It is pointed out that the compositionality constraint depends on independent restrictions on the form of the grammar; in particular on the way morphological features are incorporated. In order to prevent that this causes the compositionality constraint to lose content, we sketch a framework in which the rôle of features in syntax is restricted.

This paper focuses on the interplay between the compositionality principle and morphological features. It is part of a larger project in which we investigated some of the requirements on the organization of the grammar that the principle of compositionality gives rise to. As such, this paper may be regarded as a further discussion of some issues that were raised in Landman & Moerdijk (1983).\(^1\) There we have discussed some restrictions on rules in Montague Grammar, the most important of which was the compositionality principle. We will point out below (as we have done in (1983)) that although the compositionality principle is a wellmotivated constraint and plays a central rôle in the relation between syntax and semantics, it is a purely methodological and – a priori – empirically empty principle when it is considered independently from other restrictions on the grammar. The compositionality principle becomes a ‘real’ constraint on possible grammars only when other constraints are imposed.

Thus, each time one introduces a new component in the grammar, one should wonder whether, or how, the newly introduced elements affect the old restrictions. In particular, if a mechanism is introduced to deal in a precise and formal way with operations on morphological features like gender agreement and case assignment, it is necessary to put some constraints on what features are and what the rules can do with them. For if not, how can

---

\(^1\) Preliminary versions of this paper and of Landman & Moerdijk (1983) both appeared as part of Landman & Moerdijk (1981).
we be sure that the newly introduced elements cannot be used to bypass the restrictions imposed on other parts of the grammar, and hence make the compositionality constraint an empty requirement?

Precisely this is the problem that we aim to discuss in this paper. We will first consider what morphological features should and should not be used for. This will lead to some informal restrictions on their use. In the bulk of this paper we will on the basis of this informal discussion propose a way of formally incorporating features in the grammar.

We will start, however, with a discussion of the framework that we assume. After ten years of Montague Grammar it is not an exaggeration to say that the hardest, if not the only, core of it is the principle of compositionality of meaning. Giving up this principle requires strong arguments. In fact, it can often be heard that the principle of compositionality is the only guarantee that what one is doing is semantics (as opposed to just translating into some formal language).

The principle of compositionality is, in a very informal formulation, the principle that the meaning of a complex expression is a function of the meanings of its constituent parts and the way they are combined. The idea that some form of this principle is an essential part of any theory about how new sentences can be interpreted on the basis of knowledge about old sentences can already be found in Frege's later work (cf. the opening sentences of Frege (1923)).

In Landman & Moerdijk (1983) we considered a particular formalization of the principle as a methodological constraint on possible grammars: "the principle admits only these grammars in which the meaning of an expression $\alpha$ can be calculated by means of a restricted and well defined class of operations, from the meanings occurring in the derivation of $\alpha$, together with the operations on these meanings induced by the syntactic rules used in the derivation of $\alpha$" (see Landman & Moerdijk (1983), p. 90-1). In a compositional grammar the meaning of an expression depends only on the meanings of its parts and the semantic operations used, not on syntactic concepts (like how many t-nodes (S-nodes) the expression contains), nor on derivational concepts (like how many rules there were used in the derivation of that meaning), nor intermediate structures like formulas of the logical language used to refer to these meanings (like how many variables the translation contains). The latter is the main issue in Landman & Moerdijk (1983).

However, we have argued that the compositionality constraint in isolation is empirically empty. If we do not put independent restrictions on the syntactic and semantic rules of the grammar, the compositionality cons-
traint can be satisfied while the meaning of expressions still indirectly depends on anything whatsoever. This happens, for example, if we can encode aspects of derivations (like rule ordering) in syntactic rules or conditions there on, or if we allow 'non-semantic' operations (like 'change the meaning of father into that of chair'). The compositionality principle has no content without independent restrictions on the form of grammatical rules and their interrelation.

In Montague Grammar (MG) syntactic rules build up complex expressions out of less complex ones, while simultaneously the semantic interpretations of these expressions are built up. In Montague (1973) (henceforth PTQ) the semantic interpretation is mediated by a translation function which translates every syntactic expression into an expression of intensional type logic (IL), and associates with every syntactic operation an operation on the translations of the input expressions. Because every such translation has a modeltheoretic interpretation (an intension) and every operation on IL-translations used comes from an operation on meanings (objects in the model), the compositionality principle is satisfied.

Although in PTQ syntactic rules map strings to strings, it has become common to replace such strings by labelled bracketings, and a discussion of the reasons for this change seems unnecessary. An important consequence is that we can now refer to the 'syntactic structure' of the output of syntactic rules without making reference to derivational history, thus opening the door for syntactic conditions on rule application.

A rule of the grammar then is a triple \( R = (F^R, G^R, S^A_R) \), where \( F^R \) is a syntactic operation, \( G^R \) a semantic operation, and \( S^A_R \) a structural analysis, or alternatively, a condition on the application of the rule \( R \). We assume that the syntax is based on a categorial grammar, with a correspondence between syntactic categories and semantic types. We further adopt the following restrictions from Landman & Moerdijk (1983): there are only two kinds of basic operations, namely concatenation which builds up a complex labelled bracketing (henceforth \( \text{lb} \) \( [\mathcal{Z}_{\alpha/\beta} [\xi]_{\beta}]_{\alpha} \) out of \( \text{lb}'s \) \( [\mathcal{Z}_{\alpha/\beta}]_{\alpha} \) and \( [\xi]_{\beta} \), and (structure preserving) substitution which substitutes an lb \( [\xi]_{\alpha} \) in an lb \( [\phi]_{\beta} \) for some sublb's of \( [\phi]_{\beta} \). The semantic operation corresponding to concatenation is just functional application \( (\xi' (\,^\xi)) \), while the operation corresponding to substitution is a substitution rule of IL substituting some terms for free variables by means of \( \lambda \)-abstraction. Conditions on application are taken to refer to properties of the syntactic structure of a labelled bracketing; this includes possible reference to indexed pronouns. (In the second part of this paper we will give a formal definition of 'syntactic structure'.)
We can now return to the problems that this paper aims to discuss. Suppose we introduce a formalism for morphological features and precise morphological operations within the framework sketched above. We will then have rules that enrich labelled bracketings with morphological information; or we have lb's that are from the start enriched with morphological information, and rules that operate on that information. Now if the conditions on rule application can refer to aspects of enriched syntactic structures, it is not clear whether the earlier constraints are in any sense of the word still restrictive. For example, if we can encode information about the derivation in morphological features, and the conditions on rule application can refer to the (non-)occurrence of features, then after all applicability of our rules would be sensitive to derivational history: we could encode rule ordering in features, we could use features as syntactic filters, etc. To give an extreme example, if we have a rule which introduces a feature in a structure if that structure consists of more than 25 nodes, and we have another rule which, relative to that feature, deletes the whole structure, then we have made a filter, and we should seriously ask ourselves what content there is left for the statement that 'the meaning of the latter is a function of the meanings of its parts'.

