Volume 14 *Proceedings of NELS 14* Article 19

1984

A Fregean Restriction on Metarules

Michael Moortgat
INL

A FREGEAN RESTRICTION ON METARULES

Michael Moortgat

INL

0. Abstract*

In current versions of GPSG the use of metarules is subject to severe limitations. Gazdar, Klein, Pullum & Sag (1982) and Flickinger (1983) propose to restrict the application of metarules to rules introducing lexical heads. Flickinger calls this restriction the Lexical Head Constraint. Take the central metarule for the introduction of gaps:

(1) Trace introduction metarule (cf. Flickinger 1983:93)
<A -> H,B,W> ==> <A/B -> H,t,W> (H=head)

Given the LHC, A in this rule has to stand for a category introducing a lexical head, that is, for a category of bar level 1 in common X-bar parlance. The LHC allows the gaps in (a) below, but disallows those in (b), because they do not correspond to sisters of a lexical head.

(2)(a) Who did John [see **t**]?
In which box did John [put his pencil **t**]?
(b) *Whose did John borrow [**t** book]?
*Who did John regret that [**t** arrived late]?

Gazdar, Klein, Pullum & Sag (1982) state that metarules under this conception "express generalizations about possibilities of subcategorization". Notice that these generalizations come in two types: they can be of a local or a non-local nature. The LHC is motivated on the basis

of allowable long-distance dependencies. As it stands, the constraint is too weak to ban metarule accounts for local, lexical dependencies, such as passive (Gazdar (1982)) or Dutch verb raising (Hoeksema (1981)). It seems, then, that metarules under the current restrictions perform the same heterogeneous functions as transformations: they cannot discriminate between long-distance depedencies and lexical dependencies.

In order to exclude local dependencies from the domain of metarules, I'll develop a theory of the lexicon within GPSG. The word structure grammar (WSG) is organized in accordance with the requirements of Montague's UG: the syntax of words is associated with a model-theoretic interpretation. The relation between word syntax and semantics is governed by the principle of compositionality. The meaning of a morphologically complex expression is given as a function of the meaning of its constituent parts and the word structure rule that has put them together. Frege's principle commits me to what Bach has called the rule-to-rule approach: every syntactic rule of the WSG is associated with a semantic rule. However, following Klein & Sag (1982), rule-specific semantic stipulations will be avoided. The semantics of word structure rules must be predictable on the basis of the logical types of the constituents, and a limited set of possible modes of combination in semantic rules. Call this set **G.**

Consider now the range of phenomena known as complement inheritance. Take a verb such as **rely**. It is subcategorized for a PP[**on**], i.e. a prepositional complement with the lexical preposition **on.** The nominalization of **rely** is subcategorized for the same PP[**on**] : **reliance on NP**. Apparently, the subcategorizational property of **rely** is inherited by the derivation **reliance**. Expressions such as **reliance on NP** present a mismatch between syntactic structure and semantic scope relations - what Pesetsky (1983) has called a bracketing paradox. Semantically, the nominalization affix has scope over the combination **rely on NP**, which suggests a bracketing nom(rely(on(NP))); but syntactically, the affix forms a unit with **rely**, and the complex form **reliance** is then combined with the PP[**on**].

On the basis of complement inheritance facts, I'll argue that **G** cannot be restricted to functional application if the interpretation of word structure has to respect the compositionality requirement. The set **G** will be extended with a generalized version of functional composition. Using functional composition, a morphologically complex expression like **reliance** can get a strictly Fregean interpretation which yet accounts for the phrasal scope of the affix. This analysis will be contrasted with a GPSG metarule approach. A metarule would state the dependencies between **rely on NP** and **reliance on NP** on the phrasal X^1 level, not on the lexical X^0 level. I'll show that metarules for complement inheritance violate the compositionality principle: they introduce complex lexical nodes for which there is no corresponding constituent in the semantics, hence no semantic object in the model.

This line of argumentation will then be applied to standard metarule accounts for passive and Dutch verb-raising. They will be shown to suffer from the same defect: they introduce morphologically complex expressions (the passive participle and the verb raising cluster) the meaning of which is not computed in a strictly local manner from the meaning of their parts. But given functional composition, a Fregean alternative is again available (e.g. Bach (1983a), Steedman (to appear) for verb-raising). On the basis of these facts, the hypothesis will be put forward that metarule accounts for local dependencies can be eliminated entirely.

What I have to say owes a great deal to insights developed within generalized categorial grammar where the possibilities of functional composition have been explored for some time. I try to incorporate the semantic insights concerning functional composition into an X-bar word structure grammar without giving up the syntactic constraints of X-bar theory. The main difference between categorial grammar and phrase structure grammar concerns the nature of the type assignment function. In categorial grammar, the relation between syntactic categories and semantic types is a homomorphism. In phrase structure grammar, the type assignment function is less transparent: the semantic type of an expression cannot be directly read off from the syntactic category (cf. Gazdar, Klein, Pullum & Sag (to appear)). This weaker position is motivated by the claim that there are syntactic generalizations which must be expressed independently of semantic types and vice versa. Another way of stating it would be that the type theory used in the interpretation, and its natural combinatorics (functional application and composition) is in itself too rich, but that the word syntax, constrained by X-bar theory, prevents one from exploiting this excessive power.

1. Word syntax and semantics

Janssen (1983) gives a detailed account of the general algebraic framework which forms the background for the approach towards morphology adopted here. Within this framework, syntax and semantics are many-sorted algebras; the relation between these algebras is a homomorphism. Let A and B be the syntactic and the semantic algebra of our word grammar.

$$A = \langle (A_s)_{s \in S}, \mathbf{F} \rangle \qquad B = \langle (B_t)_{t \in T}, \mathbf{G} \rangle$$

S and T are two non-empty sets, the categories of the word syntax and its semantics. For each category there is a set of elements of that category, A_s and B_t. Finally, each algebra contains a collection of operations – the syntactic operations F and the semantic operations G. Let us assume with Janssen that the operations are mappings defined for specific categories as indicated below, i.e. that rules and operators coincide.

$$F_i: A_{s_1} \times \dots \times A_{s_n} \rightarrow A_{s_{n+1}}$$

$$G_i: B_{t_1} \times \dots \times B_{t_n} \rightarrow B_{t_{n+1}}$$

What does it mean in this framework that the interpretation obeys the principle of compositionality? Frege's principle assumes the form of a specific kind of mapping between the syntactic and the semantic algebra, viz. a homomorphism. The syntactic algebra and the semantic algebra have a similar structure, as indicated above; the interpretation function, as a homomorphism, preserves this structure. More specifically, the interpretation function h:A->B is a homomorphism if (i) h respects the categories, i.e. $h(A_s) \subset B_t$ where t is the semantic category corresponding to the syntactic category s; and (ii) h respects the operators, i.e. $h(F(a_1,\ldots,a_n) = G(h(a_1),\ldots,h(a_n))$ where G is the semantic operation of B corresponding to the syntactic operation F in A (cf. Janssen (1983:22)).

