regressive voicing assimilation, *i.e.* it never devoices. Let us have a look how the data (78) are derived.

(88) $[\text{koRoVka}]_W$ ==> [korofka] by **UCM** [=(24a)]
/ \
$[\text{koRoV}]_{W/W}$ $[\text{ka}]_W$

(89) $[\text{odvdovy}]_W$ **VA** [=(19b)]
/ \
$[\text{oT}]_{W/W}$ $[\text{vdovy}]_W$

In order to derive [trezf] from [treZV], we need a further rule:

(90) N\N + (N\N)\(N\N) ==> N\N [=(22)]
a b c
PI(c) = PI(a)+PI(b), with the value for [voice] left uninterpreted in PI(b).

If we interpret this rule according to the lines discussed above, we obtain the following derivation:

```
(91)        [tRezV]_S ==> [trezf]  by UCM
               /   \
              /    [ezV]_N
             /     /   \
            /     /    [zV]_N\N
           /     /     /     \
      /tR/_S/N  /e/_N  /z/_N\N  /v/_(N\N)\(N\N)
```

So, the data come out correctly, and we may turn to an assessment of the approach. My general impression is this: Whatever the merit of this analysis may be, it is entirely independent of categorial principles.

Consider the syllable structure first. The theory boils down to the claim that syllables have an onset, a nucleus and a coda - widely shared assumptions. The novelty of the theory is that the coda is divided into a "central" coda (N\N) and a postcoda (N\N)\(N\N). The formulation in terms of categories seems artificial to me, because consonants are not natural functors: One and the same consonant, say /t/, can both occur in the onset and in the coda. Thus, it may have the categories S/N and N\N.

As to invariance, this principle is independent of the framework as well. One could state this principle for any phonological theory whatsoever. In more conventional terms, the analysis seems to be this: 1. Unmarked obstruents are voiced at the onset. 2. Unmarked obstruents are voiced in the "central" coda. 3. **VA** applies at the level of the phonological word. 4. **UCM** applies postcyclically.

Let us comment finally on the empirical predictions of the theory. On p. 486 Wheeler writes: "So, the prediction is that languages which have a regressive voicing assimilation process operating between syllables, like Russian, must also have final devoicing."

This is an interesting prediction. It is not directly falsified by classic Arabic which has a regressive devoicing rule but no final devoicing. Consider the following alternation (I am indebted to Cristoph Correll for the examples):

(92) a. yak+tub [b] (an imperative form)
b. ka+tab+ta [ft] (a perfect form)

If we ignore the fricativization of the [b] in the second example, the (intermediate) form [katapta] is correctly derived from [kataBta] *via* the regressive devoicing rule, given Wheeler's assumptions. In order to block the derivation [yaktuB] ==> *[yaktup] *via* UCM, Wheeler presumably has to say that unmarked obstruents have to be voiced in final position. It seems to me that such a move would not be very attractive for the theory.

To conclude: Wheeler's article contains interesting proposals, but it has not convinced me that there is such a thing as categorial phonology.

4.4 Blends of different theories

4.4.1. In "Aspects of a Categorial Theory of Binding", Gennaro Chierchia tries to integrate ideas of Cooper (1983) and Bach&Partee (1981) into a framework that assumes categorial trees as deep structures. Chierchia claims that his approach can explain the following contrast:

(93) a. Mary showed the men each other [=(14)]
b. *Mary showed each other the men
c. Mary showed the men to each other

The contrast is supposed to arise because, in the categorial tree, the indirect object (IO) combines first with the verb, then comes the direct object (DO) and last the subject (SU). (Note, that Vennemann (1974) had claimed this some years before.) Thus the categorial tree of (93a) is the following:

(94) (((show each other) the men) Mary)
IO DO SU

In this tree, DO C-commands IO and can therefore bind the former, where C-command and binding are understood as in GB-theory.