More realistically perhaps, we have to make sure that it is impossible for features to be used in the way that so-called syntactic markers are. Syntactic markers in the sense meant here\(^2\) are introduced at some level of the derivation and, if necessary, deleted at another level. A marker represents a certain syntactic property that a structure has during that part of the derivation at which the marker is present. One of the main problems that such a notion of marker gives rise to is, that various aspects of derivational history can be encoded. For example, rule ordering is easily introduced by means of markers, as was explained in Landman & Moerdijk (1981). We have argued there and in Landman & Moerdijk (1983) that the syntactic properties that such markers were meant to represent are more explicitly expressed by use of structural conditions on rule application, and that such conditions do not lead to the difficulties that markers lead to.

To avoid such problems, we have to restrict the rôle played by features in several ways. In the first place, given the fact that it is necessary to enrich labelled bracketings with morphological features, our notion of syn-

---

\(^2\) Examples of such markers can be found in the papers by Bennett, Partee, Rodman, Thomason, all in Partee (1976). For discussion of markers, see Landman & Moerdijk (1981).
tactic structure seems to have changed. But when we say that structural conditions on rule application can refer to syntactic structure, we want this to be syntactic structure in the original, narrower sense: such conditions refer to category labels, occurrences of sublb's like indexed pronouns, and treestructure properties like ‘c-command’, but not to features. This means that we have to distinguish between ‘syntax proper’ and ‘enriched syntax’ as we shall do formally later on. It will be clear that when syntactic rules cannot ask for the (non-)occurrence of certain features, it is already much harder to use the information about derivational history that could possibly be contained in features in the way that, for example, syntactic markers were used.

Secondly, we have to put some restrictions on the form and function of features themselves. We take features to have a strictly morphological function: they determine the morphological form that derived sentences should finally have, and thus they are to serve as inputs for morphological operations mapping completed labelled bracketings onto final forms, phonological representations. In particular, features are not meant to encode purely syntactic structure (in the ‘narrow’ sense), and one has to take care that morphological features do not, as a byproduct, act as a syntactic filter. To this end, we should at least require that features, once added to a structure, cannot be deleted or changed. Together with our first requirement on the form of syntactic conditions on rule application this already considerably restricts the ‘coding capacities’ of features. For example, even if we take a feature to stand for the place in the derivation where it was introduced, we cannot at a later stage in the derivation make ‘syntactic’ use of this information, because at the level of syntax there is nothing we can do with that feature. This distinguishes features from the syntactic markers mentioned above.

Of course, in some sense features do contain purely syntactic information. For example, the feature nominative for case on a nounphrase informs us that this nounphrase occurs as the subject of a tensed verb. The point we want to make is that it should not be possible to use this information contained in the case feature at any level other than when the final morphological realization of this nounphrase is determined.

How is the feature structure of what we just called enriched labelled bracketings built up by the syntactic rules? Some of the features, like those for gender, we assume to be generated in the lexicon. Others, like those for case, are introduced through the application of syntactic rules. Thus, the morphological form of a sentence is completely determined by the lexical properties of its components, the syntactic structure of that sentence, and the
way it is built up. By application of syntactic rules, the feature structure is built up in a compositional way. Features are combined as if they are unstructured wholes, i.e. without using their internal structure. Thus, although the 'feature part' of a labelled bracketing must obviously have some internal structure, this structure cannot be used at the level of syntactic rules, and it is only at the level of 'phonological representation' that this structure is unraveled by the rules for realization of morphological features. At earlier levels (application of syntactic rules) morphological operations act on feature structures in a uniform way.

We will now incorporate morphological features in the grammar from Landman & Moerdijk (1983) as sketched above. We will begin by introducing a notion of enriched labelled bracketing, and describe how the purely syntactic structure can be extracted from such lb's. Furthermore, it will be indicated how morphological rules map enriched labelled bracketings onto final forms. The main discussion, however, concerns the formalization of the morphological operations that are performed by syntactic rules. We will formalize the operations of feature agreement and assignment by using the operations of union and substitution, and we will argue that these satisfy our earlier requirements. Finally, we will extend our operations to cases of non-local agreement and assignment, such as gender agreement between antecedents and anaphors. We will show that these phenomena can be handled by the same mechanism (union and substitution) as the local cases.