The algebraic framework, as a general metatheory, imposes no constraints on possible categories or possible operations of the syntactic and semantic algebras. However, a linguistically interesting instantiation of the framework will formulate such constraints. It is a direct consequence of the compositionality requirement that constraints on one algebra will have repercussions for the other. In this section, I will assume with Selkirk (1982) that the categories and operations of the word syntax are constrained by X-bar theory. In section 2., I'll investigate the consequences of this version of word syntax for the semantic algebra.

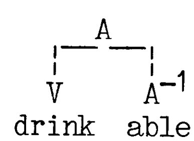
Selkirk (1982) defends the position that the basic concepts and principles of X-bar theory can be extended below the word level, and that notions from syntax such as head, maximal projection from a head, complements of a head vs. modifiers, are applicable to the syntax of words once the relevant lexical parameters have been set. The WSG, then, will contain syntactic objects of the type illustrated in (3). A rule of this form allows a word tree like (4).[1]

(3) $[\ V[2]^0\ \ A^{-1}\ ]_{A^0}$ (node admissibility format)

$A^0 \rightarrow V[2]^0\ \ A^{-1}$ (rewrite format)

where A^{-1}={able,...} and V[2]={drink,eat,...} (transitive V's)

(4)



The categories of the WSG consist of a feature specification (abbreviated as N,A,V, etc.) and a bar integer, indicating the syntactic level of the category in the bar hierarchy. As in PSG, we have to determine the maximal and minimal bar level values. Selkirk's distinction between stems and roots is irrelevant to the problem discussed here, and will be ignored. Bar level zero is taken to be the maximal projection of the WSG, i.e. the recursive level and the interface between lexicon and syntax. The minimal projection is set at -1, a departure from Selkirk. The bar level feature assumes a negative value for bound morphemes: affixes have lexical entries with categorial features like free morphemes; but since they are of level -1 they cannot enter the PS syntax as they stand; they must combine with a complement to make an X^0. The rule above gives expression to the traditional idea that affixes like **-able** are category-determining elements. In X-bar terminology, this means that the affix in (4) is the head of the complex expression: it has the same major categorial features, and it is one step down in the bar hierarchy. The notion of head can be defined for the WSG as it is for the PSG. Roughly, then, (cf. Gazdar (1982))

(5) In a rule $[\ldots Y \ldots]_X$, Y is the head of X iff
(i) $X = \langle[\alpha N],[\beta V],j\rangle$
(ii) $Y = \langle[\alpha N],[\beta V],i\rangle$
(iii) $i \leq j$

Notice that the final clause allows for two kinds of head. The bar level of the head may be smaller than that of the derivation as a whole; in that case the affix is the head (cf. **-able**). Or the bar level of the head may equal that of the derivation; in that case the base itself, not the affix, functions as the head of an expansion. The second option is needed for rules of category-neutral affixation, such as (6). In Moortgat (1981) I have argued that category-neutral affixes of this type should be treated as modifiers, syntactically and semantically.

(6) $[\ A^{-1}\ N^{0}\ ]_{N^{0}}$ where $A^{-1}=\{$ex-,arch,...$\}$

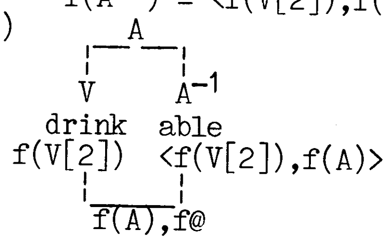
ex-president, arch-enemy,...

Given these assumptions, one can view the interaction between phrase structure grammar and WSG in the following way. The phrase structure grammar contains a set of rules introducing lexical items - the X^1 rules. An example would be (7), the A^1 rule for intransitive adjectives. A rule like (7) is associated with a set of basic expressions of the appropriate sort. Word structure rule (3) enlarges this set with a class of derived expressions of that type. The hidden assumption behind this view is that the rules of the WSG are structure-preserving, i.e. that they do not generate types of lexical items which do not fit some X^1 frame of the PSG.

(7) $[\ A^{0}\ ]_{A^{1}}$ A=$\{$blind,deaf,dead,...$\}$ (i.e. basic A's)

Let me indicate now how the WSG can be associated with an interpretation. The relation between word syntax and semantics is governed by one central principle: compositionality. The principle requires that the meaning of morphologically complex expressions be given as a function of the meaning of their parts and the way they are put together. As usual, the interpretation will make use of a translation into IL. This further requires a function f which maps the syntactic categories of the WSG onto the syntactic categories (the types) of IL; and a function t which translates the elements of the terminal vocabulary into constants of IL. (8)-(10) illustrate these requirements using the example **drinkable.**[2]

(8) $f(V[2]) = \langle f(NP),\langle f(NP),f(S)\rangle\rangle$ transitive verbs
$f(A) = \langle f(NP),f(S)\rangle$ intransitive adjectives
$f(A^{-1}) = \langle f(V[2]),f(A)\rangle = \langle\langle f(NP),\langle f(NP),f(S)\rangle\rangle,\langle f(NP),f(S)\rangle\rangle$

(9)


1. $[\text{drink}]_V$: drink'
2. $[\text{able}]_{A^{-1}}$: able'
3. $[[\text{drink}][\text{able}]_A$: able'(drink')

(10) able' $= \lambda R \lambda \mathbf{P} \Diamond [R(\mathbf{P})(\hat{P} \bigvee y P\{y\})]$ (R of type f(V[2]), **P** of type f(NP))

First consider the type assignment function. We know from the lexical type assignments in the PSG what logical types correspond to X^0 expressions. In the WSG, f has to be expanded so as to assign a type to the affixes as well. The function will be stated in such a way that the type assignments from syntax remain valid. In the example **drinkable**, we know already from the PSG that the lexical type assignments to transitive verbs and intransitive adjectives are as in (8). So let's fix the lexical type assignment to the adjective-creating affix, $f(A^{-1})$ as $\langle f(V[2]),f(A)\rangle$, i.e. a function from transitive verb denotations to intransitive adjective denotations.