What worries me a bit with this approach is this: Contrary to what one would expect of a categorial grammarian, Chierchia's account is *not* based on semantic intuitions as to the notions "direct object" and "indirect object". In (93a), the direct object is the "dative-NP", *i.e.* the "goal", whereas in (93c) the direct object is the accusative, *i.e.* the "theme". Thus, there seem to be no independent semantic criteria for grammatical functions: In direct object constructions, the direct object is the NP that can bind the other object, *i.e.* the indirect object. (This sounds almost circular.) It follows that ditransitive verbs must have two different subcategorization frames, an idea that has been accepted for a long time, *e.g.* in Lexical Functional Grammar.

Let us suppose this is so. Does it follow that this approach fares better than generative proposals? I don't think so. Chierchia's theory assumes that the indirect object is the argument that combines first with the verb. This assumption gives us the correct word order for (93a), since the categorial tree (94) has that order. But (94) is the categorial structure for (93c) as well. Therefore, we need syntactic transformations which convert this tree into (93c). Of course, we can do exactly the same in a generative approach. Thus, Chierchia's argument might falsify a particular GB-analysis, but it doesn't prove that categorial grammar can solve this puzzle better than other theories.

Nevertheless, I should add that to my mind Chierchia's framework is on the right track. He does not claim that we can do the whole of syntax by categorial principles alone. Categorial trees belong to a level of syntactic representation, a level where the binding principles are checked. Thus, this theory is a version of a multiple level approach, and the objections that have been brought forward against one-level categorial theories fail to apply. It remains to be shown, of course, that one particular level of the grammar has indeed a categorial structure.

4.4.2. Carl J. Pollard's article "Categorial Grammar and Phrase Structure Grammar: An Excursion on the Syntax-Semantics Frontier" is an informative article if one wants to compare different notations which are on the market. As for categorial grammar itself, it doesn't contain any new claims relevant for the development of the theory, as far as I can see.

4. Conclusion

Before I try to get to a general conclusion, let me first summarize the essential findings of the discussion.

First: There is the problem of overgeneration known as *Moortgat's Nuisance*. It seems to me that there is no way to formulate restrictions by means of categorial methods proper. The restrictions need to come from outside, so to speak. The sources of overgeneration are the principles of gap inheritance, in particular, the principles of type lifting (**M**, **G** and others). The only way to control gap inheritance, as far as I can see, is to simulate the different restrictions on movement that have been investigated in generative grammar, say GB-theory. That would make a branch of that paradigm.

Second: The assumption that any kind of discontinuity and long distance dependency can be simulated categorially is problematic in view of the fact that certain instances of "movement" do not have a natural semantic equivalent. A clear case were certain particle movements (cf. the discussion of Huck's article). Since the particles do not denote anything, we cannot represent them by non surfacing variables. But such a representation is needed if the categorial representation is to make sense. It seems to me that these facts show the untenability of what Bach has called the semantic import of syntactic categories. The moral to be drawn is that syntax is autonomous from semantics in at least some respects.

Third: The analysis of non-constituent conjunction by Dowty has not shown the need of categorial principles. I think the cases considered by Dowty are better accounted for within a general theory of ellipsis. On the other hand, if Dowty's account is correct, then it can also be reformulated in terms of movement.

Fourth: Keenan & Timberlake and many others have argued that categorial methods are the appropriate tool for analysing operators which change grammatical functions. One is inclined to believe that the claim is correct as far as the semantic side of these operations is concerned. As I have said this is the trivial aspect of the problem.As regards the non-trivial problems, however, categorial grammar provides us with no interesting restrictions: Any kind of imaginable grammatical-function-changing process is expressible, but only certain processes are observed. Thus, something more seems to be involved, and categorial grammar does not tell us what.