How are features introduced in the grammar? We assume that there is a language dependent set $F$, of feature sets. For English this set $F$ would contain at least the feature sets $F_g$, $F_n$, $F_p$, $F_c$ for gender, number, person, and case, respectively. Some of these are specified in the categorial lexicon as properties of lexical items (gender, for example), and some of them are introduced by syntactic rules (for example, case). To fix some terminology let us put:

\[
\begin{align*}
F_g &= \{\text{male, fem, neut, } 0_g\} \\
F_n &= \{\text{sing, plur, } 0_n\} \\
F_c &= \{\text{nom, gen, dat, acc, } 0_c\} \\
F_p &= \{1, 2, 3, 0_p\}
\end{align*}
\]

Every feature set $F_i$ contains a neutral feature $0_i$, which will be interpreted by the morphology as the empty instruction, having no effect at all.
A lexical entry like \([\text{man}]_{\text{CN}}\) will then get features male for gender, singular for number and 3 for person in the lexicon. A feature for case is kept undetermined until \text{man} shows up as part of the object the man of a transitive verb phrase, where it gets accusative, or the subject of an intransitive verb in which case it gets nominative case. We will use feature variables to indicate that a certain feature is not yet determined. One does not see all this in English of course, except on pronouns, because the English inflectional system has degenerated. Similarly, determiners in English do not show inflection for person. Still, in order to describe how the system would be applicable to languages like French or German and many others having a more elaborate inflectional system (think of gender showing on French determiners and case and gender on determiners in German), we here choose to give English a richer ('underlying') inflectional system than it actually seems to have. To express formally what sort of features a lexical entry or a more complex labelled bracketing has, we introduce feature sequences.

**definition:** A feature sequence \(f\) is a sequence \(\langle f_i \rangle_{i \in I}\), with for each \(i \in I\), either \(f_i \subseteq F_i\) or \(f_i = x_i\).

Here \(\{F_i : i \in I\}\) is a fixed enumeration of \(F\). Thus a feature sequence is a sequence whose entries are sets of features or variables ranging over such sets.

This suggests that enriched labelled bracketings will look like \([\alpha]_{C,f}\), where \(C\) is a category-label and \(f\) is a feature sequence, and \(\alpha\) is either a lexical item or sequence of labelled bracketings. At a final stage morphological rules will convert such a structure into a 'surface string' or 'morphological expanded' labelled bracketing. For example, a pronoun having features male for gender, sing for number, objective case and third person corresponds to the lb \([\text{PRO}]_{T_r} <\text{(male)}, \text{(sing)}, \text{(acc)}, \text{(3)}, \ldots>\), and is 'realized' by the morphology as the pronoun him, while a structure like \([\text{PRO}]_{T_r} <\text{(fem)}, \text{(sing)}, \text{(nom)}, \text{(3)}, \ldots>\) is realized as she.

If we pursue these ideas a bit further, however, it turns out that matters are somewhat more complex. One of the major difficulties is caused by the fact that syntactic rules should perform certain operations on feature sequences. For example, if you apply the syntactic rule combining a transi-
tive verb and a termphrase to make an intransitive verb, objective case must be assigned to (the feature sequence of) the termphrase. Thus syntactic rules have to take care of case-assignment.

The operation on feature sequences used for case-assignment should work uniformly for all transitive verbs; i.e., assigning objective case to John in [hit John]TV should work exactly the same way as assigning objective to a book in [[give Mary]TV [a book]]TV. Hence give Mary must assign objective case to a book. Now if one does not want objective case appearing out of thin air, it is only reasonable that it comes from give. It is a lexical property of give that it first takes a termphrase as indirect object and assigns it dative case, and then takes a termphrase as direct object and assigns it objective case. This means that we have to explain how this assignment of objective case ‘climbs up’ from give to the outside of the transitive verb [give Mary]TV. Note that it is essential that the morphology cannot search for information inside [give Mary]TV. It cannot look for the verb give, because these rules should be uniform, i.e. they should operate on simple TV’s in exactly the same way as they do on complex ones.

Moreover, if we combine a termphrase of the form

[[...]]TV[[...]]CN[[...]]CN[[...]]CN[[...]]CN

with a transitive verb, objective case will be assigned uniformly to all nodes: not only the top-T-node receives case, also nodes within this structure receive objective case. Traditionally this is dealt with through a rule of feature-percolation: features assigned to the top-node ‘percolate’ down to the dominated nodes, with restriction that percolation cannot cross t or T-nodes (S, NP). To give an example of this restriction, in German für assigns accusative case while wegen assigns genitive case. The grammar has to make sure that in the structure

[[für [den Aufenthalt [wegen [des schlechten Wetters]]]]]

accusative is assigned only to den und Aufenthalt, not to Wetter. We will show that the use of variables assures that feature assignment is automatically ‘clause-bound’. This seems to be an advantage of a grammar that builds up structure through concatenation rules instead of top-bottom rewrite rules.

Assignment is a feature phenomenon that is traditionally distinguished from another one, viz. agreement. For example, adjective and common noun have to agree in number, gender and person; and there is subject-verb agreement in person and number. Thus a termphrase like [[la]TV[[belle]CN[[fille]CN]CN]CN]TV is wellformed only if all three lexical
items show the same gender, number, and person, and also the same case, as can be seen from a German example like

\[(\text{das Buch . .}[[\text{des}]_T/\text{CN}[[\text{kleinen}]_{\text{CN}/\text{CN}}[[\text{Mannes}]_{\text{CN}}]_{\text{CN}}]_T.\gen.\gen.\gen.\]

In order to describe these phenomena of assignment and agreement adequately, we will make our labelled bracketings more complex.

The underlying idea is very simple: for functional categories A/B labelled bracketings \([\alpha]_{A/B}\) will not only have a feature sequence containing information about the morphological realization of \(\alpha\) itself, but also a feature sequence containing information about what features have to be assigned to an lb \([\beta]_B\), when \([\beta]_B\) is combined with \([\alpha]_{A/B}\) to build up an lb \(((\alpha)_{A/B}([\beta]_B))_A\). Thus \([\alpha]_{A/B}\) will not only have a feature sequence \(f\) as above, but also a feature sequence \(g\) for assignment. This results in labelled bracketings of the form

\([\alpha]_{(A/B)g} f\).