Consider now the semantic translation rule. The rule-to-rule hypothesis demands that we specify some mode of combining the translations of the constituents so as to obtain the translation associated with the mother node. Usually, the mode of combination is stipulated over and over again for each rule. Instead, I will follow Klein & Sag's (1982) interpretation of the rule-to-rule hypothesis. They have suggested to avoid rule-specific semantic stipulations. They claim that the particular mode of combination need not be stipulated, but that it is derivable from the types of the constituents involved, and a limited set of possible modes of combination in semantic rules. We know the types of the affix-head and of its complement. We have to combine the translations of these expressions in such a way that the result is an IL expression of the type of the derived category. Clearly, the semantic mode of combination that will guarantee this is functional application: the translation of the affix-head is applied as a functor to the translation of the base. Below the line in (9) a semantic analysis tree is given in terms of the types of the constituent expressions and the mode of combination used. f@ stands for functional application.

One might think that the appearance of able'(drink') as the translation of **drinkable** is a little unrevealing. From Montague-oriented work on lexical semantics one would expect some complex lambda-expression which directly accounts for the fact that the subject of an **able**-adjective is understood as the object of the verbal base, and that **drinkable** means something like "which can be drunk" (cf. Dowty (1979)). I think this goes beyond what a compositional semantics for morphologically complex structures can be expected to do. Frege's Principle requires that the meaning of the whole is given as a function of the meaning of the parts and the way they are combined. The parts in the example are the morphemes **drink** and **able**. These were each associated with a certain logical type, and hence with a set of possible denotations; and the semantic rule indicated how they were to be combined, viz. by f@. Compositional semantics can stop here. In particular, it is not necessary to state what specific function from transitive-verb type things to intransitive-adjective type things is associated with **able**. If one wants to do that, one is giving the lexical semantics for that morpheme. One might wish to state the equivalence in (10), which, modulo some type raising, makes able'(drink') equivalent to Dowty's semantics for **able**-adjectives. In the rest of this paper I'll ignore the lexical semantics of bound morphemes, and I'll concentrate on the recursive aspects of the interpretation of morphologically complex expressions as far as these are determined by the word syntax.

A more important difference from Dowty (1979) concerns the semantic status of derivational morphemes. Dowty introduces bound morphemes syncategorematically: "operations, not just morphemes, are assigned meaning [...]; there is no need to distinguish a bound morpheme itself [...] from the operation of attaching that morpheme (Dowty (1979:304))". Under the approach defended here, affixes have the full status of basic expressions, syntactically and semantically. They belong to a syntactic category, which gets assigned a semantic type by the type assignment function f; as basic expressions, they are elements of the domain of the translation function t which links them with constants of IL.

It is clear that the syncategorematic treatment and the basic expression approach are descriptively equivalent. But the basic expression approach has conceptual advantages: syncategorematic introduction of affixes prevents one to express obvious generalizations. Suppose a language has different affixation processes deriving, say, transitive verbs from nouns, and that the derivations are associated with different meanings. An example would be English denominal verbs: compare **bewall**, "to provide NP with N", **disarm**, "to rid NP of N", **enslave**, "to make NP into N". Because the meaning contribution of the affixes is different in each case, the syncategorematic approach has to posit distinct rules for the ornative, privative and causative cases. If affixes are treated as basic expressions, one rule suffices. The syntax-dependent aspect of the interpretation is common to the three affixation processes: the affixes are all of type <f(N),f(V[2])>; they are associated with different IL constants, however. So we can locate the difference between the ornative, privative and causative derivations exactly where it originates - in the different meaning contribution of the affixes.

2. Relatedness paradoxes and functional composition

In the following paragraph, I'll concentrate on word structure rules which cannot be succesfully interpreted with the use of simple functional application. All these examples have the following characteristics: the word syntax puts together two expressions A and B to form a derived expression C; but there is a type mismatch between the sisters A and B - it's impossible to functionally apply one to the other so that the resultant expression is of the type assigned to C.

As a first example, consider some well-known complement inheritance facts. The WSG has to contain a rule like (11) to derive words like **blindness**, abstract nouns derived from adjectives.

(11) $[\ A\ \ N^{-1}\ ]_N$ $\qquad N^{-1} = \{\text{ness}, \ldots\}$

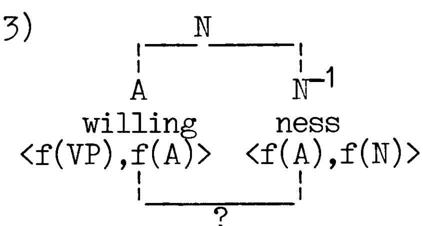
blindness	darkness
bitterness	deepness
brightness	drunkenness

Assume that the type of simple intransitive adjectives and common nouns is <f(NP),f(S)>. The suffix will consequently be assigned the type <f(A),f(N)>, a function from A-type things to N-type things. The lexical semantics of **-ness** is of no relevance for the problem at issue, so t(ness) is just ness'. **Blindness** is translated as ness'(blind'); the semantic mode of combination is functional application.

Consider now adjectives like **willing** in (12). The lexical types assigned to these adjectives are not just <f(NP),f(S)>; they are subcategorized for a PP or a VP[INF]. Semantically, they are functions from the type of things their complements denote to simple adjective denotations. But abstract nouns can be derived from these adjectives just as from the simple adjective **blind**. In such cases, the subcategorization of the base is inherited by the derived expression. We cannot invoke functional application to combine these expressions semantically: the argument of **ness** is not of the appropriate type. Yet, clearly, we want to formulate one rule of **ness**-affixation, and hence one type assignment to **ness**, which generalizes over these cases.

(12)	willing [to VP]	willingness [to VP]
	related [to NP]	relatedness [to NP]
	fond [of NP]	fondness [of NP]
	apt [to VP]	aptness [to VP]
	contented [with NP]	contentedness [with NP]
	kind [to NP]	kindness [to NP]
	distinct [from NP]	distinctness [from NP]

(13)



The examples in (12) fall under an extension of Williams' (1982) notion of relatedness paradoxes. **Willingness to go** is morphologically related to **willing to go**, although the phrase **willing to go** forms no proper subconstituent of it: this phrase is broken up by the affix. There have been various attempts recently to resolve the relatedness paradoxes posed by complement inheritance.

Along the lines of Lieber's (1980) feature grammar, there is a purely syntactic approach to inheritance. Toman (1981), and Lieber (1983) have proposed that when affixes do not carry subcategorizational features themselves, these features can percolate upward from the non-head to the level of the derivation. It is clear that feature percolation does not solve the relatedness paradox. The percolation device incorrectly suggests that the VP complement in **willingness to go** is an argument of **willingness**, whereas the paradox resides exactly in the mismatch between syntactic structure and semantic scope: semantically, the affix **ness** has scope over the combination **willing to go**, but syntactically, the affix forms a unit with **willing**.