Fifth: Strong surface compositionality as advocated by Steedman and others - no gaps, no variable binding in natural language - is an illusion, as I have argued. Clearly, we need bound variables. The most perspicuous way is to have a transparent representation where everything appears what we need. In other words, we should not be afraid to admit empty categories in syntax (and, perhaps, in morphology, too).

Sixth: Concerning the bindings facts of natural language, categorial grammar cannot solve them better than any other theory. I have made this point in connection with Chierchia's article.

Seventh: A natural treatment of reanalysis was said to be another nice feature of categorial grammar. This may still be so. But the evidence that favores categorial grammar with respect to other theories seems scarce to me, as I have argued in my comments on Moortgat's article.

Eighth: The solution of bracketing paradoxes is assumed to be a merit of categorial grammar. But we have to add a qualification. Something along these lines might be in order for morphology, but the application in syntax is much more restricted than Moortgat's article suggests. Baker's

(1988) "Uniformity of theta assignment hypothesis" may be regarded as the claim that there is no functional composition in the syntax at all. On the other hand, Di Scullo & Williams(1987) allow functional composition in syntax in certain constructions, *e.g.* small clauses. Even if they were right on this, it still would not follow that we need syntactic categories of the categorial sort. functional composition is a semantic principle that can be formulated without commitment as to the syntactic representation of the functions involved. Finally, the treatment of the German and Swiss verbal complex surely involves syntactic principles proper. Thus, it is doubtful whether there is any such a thing as categorial syntax or morphology.

Ninth: I have not discovered convincing evidence that we need categorial principles in phonology.

My overall conclusion is that categorial grammar, as it stands, is too simple to be a serious candidate for a comprehensive linguistic theory. If there is any truth in the findings of the so called GB-theory (say, Chomsky's (1986) *Barriers*), then things are much more complicated. One should just try to reformulate the central distinctions of that theory (*e.g.*, theta marking, L-marking, blocking category, barrier, government, antecedent government, proper government, degrees of subjacency, the notion of subject, and so on) in terms of categorial grammar and see what kind of problems arise.

Of the concepts just mentioned, only functional properties like valency and modification are encoded in the trees, *i.e.* properties to be covered by the theory of theta marking. The other relevant distinctions have to be added to the trees from outside in form of feature conventions, metalinguistic rules, or whatever. These additions are as important as the categorial principles, however, because only after their introduction are we in a position to formulate interesting constraints for natural languages. categorial grammar has nothing to say about them. To call the so enriched theory *categorial grammar*, is largely a matter of taste. A better name would be *grammar*.

The difficulties of categorial grammar stem to my mind from an overemphasis of the semantic side of syntax. Clearly, there is a close parallelism between syntax and semantics. But it is not quite as strict as genuine categorial grammarians tend to believe. It is a good methodological principle to go as far as possible. But I think I have shown that the the parallelism has been stretched as far as it can go. Clearly, we have to weaken the relation somewhat. The problem is, however, that we cannot weaken the relation without giving up the basic tenet of the theory, *viz*. Bach´s principle of semantic commitment of syntactic categories. If we make that step, we are not categorial grammarians anymore, but just grammarians.

Some of the insights of categorial grammar will remain though. For instance, I believe that theta theory should be reformulated in the spirit of categorial grammar. Theta marking in syntax would just be functional application then (*vide* von Stechow 1989). Similar remarks hold for the representation of quantifier scope. For that purpose, it is very useful to have a categorial representation – a level "transparent logical form" – that uses a language with λ-abstraction in style of Montagues *Intensional Logic*, or Lewis' (1972) and Cresswell's (1973) λ-categorial languages. The principles governing argument inheritance in morphology should be reformulated on that level of representation, too, using the methods of extended categorial grammar with the provisos made in the discussion. In other words, the principles of categorial grammar boil down to no more than principles of semantic compositionality.

References

Aronoff, M. (1976). *Word Formation in Generative Grammar.* Cambridge/Mass.: MIT Press.