As the give-example above illustrates, this should go arbitrarily deep in order to get lb's like \([\text{give}]_{(\text{IV/Tl})/Tgh}\) where \(g\) contains the information that the indirect object receives dative, and \(h\) the information that the direct object receives objective. Formally, we then get

definition: 1. prelabels are defined inductively by
   (i) if \(C\) is a basic category\(^4\), then \(C\) is a prelabel.
   (ii) if \(A\) and \(B\) are prelabels and \(g\) is a feature sequence, then \(A/Bg\) is a prelabel.
2. labels are defined by:
   if \(P\) is a prelabel and \(f\) is a feature sequence, then \(Pf\) is a label.
3. a labelled bracketing is a structure \([\alpha]_\ell\) where \(\ell\) is a label, \(\alpha\) is lexical, or \(\alpha\) is \(AB\) where \(A\) and \(B\) are labelled bracketings.

It is easy to see what this definition says, if you recall the definition of categories of a categorial grammar: basic categories are categories, and if \(A\) and \(B\) are categories then \(A/B\) is a category. You can then make a label out of

\(^4\) This, of course, is only meaningful if it has been specified what the basic categories are. Here we take them to be T, t, IV, PP, CN. We assume that a more elaborate theory can motivate a choice.
a category by inserting feature sequences as subscripts to all categories occurring in this category, except for those immediately to the left of a slash. For example, take the category \(((t/T)/IV)\) and insert feature sequences to get \(((t/T_f)/IV_{g+h})\), which is a label. (To show formally that it is a label we have to show that \(((t/T_f)/IV_g)\) is a prelabel, i.e. that \(t/T_f\) and \(IV\) are prelabels; \(IV\) is a basic category, hence a prelabel; and \(t/T_f\) is a prelabel because \(t\) and \(T\) are basic categories.)

Note that given this definition of prelabels we do not have categories as such directly available in the grammar. This runs against principles stated earlier, saying, for example, that conditions on applicability of syntactic rules should not refer to features, and hence not to labels, while obviously they have to refer to categories. To avoid this problem we have to abstract categories form labels.

**Definition:** for a prelabel \(P\) we define \(c(P)\), the category associated with \(P\), inductively by:

(i) If \(P\) is a basic category \(c(P) := P\)

(ii) If \(P\) is \(A/B_g\), then \(c(P) := c(A)/c(B)\)

For a label \(\ell\) we then define \(c(\ell)\) to be \(c(P)\), where \(\ell = P_g\).

The structure that results from a labelled bracketing \(A\) if we replace all labels \(\ell\) in \(A\) by their associated categories \(c(\ell)\) will be called the *syntactic structure* of \(A\). Thus, saying that structural conditions on rule application, only refer to syntactic structure of lb’s implies that they cannot ask for the (non-)occurrence of morphological features. In this way enriched structure and purely syntactic structure are distinguished.

Another notion that will prove useful below is that of *relevant feature*. The idea is simple: verbs in English do not have case, so the feature for case is not relevant for the category IV. We let the grammar assign to each category \(C\) a set \(r(C) \subseteq I\) of relevant features for \(C\), presumably somewhere in the lexicon. Of course the assignment of sets of relevant features to categories is language dependent. For expository purposes (cf. the remark on French and German determiners above) we will not strictly adhere to this assignment, and suppose more features to be relevant for a certain category (for example, gender for determiners) than can be justified for English. (This is by no means essential: it is easy to reformulate the rules, for example, without having gender and person as relevant for determiners.)

Before we formulate morphological operations we will briefly consider the character of morphological rules. Morphological rules are rules
that adjust the form of lexical strings of a certain category with a certain feature sequence. One may think of the morphological component as a machine interpreting certain instructions expressed by the categories and the feature sequences. A particular combination of instructions will result in a complex instruction, something like 'realize a pronoun (of category T) which is male, singular, third person, with objective case as *him*'. Assuming the morphological component to look like this, it is very reasonable to say that morphological rules do not 'react' to every instruction. For example, they cannot realize a termphrase as having simultaneously nominative and objective case. Neither will they produce a proper output if an instruction is lacking where there should be one, as e.g. 'realize a pronoun without person'. Wellformed outputs of this process taking place in the morphological component then are structures with all features 'interpreted' (by the machine, whence all lexical items being expanded; note that the machine – as indicated below – works inwards).

Implicit in all this are properties of morphological rules like the following:

Let \([\alpha]\) be an Ib with outer feature sequence \(f = \langle f_i \rangle_{i \in I}\) and let \(M\) be a morphological rule 'interpreting' the k-th feature \(f_k\). Then

(a) If \(f_k = \phi\) and \(k \in r(c(P))\) then \(M\) deletes \(\alpha\).
(b) If \(f_k \neq \phi\) and \(f_k\) is not a singleton (a set with exactly one member) then \(M\) is not applicable.
(c) If (a) and (b) do not apply then

(i) If \(\alpha\) is lexical, then

- either 1. \(M\) maps \(\alpha\) onto itself, if \(k \notin r(c(P))\), (so, irrelevant features are ignored)
- or 2. If \(k \in r(c(P))\), \(M\) expands \(\alpha\) as prescribed by the instruction relative to \(f_k\).

(ii) If \(\alpha\) is of the form \(\beta\gamma\) (Ib’s), \(M\) maps \(\alpha\) onto the concatenation of \(M\) applied to \(\beta\) and \(M\) applied to \(\gamma\).

(a) states that an Ib lacking a relevant feature is realized as empty in the morphology. In (b) the non-applicability is of course caused by an ambiguous instruction. As indicated above, the morphology can’t process such an instruction and the sentence is ruled out. We will give examples of this below. (c (i) 2) is the place where classical morphological paradigms fit into our grammar.
Let us now go back to labelled bracketings and the question of what syntactic rules can do with them. We start with concatenation.