A different approach is defended by Williams (1982) and Pesetsky (1983). These authors correctly identify the semantic scope problem, but they resolve it by abandoning the surface compositional approach. They state the relation between **willing to go** and **willingness to go** by invoking a second level of syntactic representation; Pesetsky obtains this by raising the affix in LF, Williams by the comparable device of head-stripping. The rejection of compositionality presents a weakening of the relation between syntax and semantics which is unavailable in the framework defended here.

Let us investigate then whether complement inheritance can be captured in the GPSG framework in terms of a metarule relating A^1 to N^1. Such a rule could be formulated as in (14):

$$(14)\quad < [\ A \quad W\ ]_{A^1}\ ,\ P > \;==>\; < [\ \ [[A]ness]_N \quad W\ ]_{N^1}\ ,\ ness'(P)\ >$$

The rule says that for an A^1 expansion with a variable W ranging over subcategorized complements (possibly null), and a translation P, there is also an N^1 expansion, with **ness** right-adjoined to the A head and the same subcategorization W. The affix **ness** has scope over the (possibly complex) A^1 translation P. A metarule like (14) is not ruled out by the Lexical Head Constraint: it is stated on PS rules introducing lexical heads. But does it obey the compositionality requirement?

When the requirement is extended to the lexicon, the metarule suffers from the same defect as the Williams-Pesetsky approach: it violates Frege's principle. The WSG builds a complex expression **willingness** dominated by the lexical node N. The compositionality principle requires that we assign this complex syntactic expression a meaning as a function of the meaning of its parts, **willing** and **ness**. In other words, we cannot postpone the interpretation of the complex derived N node until the meaning of the adjectival base has been combined with its complement VP. But this is exactly what the metarule attempts to do. This brute force method is unavailable given a tight interpretation of the Fregean principle: every category of the word syntax must receive a semantic value computed locally from the semantic values of its components.

In the remainder of this paper I'll try to solve the problem of complement inheritance in a compositional way, without making an appeal to phrasal metarules. At this point, it may be useful to recall the discussion of the balance between the syntactic and the semantic algebra from the previous section. We chose to constrain the operations of the word syntax to a CF word structure grammar of the X-bar type. Given this constraint on the permissible syntactic operations, complement inheritance facts seem to defy a compositional interpretation. Faced with this problem, one can either loosen the constraints on the permissible syntactic operations of the word grammar, or one can explore semantic possibilities, and give up the implicit assumption that functional application is the only semantic operation corresponding to syntactic concatenation. The second approach will be taken here.[5] If there is a meaning correspondence between two expressions, this correspondence must be established in the semantics; there is no compelling reason to express the correspondence in the syntax as well (cf. Janssen (1983)). Let us investigate the consequences of enlarging the set of possible modes of combination in semantic rules with a generalized version of functional composition.

North East Linguistics Society, Vol. 14 [1984], Art. 19

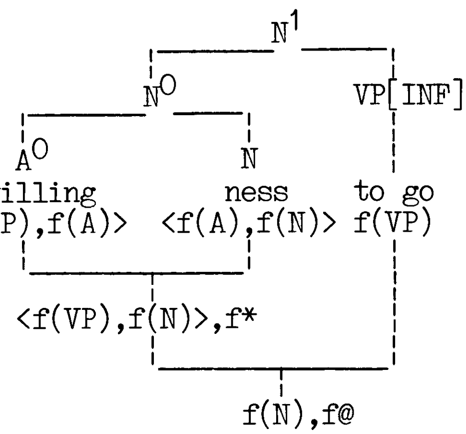
(15) Functional composition
If X' is a meaningful expression of type $\langle b,c\rangle$ and Y' is a meaningful expression of type $\langle a_1,\langle\ldots\langle a_n,b\rangle\ldots\rangle$, then the composite of Y' and X', f*(X')(Y') is a meaningful expression of type $\langle a_1,\langle\ldots\langle a_n,c\rangle\ldots\rangle$

Definition f* :

$$\lambda v_{\langle b,c\rangle}\lambda v_{\langle a_1,\langle\ldots\langle a_n,b\rangle\ldots\rangle}\lambda v_{a_1}\ldots\lambda v_{a_n}[v_{\langle b,c\rangle}(v_{\langle a_1,\langle\ldots\langle a_n,b\rangle\ldots\rangle}(v_{a_1})\ldots(v_{a_n}))]$$

The version of composition given here is a generalization of Klein & Sag's (1982) combinator f_R, which is only defined for the types involved in raising constructions. This type restriction is given up in (11), where **a,b,** and **c** stand for arbitrary types. It is a generalization of the standard version also in allowing for the inheritance of more than one argument. In this respect it is equivalent to the liberal versions of composition used within the categorial framework, e.g. Bach (1983a) or Steedman (to appear).[4] Figure (16) illustrates how f* works in the case of **willingness.**

(16)



1. $[\text{willing}]_A$: willing'
2. $[\text{ness}]_{N^{-1}}$: ness'
3. $[\text{to go}]_{VP}$: go'
4. $[[\text{willing}][\text{ness}]]_N$: f*(ness')(willing')
   $\lambda P[\text{ness}'(\text{willing}'(P))]$ (Def.f*; P of type f(VP))
5. $[[\text{willingness}][\text{to go}]]_{N^1}$: $\lambda P[\text{ness}'(\text{willing}'(P))](\text{go}')$
   ness'(willing'(go'))

The lexical type of **willing** is a function from VP type things to intransitive A type things. The lexical type of **-ness** is a function from intransitive A type things to N type things. Functional application cannot combine these two expressions. But f* assigns **willingness** to the set of meaningful expressions which are functions from VP type things to N type things. A derived expression like **willingness** combines with the VP complement by f@. And the translation assigned to **willingness** by f* guarantees that the VP complement ends up as an argument of the adjectival base, not of the derived noun. So Frege's Principle is strictly obeyed, without giving **-ness** the wrong scope.