Bach, E. & B.H. Partee (1980). Anaphora and Semantic Structure. In: K.J. Kreiman & A.E. Ojeda (eds.), *Papers from the Parasession on Pronouns and Anaphora.* Chicago Linguistic Society.

Baker, M. Ch. (1988). *Incorporation. A Theory of Grammatical Function Changing.* Chicago und London : The University of Chicago Press.

Baker,M.Ch. (1989). Morphological and syntactic objects: a review of Di Sciullo and Williams' *On the definition of word. Yearbook of Morphology,* 259-283.

Bar-Hillel, Y., C. Gaifman & E. Shamir (1960). On Categorial and Phrase Structure Grammars, *Bull. Res. Concil Israel* F9, 1-16.

Chomsky, N. (1981). *Lectures on Government and Binding.* Dordrecht : Foris.

Chomsky, N. (1982). *Some Concepts and Consequences of the Theory of Government and Binding.* Cambridge/Mass. : MIT Press.

Chomsky, N. (1986). *Barriers.* Cambridge/Mass. : MIT Press.

Cooper, R. (1983). *Quantification and Syntactic Theory.* Dordrecht : Reidel.

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

Di Sciullo, A. M. und Williams, E. (1987). *On the Definition of Word*. Cambridge/Mass. und London : MIT Press.

Fabb, N. (1984). *Syntactic Affixation*. MIT Dissertation.

Gazdar, G., E.Klein, G. Pullum & I. Sag (1985). *Generalized Phrase Structure Grammars*. Oxford : Blackwell.

Haegeman, L. & H. van Riemsdijk (1986). Verb Projection Raising, Scope and the Typology of Verb Movement Rules. *Linguistic Inquiry* , 17.3, 417-466.

Hayes, B. (1984). The Phonetics and Phonology of Russian Voicing Assimilation. In: M. Aronoff & R.T. Oehrle (eds.), *Language and Sound Structure*. Cambridge, Mass. : MIT Press.

Höhle, T. (1982). Über Komposition und Derivation: Zur Konstituentenstruktur von Wortbildungsprodukten im Deutschen. *Zeitschrift für Sprachwissenschaft* 1, 76-112.

Kayne, R. (1984). *Connectedness and Binary Branching*. Dordrecht : Foris.

Kimenyi, A. (1980). *A Relational Grammar of Kinyarwanda*. Berkeley: University of California Press.

Lewis, D. (1972). General Semantics. *Synthese*, Vol. 22, 18 - 67.

Montague, R. (1974). The Proper Treatment of Quantification in Ordinary English. In: R.H. Thomason (ed.), *Formal Philosophy. Selected Papers of Richard Montague*. New Haven and London : Yale University Press.

Pesetsky, D. (1985). Morphology and Logical Form. *Linguistic Inquiry* 16, 193-246.

Quine, W.V.O.: Variables explained away. In: *Proceedings of the American Philosophical Society* 104, 343-347.

Rooth, M. & Partee, B. (1982). Conjunction, Type Ambiguity, and Wide Scope of *Or*. In: D. Flickinger et al. (Hrsg.), *Proceedings of the 1982 West Coast Conference on Formal Linguistics*. Stanford University : Department of Linguistics.

Schönfinkel, M. (1924). Über die Bausteine der mathematischen Logik. *Mathematische Annalen* 92, 305 - 316.

Stechow, A. von (1989). Syntax und Semantik. *Arbeitspapiere der Fachgruppe Sprachwissenschaft* 1, Universität Konstanz.

Szabolcsi, A.(1987). Bound variables in syntax (are there any?). To appear in: R. Bartsch et alii (eds.), *Proceedings of fht 6th Amsterdam Colloquium*. Dordrecht : Foris.

Venneman, Th. (1974). Topics, Subjects and Word Order: Form SXV to SVX via TVX. In: J. Anderson &Ch. Jones (eds.), *Historical Linguistics*, Vol. II, Amsterdam 1974.