Before we can say anything about what concatenation rules look like when operations on feature sequences are included, we first have to consider which operations can be used. Since the members of a feature sequence \(\langle f_i \rangle_{i \in I}\) are sets or variables for sets, we can, in principle, perform all sorts of set theoretic operations on them. However, since we do not have the operations refer to the internal structure of features, we have to restrict ourselves to purely algebraic operations. Here some wellknown operations present themselves, viz. the boolean operations of intersection, union, and complementation, respectively,

\[
\begin{align*}
\text{f}, g &\mapsto f \cap g = \{ x : x \in f \text{ and } x \in g \} \\
\text{f}, g &\mapsto f \cup g = \{ x : x \in f \text{ or } x \in g \} \\
\text{f}, g &\mapsto f - g = \{ x : x \in f \text{ but not } x \in g \}.
\end{align*}
\]

However, even the boolean operations are already too strong for our purposes. A difference between the union operation on the one hand and the operations of intersection and complementation on the other, is that the union operation never deletes members of \(f_i\) or \(g_i\), while the intersection and complementation operations may do so. Now, it was stated above that features could not be deleted or changed. For this reason we exclude complementation and intersection:

*The only Boolean operation on feature sets is the union operation.*

For ease of notation we define for feature sequences \(f = \langle f_i \rangle_{i \in I}\) and \(g = \langle g_i \rangle_{i \in I}\),

\[
f \cup g = \langle f_i \cup g_i \rangle_{i \in I}.
\]

For assignment we will use the variables \(x_i\) in feature sequences. These variables have a function similar to that of indexed pronouns in PTQ for which ‘real’ termphrases are substituted. For example, the variable \(x_c\) should be thought of as keeping case undetermined in a labelled bracketing like \([ [a] c \langle \ldots x_c \ldots \rangle [B] d \langle \ldots x_c \ldots \rangle ] T \langle \ldots x_c \ldots \rangle\) until this lb is combined with another lb, say a transitive verbphrase, to build an IV. The TV then assigns objective case to the termphrase, which means that the TV substitutes ‘real’ values for the variables \(x_c\) in the argument lb.

Formally, we define for a labelled bracketing \(A\) and a feature sequence \(g\), \(A [g] \) to be the labelled bracketing which comes from \(A\) by replac-
ing each $x_i$ in $A$ by $g_i$. Similarly, for feature sequences $f$ and $g$, $f[g]$ is the
feature sequence which comes from $f$ by replacing each $x_i$ in $f$ by $g_i$. We will
use the notation $A[g_i/f_i]$ for substitution of only feature $f_i$.

Note that substitution of features of sort $i$ is automatically uniform in
this system, because for each $i$ we have introduced only one variable $x_i$ (and
not, for example, distinct variables $x_i, y_i, z_i, \ldots$).

The union operation and the substitution operation are the only oper-
ations on features in the system. The union operation will be used to capture
agreement phenomena, while the substitution operation will handle assign-
ment. This means that we will distinguish assignment from agreement by
introducing different kinds of concatenation rules. As far as syntax proper is
concerned the rules are identical (just concatenation). For agreement, we
define a concatenation rule $\oplus_1$ by

\[(\alpha)_{(A/B)\beta h} \oplus_1 (\beta)_{B_k} \Rightarrow [(\alpha)_{(A/B)\beta h} (\beta)_{B_k}]_{A_k[e]} \cup h.\]

The only instances of (1) we consider here are those which have $g = x = \langle x_i \rangle_{i \in i}$. In fact, only these cases can properly be called agreement,
since then $[\alpha]_{(A/B)\beta h}$ does not assign anything to $[\beta]_{B_k}$, because substitution
of variables has no effect at all. By abuse of notation we will omit the
sequence of variables to be ‘assigned’ and write these instances of $\oplus_1$ as

\[(2) [\alpha]_{(A/B)\beta h} \oplus_1 (\beta)_{B_k} \Rightarrow [(\alpha)_{(A/B)\beta h} (\beta)_{B_k}]_{A_k[e]} \cup h.\]

(abuse, since $(A/B)\beta h$ is not a labelled bracketing).

When referring to the agreement rule $\oplus_1$, we will alway mean (2), i.e. the
special case of (1) with $g = x$.

To illustrate the agreement rule let us consider two combinations of
old and man: Putting

\[[\text{old}]_{(\text{CN/CN})}, \langle \text{male}, \text{[sing]}, x_e, \phi, \ldots \rangle \quad (\text{abbreviated as 'old'})\]

and

\[[\text{man}]_{\text{CN}, \langle \text{male}, \text{[sing]}, x_e, \{3\}, \ldots \rangle} \quad (\text{abbreviated as 'man'})\]

together by means of $\oplus_1$ gives us

\[\langle \text{old 'man'} \rangle_{\text{CN}, \langle \text{male} \cup \text{male}, \text{[sing]} \cup \text{[sing]}, x_e \cup x_e, \{3\} \cup \phi, \ldots \rangle}\]

which is, of course, the same as

\[\langle \text{old 'man'} \rangle_{\text{CN}, \langle \text{male}, \text{[sing]}, x_e \{3\}, \ldots \rangle}\]
On the other hand, if we combine ‘man’ with

\[(\text{old})(\text{CN}/\text{CN}) \langle \text{fem}, \text{plur}, x_c, \phi, \ldots \rangle \]  
(call it “old’)

then \(\Theta_1\) gives us

\[\langle \text{“old”} “\text{man}” \rangle_{\text{CN}}, \langle \text{fem} \cup \text{male}, \text{plur} \cup \text{sing}, x_c \cup x_c, \phi \cup \{3\}, \ldots \rangle \]

which is identical to

\[\langle \text{“old”} “\text{man}” \rangle_{\text{CN}}, \langle \text{fem, male}, \text{plur, sing}, x_c, \{3\}, \ldots \rangle \]

This complex lb cannot be interpreted by the morphology, because there are some non-singleton feature sets. In general, morphological rules are applicable only if number and gender of adjective and common noun agree.