Inheritance of complements under derivation is a widespread phenomenon. I won't enter here into the many problems involved in determining the exact range of inheritance. I have discussed some of these problems elsewhere (Moortgat (1981)), where I suggest that there is a semantic dichotomy in the WSG based on the distinction between transparent and nontransparent affixes, or in terms of the present framework, between affixes allowing f* and affixes restricted to f@. It has been observed, for example, that **-able** doesn't show inheritance effects (**breakable**, but : ***persuadable to go**, ***considerable a fool**, ***puttable on the desk**; see Aronoff (1976:49)), which means that it can only combine with a basic transitive expression of type f(V[2]) via f@. Further research will have to demonstrate whether membership of the class of transparent affixes can be derived from other properties.[5]

I motivated the introduction of functional composition on the basis of complement inheritance facts in derivational morphology. In the remainder of this section, I would like to investigate some related areas in the lexicon where standard metarules violate Frege's principle but where functional composition again offers a Fregean alternative. Let's first discuss the controversy concerning phrasal vs. lexical accounts of passive. Bach (1980) and Keenan (1980) have argued convincingly that passive cannot be adequately treated as a purely lexical rule. The scope of passive is the transitive verb phrase; the head of a transitive verb phrase may be a basic transitive verb (**kill**) or a complex one (**persuade to go**). The initial GPSG answer to these arguments has been to derive passive VP's from active VP's by a metarule, stated on the phrasal level (Gazdar (1982:161)). Nevertheless, passive morphology is a word-based process and shows up on the verbal head of the transitive verb phrase. With functional composition the passive participle can be given an interpretation which strictly obeys Frege's principle, and yet accounts for the phrasal scope of the passive rule. The passive morpheme will be introduced by (17). It is assigned the required "phrasal" type <f(V[2]),<f(NP),f(S)>>. But syntactically, the passive morpheme combines with a verb stem. If the type of this verb stem is more complex than f(V[2]), functional composition will be used to assign the passive participle an interpretation. The derivation of **persuaded** in (18) illustrates this. Notice that (18) accounts only for the short passive. I won't discuss the interpretation of the **by**-phrase here, which I regard (following Wasow (1980) and Keenan (1980)) as an optional modifier extraneous to passivization proper.

(17) $[\ V[TRANS]\ \ V[PASS]^{-1}\ ]_{V[PASS]^0}\qquad V[PASS]^{-1}=\{ed,\ldots\}$

$f(V[PASS]^{-1}) = \langle f(V[2]),\langle f(NP),f(S)\rangle\rangle$
$t(ed) = ed' = \lambda R\, \lambda \mathbf{P}[R(\mathbf{P})(\hat{P}\bigvee yP\{y\})]$

(18)

```
                 ___V[PASS]___
                |             |
                V          V[PASS]^-1
            persuade          ed
       <f(VP),f(V[2])>   <f(V[2]),<f(NP),f(S)>>
                |             |
                |_____________|
                       |
            <f(VP),<f(NP),f(S)>>,f*
```

1. $[\text{persuade}]_V$: persuade'
2. $[\text{ed}]_{V[PASS]^{-1}}$: ed'
3. $[[\text{persuade}][\text{ed}]]_{V[PASS]}$: f*(ed')(persuade')

$$f^*(\lambda R\,\lambda \mathbf{P}[R(\mathbf{P})(\hat{P}\bigvee yP\{y\})])(\text{persuade}')\quad (\text{Def.ed}')$$
$$\lambda P\,\lambda \mathbf{P}\,[\text{persuade}'(P)(\mathbf{P})(\hat{P}\bigvee yP\{y\})]\quad (\text{Def.f}^*)$$

I might stress here that in a semantics which allows functional composition, the choice between a metarule account and a strictly lexical account is not free. One can put it in terms of Keenan's (1980) Minimal Domain Principle which requires that "in defining a particular derivational function [...] we take the domain of definition of the function to be the smallest class (level) of structures on which it can be defined giving adequate results" (1980:184). If functional application is the sole mode of combination in the semantics, the domain of the passive participle is indeed too small to give an adequate account for the phrasal scope of passive. If the semantics has functional composition at its disposal, the passive participle is the smallest level of structure on which the rule can be defined compositionally; therefore, the Minimal Domain Principle requires that we state the rule at this level, and not at a higher level.

The last area of the WSG I want to discuss is verb raising in Dutch (VR). There have been various non-transformational analyses of this construction. Hoeksema (1981) has defended a GPSG metarule approach. Within categorial grammar, different accounts based on functional composition have been proposed, e.g. Bach (1983a), Steedman (to app.), Zwarts (1983). I don't intend to review the complex array of facts surrounding the VR construction. But I want to point out what the consequences are of the compositionality requirement in the WSG.

(19)	dat Jan wil vertrekken that Jan wants leave	[[wil][vertrekken]]
	dat Jan Marie wil straffen that Jan Marie wants punish	[[wil][straffen]]
	dat Jan Marie een boek wil geven that Jan Marie a book wants give	[[wil][geven]]
	dat Jan het boek op de tafel wil leggen that Jan the book on the table wants put	[[wil][leggen]]

A typical set of VR structures is displayed in (19). The cluster has a predictable number of arguments, which is a function of the number of V's in it and the subcategorization of these V's. The problem is how to calculate this function. As in the cases of passive and inheritance discussed above, the metarule approach states the relationship between the subcategorization of the VR cluster and the constituent verbs on the phrasal level. The rule for equi-verbs like **willen** (**want**) is given as (20) by Hoeksema (simplified somewhat).

(20) $\langle[\ W\ \ V\ ]_{VP}, P\rangle \Longrightarrow \langle[\ W\ \ [V[m]\ V]_V]_{VP}, V[m]'(P)\rangle$

$V[m] = \{\text{wil},\ldots\}$ (want,...)

The input to the metarule is a VP with a head subcategorized for W (possibly null); the output is a VP with the same subcategorization W, but with a complex VR cluster consisting of the original verb with the equi-verb (**wil,...**) left-adjoined to it. The input VP translates as P. The equi-verb is treated as a functor having the VP translation P as an argument. The metarule approach violates the strict interpretation of Frege's Principle we adopt here. The VR structures are headed by a complex verbal cluster $[V\ V]_V$ dominated by a node of bar level zero. In (19) the clusters are given in brackets. Now look at the semantics of the output of the VR metarule. There is no constituent in the semantics corresponding to the syntactic cluster $[V\ V]_V$: the equi-verb is given its phrasal scope over the translation of the input VP, although it does not combine with the VP syntactically, but with a proper subpart of it, its verbal head. Again, the semantic part of the metarule performs an operation which is reminiscent of Williams' head-stripping or Pesetsky's affix raising in LF.