To give another example, combine

\[\langle \text{“old”} “\text{man}” \rangle_{\text{CN}}, \langle \text{male}, \text{sing}, x_c, \{3\}, \ldots \rangle \]  
(write ‘old man’)

with a determiner

\[\langle \text{the} \rangle_{\text{T}/\text{CN}}, \langle \text{male}, \text{sing}, x_c, \phi, \ldots \rangle \]  
(write ‘the’)

by means of \(\Theta_1\). We then get

\[\langle \text{the} “\text{old man}” \rangle_{\text{T}}, \langle \text{male}, \text{sing}, x_c, \{3\}, \ldots \rangle \]

As we have said, this doesn’t show up in English, but e.g. in French, we derive in exactly the same way that termphrases like la belle fille are well-formed only if all three lexical items show the same number and gender. In other words, things like *le beau fille, *les belle fille, etc. all are automatically excluded by our system.

Dealing with agreement in this way implies that there have to be different lexical entries for man:

\[\langle \text{man} \rangle_{\text{CN}}, \langle \text{male}, \text{sing}, x_c, \{3\}, \ldots \rangle \]  
(“man’)

\[\langle \text{man} \rangle_{\text{CN}}, \langle \text{male}, \text{plur}, x_c, \{3\}, \ldots \rangle \]  
(“men’)

Adjectives like old will have more lexical forms: all of the forms

\[\langle \text{old} \rangle_{\text{CN}/\text{CN}}, \langle \{y\}, \{z\}, x_c, \phi, \ldots \rangle \]

with \(y \in F_g\) and \(z \in F_n\). This of course implies that the lexicon would contain lots of redundancies. Hence we cannot regard it as a rather unstructured
whole. There should be rules, expressing for example that for modifiers (things of some category A/A) the lexicon contains all variants with arbitrary (singleton-) values for relevant features determined by lexical properties (like gender and number; this in contrast to such features as case, which are determined by structural properties). It is evident that all sorts of regularities concerning the connection between relevant features, categories, and concatenation rules needs further investigation.

We now continue with the concatenation rule dealing with assignment. We define a concatenation operation $\Theta_2$ by

$$
(3) \quad [\alpha]_{(A/B_{\phi}h)} \Theta_2 [\beta]_{B_k} \Rightarrow [[\alpha]_{(A/B_{\phi}h)} [\beta]_{B_k} [g]]_{A_h}
$$

Here $[\alpha]_{(A/B_{\phi}h)}$ acts as a function; h determines the morphological form of $\alpha$, and is inherited by the top-node, which means that it also determines the morphological properties of the output Ib. Feature sequences directly to the right of a slash are assigners: g is assigned to the argument $[\beta]_{B_k}$, i.e. every $x_i$ in $[\beta]_{B_k}$ is replaced in the output by the corresponding $g_i$.

To give an example of $\Theta_2$, recall the structure ‘the old man’, made out of the old and man by two applications of $\Theta_1$. The tree below represents this structure.

$$
\begin{align*}
\text{T}, &\langle \{\text{male}\}, \{\text{sing}\}, x_0, \{3\}, \ldots \rangle \\
(\text{T}/\text{CN}), &\langle \{\text{male}\}, \{\text{sing}\}, x_0, \phi, \ldots \rangle \\
\text{CN}, &\langle \{\text{male}\}, \{\text{sing}\}, x_0, \{3\}, \ldots \rangle
\end{align*}
$$

If we now combine this by $\Theta_2$ with the lb

$$
[\text{with}]_{(PT, \langle \phi, \phi, \{\text{obj}\}, \phi, \phi \rangle \langle \phi, \phi, \phi, \phi, \ldots \rangle} \text{ (obj stands for objective (oblique) case)}
$$

we get the following result: $\langle \phi, \phi, \phi, \phi, \ldots \rangle$ climbs to the topnode, and $\langle \phi, \phi, \{\text{obj}\}, \phi, \ldots \rangle$ is substituted in the argument. Every $x_c$ is then replaced by $\{\text{obj}\}$, and every $x_i \neq x_c$ is replaced by $\phi$. The following tree represents ‘with the old man’. As you can see, case goes down to all lexical items in the complement, which illustrates percolation of features.
As another example, let us consider assignment in a structure like [give Mary a book]_IV, mentioned above. Dative is assigned to Mary and accusative to a book if we represent give in the lexicon as

\[(4) \left[ \right. \left[ \text{give} \right]\right.\]

A first application of \(\Theta_2\) to (4) and \([\text{Mary}]_T, \langle \ldots x_c \ldots \rangle\) assigns dative to Mary by substituting \{dat\} for \(x_c\): we obtain

\[(5) \left[ \right. \left[ \text{give}\right]\left[ \text{Mary}\right]\right.\]

Another application of \(\Theta_2\) to (5) and \([\text{a book}]_T, \langle \ldots x_c \ldots \rangle\) yields

\[(6) \left[ \right. \left[ \right. \left[ \text{give}\right]\left[ \text{Mary}\right]\right.\left[ \text{a book}\right]\right.\]

Several options present themselves for assigning nominative to subjects of tensed clauses. For example, one could introduce an AUX-cateogry ((t/T)/IV) along the lines of Bach (1980), (containing tense, modals, etc.) and let the AUX-node assign nominative to the subject if it is tensed, and no case (i.e. assign \(\phi\)) if it is not tensed. Presumably, the infinitive marker to will then be given as

\[(7) \left[ \right. \left[ \right. \left[ \text{to}\right]\left[ \text{to}\right]\right.\]

Thus \([\text{to}]((t/T), \langle \ldots \phi \ldots \rangle)/IV, \langle \phi, (o_n), \phi, (o_p), \ldots \rangle, \langle \ldots \rangle\) is an AUX which does not assign case to its subject. For example, if one combines this AUX with \([\text{walk}]_IV, \langle \phi, x_n, \phi, x_p, \ldots \rangle\) by \(\Theta_2\), one obtains

\[(8) \left[ \right. \left[ \right. \left[ \text{to walk}\right]\right.\]

(Note that in (8), \text{walk} has neutral features for number and person, as it
should be in English infinitives.) If we concatenate (8) with
\[ \text{[PRO]}_{T, (\text{male}), (\text{sing}), x_c, (3), \phi} \]
we derive

\[(9) \quad \text{[PRO]}_{T, (\text{male}), (\text{sing}), \phi, (3), \phi} \quad \text{'to walk'} \]

Here the termphrase PRO has \(\phi\) at a relevant entry, and will therefore be deleted. Thus subjects of infinitives are present in syntax and semantics, but not in the morphologically expanded form of sentences. Morphological deletion of this kind also plays a role in the discussion of wh-substitution in Landman & Moerdijk (1981), section 3.3.

Thus we have replaced percolation by substitution. We do not have to assume any further restrictions on substitution (like clause-boundness). Substitution will go arbitrarily deep and clause-boundness of, for example, case assignment is automatically fulfilled: in a prepositional or a sentential structure all terms are argument of some case assigner, so at that level all case variables are replaced by sets of case features. If such a structure is again embedded in, say, a prepositional structure, case assignment does not affect the features in the embedded structure, because there are no variables left in that structure. In the example discussed above, the rule that combines \(\text{wegen}\) and \(\text{der}^{(\ldots x_c \ldots)} \text{schlecht}^{(\ldots x_c \ldots)} \text{Wetter}^{(\ldots x_c \ldots)}\) assigns genitive case:

\[(10) \quad \text{wegen des}^{(\ldots (\text{gen}) \ldots)} \text{schlechten}^{(\ldots (\text{gen}) \ldots)} \text{Wetters}^{(\ldots (\text{gen}) \ldots)}\]

If we now build up the termphrase:

\[(11) \quad \text{f\"{u}r}^{(\ldots x_c \ldots)} \text{Aufenthalt}^{(\ldots x_c \ldots)} \text{wegen}^{(\ldots (\text{gen}) \ldots)} \text{schlechten}^{(\ldots (\text{gen}) \ldots)} \text{Wetters}^{(\ldots (\text{gen}) \ldots)}\]

and we combine this with \(\text{f\"{u}r} + \text{acc}\), then accusative is substituted for all case variables, which are only the variables in the highest clause:

\[(12) \quad \text{f\"{u}r den Aufenthalt wegen des schlechten Wetters}\]

The system as presented here satisfies the general requirements on features that we have informally discussed in the beginning of this paper: the operations on features are restricted to union and substitution for variables; features can only be added to lb's, not changed or deleted. We have indicated that distinctions between functions and arguments, and hence distinctions between argument places and non-argument places can be introduced relative to our morphological operations. The system moreover is sufficiently rich to formalize all sorts of feature phenomena (for instance, both subject-verb agreement and object-verb agreement could be formalized). We will however not go into this any further.
Up to now we have only discussed local (i.e. clause-bound) agreement and assignment. But there are also non-local morphological phenomena, and these pose some problems for classical feature theories. Among such phenomena the most important occur in connection with anaphoric relations: an antecedent and its anaphors show gender, number and person agreement; the pronominal form of anaphors (reflexive or not) can be regarded as a form of assignment. It is not clear how an operation like feature percolation, devised to deal with local assignment, can be generalized to non-local assignment without additional (ad hoc) constraints that ensure that the features are assigned to the proper anaphoric pronouns.

One of the main topics in the discussions on pronouns seems to be the question at which level of the analysis the anaphor in a sentence like (13) should be linked to the antecedent (if at all):

(13) *John says that he likes Mary.*

This could take place in the syntax, in the semantics or at an intermediate level (logical form, discourse structure, etc.). We have argued in our (1983) that precisely the phenomenon of non-local agreement shows that this linking should take place at a level at which syntactic information is still present and can be used. In a compositional semantics the only place in the grammar where that is the case is the level of syntactic rules: syntactic rules should link anaphors to antecedents because syntactic rules should perform the operations of non-local agreement and assignment. A strong argument in favor of this is gender agreement between antecedents and anaphors in German: unlike in English the gender here is grammatical gender. At a semantic level (or even a quasi-semantic level like discourse structure) the information that ‘das Mädchen’ is neuter and not female is lost, there is only a female girl left. If anaphoric linking (and hence gender agreement) takes place at this level, one should expect anaphors bound to ‘das Mädchen’ to show female gender. Their gender is neuter, however.

Thus there should be a rule in the syntax, linking anaphors to antecedents and at the same time assigning these expressions some morphological properties. Such a rule is our ANAₙ (Landman & Moerdijk (1983)), which substitutes a term-structure as antecedent in a sentential structure relative to a certain configuration of indexed pronouns (syntactic variables) in that sentence, one of which has a certain structural property, that of being a possible n-antecedent in that sentence (for definitions, see Landman & Moerdijk (1983)).

The result of substitution is a sentential structure in which the pos-
sible n-antecedent $\alpha$ is replaced by the substituted term (the antecedent) and the other occurrences of that indexed pronoun (the anaphors) drop their index and adjust their morphological form relative to both morphological properties of the antecedent and the structural relation between antecedent and anaphor.