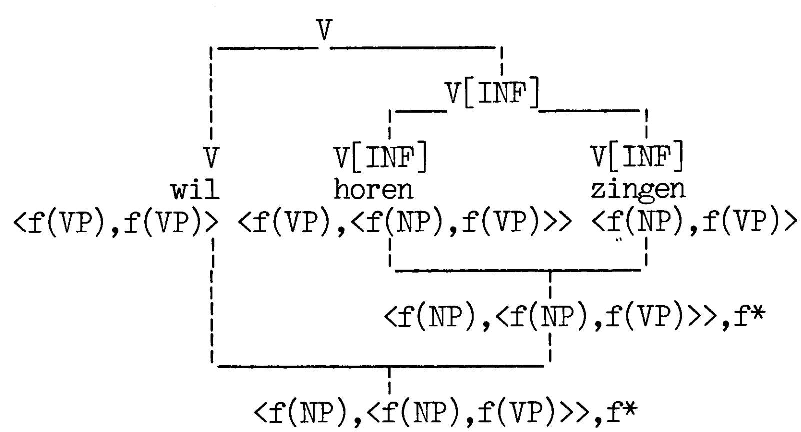
So let's investigate an alternative route, and derive the verb clusters in the lexicon, a possibility suggested in Pullum & Gazdar (1982:fn13). Given the type assignments that are needed anyway, functional composition will compute the valency of a VR cluster of arbitrary complexity. Below I present the word structure rules for the introduction of equi-verbs and verbs like **horen** (**hear**), lexical type assignments and translations for these verbs, and a sample tree. I use f(VP) to abbreviate <f(NP),f(S)>. For more details the reader is referred to the categorial works mentioned above.[6]

(21) $[\ V[m]\quad V[INF]\ ]_V$ where V[m]={wil,...}
$\quad\ \ [\ V[n]\quad V[INF]\ ]_V$ where V[n]={laat,hoor,...} (let,hear,...)

(22) f(V[m]) = <f(VP),f(VP)>
$\quad\ \ $ f(V[n]) = <f(VP),<f(NP),f(VP)>>

(23) t(wil)= $\lambda P\ \lambda \mathbf{P}\ \mathbf{P}\ \{\hat{x}[wil'(P(x^*))(x^*)]\}$
$\quad\ \ $ t(hoor)= $\lambda P\ \lambda \mathbf{P}\ \lambda \mathbf{Q}\ [hoor'(P(\mathbf{P}))(\mathbf{Q})]$

(24)



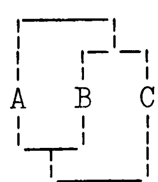
1. $[wil]_V,[horen]_V$: cf. (23)
2. $[zingen]_V$: zing'
3. $[[horen][zingen]]_V$: $f^*(\lambda P\ \lambda \mathbf{P}\ \lambda \mathbf{Q}[hoor'(P(\mathbf{P}))(\mathbf{Q})])(zing')$
   $\lambda \mathbf{R}\ \lambda \mathbf{P}\ \lambda \mathbf{Q}[hoor'(zing'(\mathbf{R})(\mathbf{P}))(\mathbf{Q})]$ (Def.f*)
4. $[[wil][horen\ zingen]]_V$:
   $f^*(\lambda P\ \lambda \mathbf{S}\ \mathbf{S}\ \{\hat{x}[wil'(P(x^*))(x^*)]\})(\lambda \mathbf{R}\ \lambda \mathbf{P}\ \lambda \mathbf{Q}[hoor'(zing'(\mathbf{R})(\mathbf{P}))(\mathbf{Q})])$
   $\lambda \mathbf{R}\ \lambda \mathbf{P}\ \lambda \mathbf{S}\ \mathbf{S}\ \{\hat{x}[wil'(hoor'(zing'(\mathbf{R})(\mathbf{P}))(x^*))(x^*)]\}$ (Def.f*)

Some explications are in order. The WS rule for V[m]'s like **wil** specifies that the verb it combines with must be a bare infinitive. This infinitive can be a basic one, or it can be a cluster itself, formed by another WS rule, e.g. the one for **horen** type verbs. The Head Feature Convention guarantees that the [INF] feature required by **wil** goes down to the head of its complement, if this is a cluster as well, as in the example (24). Notice that the control properties of the equi-verb **wil** are derived via a complex translation for this lexical item. The translation in (23) accounts for the fact that the surface subject is interpreted as the understood subject of the infinitival complement, and that it receives wide scope with respect to the complement verb (cf. Dowty (1978)). The cluster **wil horen zingen** is looking for three NP arguments (two within VP); it can consume these arguments by simple functional application.

3. Arguments vs. modifiers

The cases considered so far all had the general form of (25). Syntactic arguments suggested the upper bracketing, where B and C form a constituent which is then combined with A. But the semantic scope relations are captured more adequately by the lower bracketing: the element C has scope over the combination of A and B. The apparent mismatch between syntax and semantics was resolved by allowing functional composition as a mode of combination in semantic rules. Crucial for this account was the fact that the element A was an argument subcategorized for by B. The type assigned to B contained the information about the missing argument A, and f* passed this information on to the node dominating B and C.

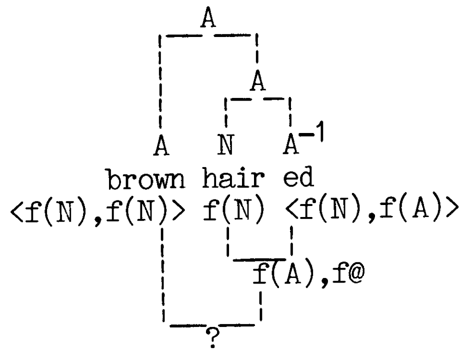
(25)



There is another type of mismatch between syntax and semantics which is structurally indistinguishable from the above, but semantically quite different. Again we have a structure (25) where the syntax forces us to group B and C, but where C has scope over the combination of A and B. The difference is, that now the element A is a modifier of B, instead of an argument. Semantically, then, A is the functor taking B as its argument.

Examples of this type of misplaced modifiers have been cited in the morphological literature as evidence against the validity of the surface compositional approach in this area of grammar. Williams (1981) and Pesetsky (1983) discuss relatedness paradoxes like **unhappier** and **atomic scientist** in which **unhappy** and **atomic science** seem to form a semantic, but not a syntactic constituent. Consider also synthetic compounds of the type **brownhaired.** The most plausible syntactic analysis is given in (26); this syntactic structure cannot be interpreted compositionally as "with (brown (hair)))", given the semantics developed so far. Notice that functional composition does not work here because the elements **un-**, **atomic** and **brown** are not arguments of **happy**, **science** and **hair**, but rather functors themselves.

(26)



For a general solution, I would like to investigate the technique of type lifting, as developed in Lambek's (1961) work on categorial grammar. In Lambek's type calculus, the function-argument relations are not fixed. Argument categories of type f(X) automatically also belong to the functional type <<f(X),f(Y)>,f(Y)>, i.e. they can also be functions from functors taking f(X) type arguments into the range of these functors. The meaning of the higher type <<f(X),f(Y)>,f(Y)> expression in these cases is predictable on the basis of the simple f(X) expression: the lifted type expression reverses the function-argument relation. Now let us return to the problem at hand, the misplaced modifiers in morphology. Suppose the grammar contains the following instance of the general type-lifting rule mentioned above.[7]

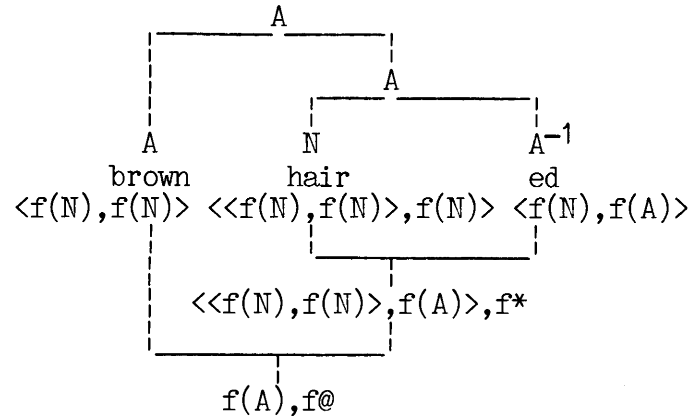
(27) If $X \in ME_{f(X)}$, then also $X \in ME_{<<f(X),f(X)>,f(X)>}$

where $t(X_{<<f(X),f(X)>,f(X)>>}) = \lambda v_{<f(X),f(X)>}[v_{<f(X),f(X)>}(X')]$

What this rule says, is that a category X of type f(X) will automatically also be of type <<f(X),f(X)>,f(X)>, i.e. a function from X-modifiers to X-type things. And the expression of the derived type is not just any function of that type: its interpretation is completely determined by the interpretation of the original expression. In fact, the value of the lifted type expression applied to its argument (the X-modifier) is the same as the value of the X-modifier applied directly to the lower type expression.

Consider now the synthetic compound **brownhaired**. On the basic type assignment to **hair** (f(N)), the derivation blocks, as illustrated above. But it works out, given the raised type assignment <<f(N),f(N)>,f(N)> and functional composition. The first step of the derivation forms the composite function of the affix-functor and the noun, interpreted here as a function on N-modifiers. The resulting function can be applied to the left member of the compound. After λ-conversion we end up with the right scope relations among the elements involved. **Brownhaired** comes out as meaning "with (brown (hair))", not as "brownly (with (hair))".

(28)



1. [hair][ed]$_A$: f*(ed')(λK[K(hair')])
λMed'(λKK(hair')(M))
λMed'(M(hair'))
2. [brown][haired]$_A$: λMed'(M(hair'))(brown')
ed'(brown'(hair'))

It might be useful to reflect here on the motivation for the introduction of the type lifting rule (27). On the basis of complement inheritance facts, functional composition was added to the set of possible modes of combination in semantic rules. Functional composition makes a very strong prediction: it predicts that the only type of expressions that can be interpreted non-locally will be arguments. As it stands, this prediction appeared to be too strong: this paragraph discusses non-locally interpreted modifiers. The strategy followed here has been to convert these modifiers into arguments by raising the type of the modified expressions. Since this obviously weakens the prediction made by functional composition, additional empirical support for the type lifting rule would be welcome.

The general possibility for function-argument reversal, as it is present in the Lambek calculus, is an exceedingly powerful device, and one would like to find a natural way to constrain it. The most obvious drawback of unlimited type lifting is that one looses an explanation for the morphosyntactic effects of the function-argument status. These effects are captured by the Control Agreement Principle (CAP), the GPSG descendant of Keenan's (1974) Functional Principle. The central claim of the CAP is that functions can show agreement with their argument, but not vice versa. Suppose now that lifted types can only be used when licenced by the CAP, i.e. that the function-argument relations they establish may not conflict with the predictions made by the CAP.

Let us investigate the lifted type assignment to the noun in the example **brownhaired** in this light. English shows no effects of the CAP in this respect, since there is no agreement between adjectives (functions of type <f(N),f(N)>) and their arguments. So we have to turn to the Dutch counterparts of this type of synthetic compound. Dutch adjectives, as noun-modifiers, agree in gender with their arguments. Compare **een stomp**[+Neuter] **voorwerp**[+Neuter] "a blunt object", and **een stomp+e**[-Neuter] **hoek**[-Neuter] "an obtuse angle". I claimed that in compounds like **stomphoekig** (**obtuse-angled**) the adjective functions as an argument of the noun with the lifted type <<f(N),f(N)>,f(N)>. The CAP predicts, then, that in these compounds the adjective may not agree with the noun - a prediction which is borne out: ***stompe-hoekig.**

3. Conclusion

The cases of complement inheritance, passive, and verb raising nicely illustrate the interaction between syntax and semantics in a GPSG framework. The crucial assumption was that the interpretation of morphologically complex structures has to obey the compositionality principle. Metarule analyses of inheritance, passive and verb raising were shown to violate the compositionality requirement for the WSG: they all introduced uninterpreted complex lexical nodes, i.e. syntactically complex nodes for which there was no corresponding object in the model. In order to maintain compositionality in the WSG, the set of possible modes of combination in semantic rules was expanded with a version of functional composition.

Functional composition makes the prediction that the only type of expressions that can be interpreted non-locally will be arguments. This prediction appeared to be too strong. In the last section, we investigated relatedness paradoxes involving misplaced modifiers. Here again, one can either enrich the syntactic algebra of the WSG or the semantic algebra. I chose the second approach, and suggested a limited use of type-lifting which converts the misplaced modifiers into arguments. Whereas functional composition seems to have promising explanatory content in the area of complement inheritance, the type lifting strategy remains as yet an auxiliary hypothesis. The misplaced modifier problem, then, is a choice topic for further research on the interaction between lexicon and syntax.

FOOTNOTES

*Versions of this paper were read at the UCLA Linguistic Institute (July 27,1983) and at the Colloquium "Meaning and the Lexicon" (Kleve, August 30, 1983). For comments and criticism I would like to thank Emmon Bach, Jack Hoeksema, Theo Janssen, Ed Keenan, Ewan Klein, Alice ter Meulen, Dick Oehrle, Barbara Partee, Tom Roeper, Sue Schmerling and Anna Szabolcsi. All errors are my own. The visit to UCLA was partially supported by a grant from the Netherlands Organization for the Advancement of Pure Research (ZWO), which I hereby acknowledge.

[1] In the remainder, syntactic rules are written down in the form of node admissibility conditions. By convention, X stands for X^0. V[n] stands for n-place predicate.

Notice that the word structure rules introduce sets of affixes, not just one affix. In this respect, a WSG rule like (3) is just like a PS rule introducing the set of transitive verbs: the PSG doesn't contain a special rule introducing **hit**, and another one for **kill**. The word structure rules given here abstract away from issues of productivity. One could say that a fully productive affix is a total function, giving a value for every element in its domain; whereas affixes of limited productivity can be regarded as partial functions: for some bases, the base+affix combination is undefined.