One of the properties that anaphors get in this way is their reflexive or pronominal form. To express this more formally we introduce a feature set $F_{pc}$ containing features for pronominal character, $F_{pc} = \{\text{refl, pron, o}_{pc}, \ldots\}$. Another aspect of the substitution rule is that it takes care of feature agreement between anaphors and antecedent: they agree in number, gender and person. We will formalize this by modifying our rule $\text{ANA}_n$ from (1983). In that paper this rule was formulated, without features, as follows:

(14) $\text{ANA}_n$: 
Let $\xi$ be an lb of category $T$, $\phi$ be an lb of category $t$, $\alpha$ a possible n-antecedent in $\phi$ (an occurrence of $\text{PRO}_n$ not c-commanded by any other occurrence of $\text{PRO}_n$ in $\phi$) then $\text{ANA}_{n,\alpha}(\xi, \phi) = \phi'$, where $\phi'$ comes from $\phi$ by replacing $\alpha$ by $\xi$ and replacing the $i^{th}$ pronoun $\text{PRO}_n$ other than $\alpha$ by:
- PRO-self if this $i^{th}$ pronoun is bound in $\phi$ (c-commanded by an occurrence of $\text{PRO}_n$ within the same t-node)
- and by PRO otherwise.
Condition: $\xi = \text{PRO}_k$ iff $k = n$
Translation: if $\xi$ translates into $\xi'$ and $\phi$ into $\phi'$ then $\text{ANA}_{n,\alpha}(\xi, \phi)$ translates into $\xi' (\wedge \lambda x_n \phi')$

With features the rule now looks like this:

(15) $\text{ANA}_n$
Let $[\xi]_T, \langle f_g, \ldots, f_{pc} \rangle$ be an lb of category $T$, $[\phi]_t, \delta$ an lb of category $t$, $\alpha$ a possible n-antecedent in $\phi$, $\alpha = [\text{PRO}_n]_T, \langle g_g, \ldots, g_{pc} \rangle$ then $\text{ANA}_{n,\alpha}(\xi, \phi) = [\phi']_t, \delta$, where $\phi'$ comes from $\phi$ by replacing a by $[\xi]_T, \langle f_g \cup g_g, f_n \cup g_n, f_{pc} \cup g_{pc} \rangle [f_c/g_c, f_{pc}/g_{pc}]$ and replacing the $i^{th}$ pronoun other than $\alpha$, $[\text{PRO}_n]_T, \langle h_h, \ldots, h_{pc} \rangle$ by $[\text{PRO}]_T, \langle f_g \cup h_g, f_n \cup h_n, f_{pc} \cup h_{pc} \rangle [f_c/h_c, f_{pc}/h_{pc}]$ where $c = \{\text{refl}\}$ if this $i^{th}$ pronoun is bound in $\phi$ and $c = \{\text{pron}\}$ otherwise.
Condition: as above
Translation: as above

What this modification says is, that in substituting a termphrase $\xi$ in a sen-
sentence by means of ANAₙ,α the resulting termphrases have two kinds of features. The first kind are those which are determined by (or equivalently, have to agree with) lexical properties of ξ: these are the features for gender, number, and person. The second kind are those features which are determined by the structural positions of the antecedent and its anaphors in φ: these are the features for case and pronominal character. (The substitution of case will always be empty if the pronoun has a position in the sentence as argument of a case-assigner.)

Features for number, gender, and person are base-generated on pronouns as they are on lexical terms. The union operation takes care of agreement in number, gender, and person between antecedent and anaphors. If there is no agreement, they receive a doubleton at their relevant place in the feature sequence, and the sentence is out.

As an example, let us take sentence (16):

(16) PROₙ said [that PROₙ liked PROₙ]

α β γ

Feature sequences: α, β have ⟨{male}, {sing}, {3}, {nom}, xₚc⟩
γ has ⟨{male}, {sing}, {3}, {obj}, xₚc⟩.

The only possible n-antecedent in (16) is α so let us apply ANAₙ,α(ξ, (16)) where ξ = [John]ₚ, ⟨{male}, {sing}, {3}, xₚc⟩.

What happens is this: α is replaced by
[John]ₚ, ⟨{male} ∪ {male}, {sing} ∪ {sing}, {3} ∪ {3}, {nom}, xₚc⟩

[xC/{nom}, {0pC}/xₚc], which is [John]ₚ, ⟨{male}, {sing}, {3}, {nom}, {opC}⟩.

β is replaced by

[PRO]ₚ, ⟨{male} ∪ {male}, {sing} ∪ {sing}, {3} ∪ {3}, {nom}, xₚc⟩

[xC/{nom}, {pron}/xₚc] = [PRO]ₚ, ⟨{male}, {sing}, {3}, {nom}, {pron}⟩

{pron}, because β is not bound in φ. γ is replaced by

[PRO]ₚ, ⟨{male} ∪ {male}, {sing} ∪ {sing}, {3} ∪ {3}, {obj}, xₚc⟩

[xC/{obj}, {refl}/xₚc] = [PRO]ₚ, ⟨{male}, {sing}, {3}, {obj}, {refl}⟩

{refl}, because γ is bound. The morphology will spell this out as (17).

(17) John said that he liked himself.

If in (17), α, β, and γ would have female instead of male on their genderplace, the substitution of John would have resulted in doubleton features, and the sentence is ruled out.
This concludes our discussion of morphological features. Summarizing, our grammar now produces labelled bracketings, exhibiting syntactic information in category labels, and morphological information in feature sequences. Syntactic rules are of two kinds: concatenation, which operates uniformly on all Iβ's of certain specified categories, and substitution, which is restricted by structural conditions on its input. Features and operations on features are restricted to avoid encoding of derivational history: Iβ's contain features and variables (without internal structure), the feature structure is built up simultaneously with syntactic structure through the operations of union and substitution. The system satisfies the requirements set in the introduction, hence the morphology as presented here does not interact with compositionality. As a pleasant consequence, it has been shown how within such a grammar local and non-local feature phenomena can be treated on a par.

REFERENCES