[2]For the sake of concreteness, type assignments from Klein & Sag (1982) will be adopted here. Below is a partial listing. <f(NP),f(S)> is shorthand for <<s,<<s,<e,t>>,t>>,<s,t>>, etc. For any two types a and b, the denotation of <a,b> is the set of all functions from the denotations of a to the denotations of b.

f(S)	=	<s,t>	
f(NP)	=	<s,<<s,<e,t>>,t>>	
f(N1)	=	<s,<e,t>>	
f(VP)	=	<f(NP),f(S)>	
f(AP[+PRD])	=	<f(NP),f(S)>	(predicative AP's)
f(V[1])	=	<f(NP),f(S)>	(intransitive verbs)
f(V[2])	=	<f(NP),<f(NP),f(S)>>	(transitive verbs)

[3]A richer inventory of permissible operations in the syntactic algebra is defended in Schmerling (1983). In this paper, I limit the attention to derivational morphology. It is not clear whether the surface compositional approach can be extended to inflectional morphology without undue stretching of the semantic algebra. See Carlson (1983) for relevant discussion.

[4]If Y' is required to be of type <a,b>, only one argument can be inherited. In its effect, the liberal formulation of (15) is equivalent to Geach's (1972) rule. For a function category a/b, this rule allows one to divide numerator and denominator by the same category, resulting in a category (a/c)/(b/c). Repeated application of the rule gives the effects of (15). Notice that composition is defined for the semantic calculus of types, not for syntactic categories. This will prove useful later on (cf. note 6).

[5]Category-neutral inflectional affixation seems to imply transparency: it is difficult to imagine a tense morpheme which could only be applied to basic intransitive verbs of type <f(NP),f(S)> and not to verbs corresponding to more complex functions. But in derivational morphology, the situation is unclear. In Moortgat (to appear) it is suggested that **of**-phrases in nominalizations such as **the finder of the treasure** should not be treated as arguments inherited through f*, but rather as free modifiers which get their argument interpretation indirectly.

[6]In the phrase structure framework adopted here, the syntax of (24) explicitly builds a lexical $[V\ V]_V$ cluster; this syntax forces the use of f* in the semantics, in order to satisfy the rule translation principle. In a categorial framework, the syntactic and the semantic aspect of the phenomenon are not so easily separated. Take **wil** of category VP/VP and **een lied zingen**, a complex expression of category VP. The CG combination VP/VP . VP = VP gives rise to the ungrammatical sentence below, and must be excluded by stipulation.

dat $[\text{ik}]_{NP}$ $[\text{wil}]_{VP/VP}$ $[\text{een lied zingen}]_{VP}$
that I want a song sing

[7]For a radical use of function-argument reversal in syntax, see Steedman (to appear).

REFERENCES

ARONOFF,M. (1976) Word Formation in Generative Grammar. Cambridge MA.

BACH,E. (1980) In defense of passive. L&P 3.

BACH,E. (1983a) Werkwoordclusters in het Nederlands en het Duits. A'dam lecture.

BACH,E. (1983b) On the relationship between word-grammar and phrase-grammar. NL&LT 1.

CARLSON,G. (1983) Marking constituents. In Heny & Richards (eds.) Linguistic Categories: Auxiliaries and Related Puzzles. Vol. 1. Dordrecht.

DOWTY,D. (1978) Governed transformations as lexical rules in Montague grammar. LI9.

DOWTY,D. (1979) Word Meaning and Montague Grammar. Dordrecht.

FLICKINGER,D. (1983) Lexical heads and phrasal gaps. In Barlow, Flickinger & Westcoat (eds.) WCCFL 2. Stanford.

GAZDAR,G. (1982) Phrase Structure Grammar. In Jacobson & Pullum (eds.) The Nature of Syntactic Representation. Dordrecht.

GAZDAR,G., E.KLEIN, G.PULLUM & I.SAG (1982) Coordinate structure and unbounded dependencies. In Barlow, Flickinger & Sag (eds.) Developments in GPSG. Stanford.

GAZDAR,G., E.KLEIN, G.PULLUM & I.SAG (to appear) English Syntax.

GEACH,P. (1972) A program for syntax. In Davidson & Harman (eds.) Semantics of Natural Language. Dordrecht.

HOEKSEMA,J. (1981) Verbale verstrengeling ontstrengeld. Spektator 10.

JANSSEN,Th. (1983) Foundations and applications of Montague Grammar. Amsterdam.

KEENAN,E. (1974) The Functional Principle. In CLS 10.

KEENAN,E. (1980) Passive is phrasal (not sentential or lexical). In Hoekstra, van der Hulst & Moortgat (eds.) Lexical Grammar. Dordrecht.

KLEIN,E. & I.SAG (1982) Semantic type and control. In Barlow, Flickinger & Sag (eds.) Developments in GPSG. Stanford.

LAMBEK,J. (1961) On the calculus of syntactic types. In Jakobson (ed.) The Structure of Language and its Mathematical Aspects. Providence.

LIEBER,R. (1980) On the Organization of the Lexicon. MIT Diss.

LIEBER,R. (1983) Argument linking and compounds in English. LI 14.

MONTAGUE,R. (1970) Universal Grammar. Theoria 36.

MOORTGAT,M. (1981) Subcategorization and the notion 'lexical head'. In Daalder & Gerritsen (eds.) Linguistics in the Netherlands. Amsterdam.

MOORTGAT,M. (to appear) Functional composition and complement inheritance. In the Proceedings of the Colloquium "Meaning and the Lexicon" (Kleve, 30/8-2/9 1983).

PESETSKY,D. (1983) Morphology and logical form. USC MS.

PULLUM,G. & G.GAZDAR (1982) Natural languages and context-free languages. L&P 5.

SCHMERLING,S. (1983) A new theory of English Auxiliaries. In Heny & Richards (eds.) Linguistic Categories: Auxiliaries and Related Puzzles, Vol. 2. Dordrecht.

SELKIRK,L. (1982) The Syntax of Words. MIT Press.

STEEDMAN,M. (to appear) A categorial theory of intersecting dependencies in Dutch infinitival complements. In De Geest & Putseys (eds.) Proceedings of the International Conference on Complementation. Dordrecht.

TOMAN,J. (1981) Wortsyntax. Ph.D. Diss.

WASOW,T. (1980) Major and minor rules in Lexical Grammar. In Hoekstra, van der Hulst & Moortgat (eds.) Lexical Grammar. Dordrecht.

WILLIAMS,E. (1981) Argument structure and morphology. LR 1.

WILLIAMS,E. (1982) On the notions 'lexically related' and 'head of a word'. LI 12.

ZWARTS,F. (1983) Categoriale grammatica - Uhlenbeck revisited. Amsterdam TIN lecture.