THE FG EXPRESSION RULES: A DYNAMIC MODEL

Dik Bakker
Universiteit van Amsterdam

ABSTRACT

In Functional Grammar, most attention has been paid to the semantic, pragmatic and discourse aspects of underlying representations. Syntax, morphology and phonology, together: the expression rules (ER), have been relatively neglected. However, apart from being an integrated part of the overall FG model, they are the major instrument to verify and falsify the structure and contents of underlying representations. In this contribution it will be shown that the expression rules as they stand can not properly fulfill that task: they both undergenerate and overgenerate. To make up for at least part of these problems, a dynamic model of the expression rules will be presented that integrates the several aspects of the original proposal, viz. grammatical morpheme generation, order templates and placement rules.

KEY WORDS: Functional Grammar, expression rules, templates, typology.

RESUMEN

En Gramática Funcional, la atención se ha centrado en los aspectos semánticos, pragmáticos y discursivos de las representaciones subyacentes. Por el contrario, el estudio de la sintaxis, morfología y fonología, es decir, de las reglas de expresión (RE) ha recibido menos atención. Sin embargo, aparte de ser una parte integrante de la organización general del modelo de la GF, las reglas de expresión constituyen la herramienta principal para validar o refutar la estructura y los contenidos de las representaciones subyacentes. Este artículo pretende demostrar que las reglas de expresión, en su formulación actual, no cumplen debidamente esta misión, ya que generan tanto de más como de menos. Con la intención de solucionar al menos parte de estos problemas, se presentará un modelo dinámico de las reglas de expresión que integra varios aspectos de la propuesta original, a saber, generación de morfema gramatical, las plantillas del orden de palabras y las reglas de colocación.

PALABRAS CLAVE: Gramática Funcional, reglas de expresión, plantillas, tipología.

1. INTRODUCTION

Since its first proposal in Dik (1978), the main effort in Functional Grammar (FG) has been directed to underlying representations. This has provided the theory with the respective versions of the layered clause model, among other things (cf. Hengeveld 1989, Dik 1997a). More recently, attention has shifted to yet deeper, or higher, levels of linguistic description: pragmatics and discourse (cf. Hannay 1991; Hengeveld 1997; Kroon 1997). These developments are quite natural, given the major tenures and requirements of adequacy held by the theory. However, it is the conviction of the current author that the expression rules —in more down-to-earth terms: syntax, morphology and phonology— have been relatively neglected. While examples of underlying representations abound in the literature, no fully worked out example of a complete expression may be found at all. There may be historical, or psychological, or even practical reasons for this state of affairs. There are, however, several more compelling reasons why it should be deplored, and made up for. Firstly, FG aims at being a complete theory of language, not a semantic theory, or a discourse theory. This implies that all components distinguished within the grammar models or speaker/hearer models proposed by the theory should be developed in a uniform fashion, with a clear idea about the nature of the interfaces between them, and the distribution of the respective grammatical features over them. Only then the theory may be put to the empirical test in the full sense of the word. Secondly, underlying representations in FG contain the semantic and pragmatic information both necessary and sufficient for the computation of the form of the corresponding sentence. This means that an underlying representation may only be correct if for each of its elements we are able to show that it is (co-)responsible for the triggering of at least one expression rule. In other words: we can only be sure about our theory of well-formed underlying representations if we have a complete theory of well-formed expressions. A third reason for having a closer look at expression rules than has been done so far is that, over the last decades, a great amount of knowledge has been acquired concerning the syntax of natural language (cf. Culicover 1997 for a recent introduction to syntactic theory). This enterprise has taken place, to a large extent, within the so-called formal paradigm, and many solutions to problems would be, or are already, catered for outside the realm of syntax in a functional theory. However, other phenomena and insights should be incorporated, mutatis mutandis, in the FG expression component. Finally, and more practically, it has been shown, above all in computational implementations of FG, that the traditional tripartite decomposition of the expression component into generation, linearization and phonematization cannot work properly, even in relatively simple cases (cf. Bakker 1994: 266f and section 3 below).

In FG, syntax and morphology boil down to ordering rules, and do not include any type of structure on any level of the sentence or of (complex) words. All decisions about form and order are taken, at least in principle, on the basis of the semantic and pragmatic information contained in underlying representations, disregarding possible structural aspects. Thus, in discussions on the autonomy of syntax or morphology, FG in its current practice takes a rather extreme functionalist position in terms of Croft (1995: 509).[1] The necessity of so-called auxiliary opera-

[1] Van Valin and LaPolla (1997) devote a 700 page book to the development of a syntactic component for another functionally based theory, Role and Reference Grammar (RRG). In order to

tors in FG makes it clear that a purely one-to-one relationship between function and form is not tenable. Some of the formal universals of the type of Greenberg (1963) might be other manifestations of function independent principles of expression, with implications for the expression rules.

In this paper I will propose a partial reorganization of the Expression Rules of FG (ER), my goal being to make up for at least part of the problems mentioned above. The proposal takes into consideration several of the existent principles of expression and constraints on it. For instance, the familiar templates and placement rules will reappear, be it in a disguised way. The strict separation between function and form is maintained in the sense that no interpretation takes place once expression has sparked off. Transformations will be excluded in the sense that no structures will be generated that have to be altered at a later stage; i.e. there is no movement and no deletion. Filtering, though not ruled out here in principle and in face of the linguistic facts, is restricted as much as possible. A major difference in comparison to the current organization of the expression component is that Stage I —computation of forms— and Stage II —linearization— will be conflated. This integrated version of the expression rules has been introduced in Bakker (1994); it will be worked out further here.[2] Stage III —phonological realization— will only be touched upon in passing, much as it is in the original version of the expression rules (henceforth: the standard model). It remains an open question whether and to what extent functionally oriented proposals to phonology (cf. Boersma 1998) may be integrated into the model of the expression component that I will propose. As another major deviation from the original expression rules, the flat templatewise organization will be replaced by a hierarchical structure, much like underlying representations (UR), or constituent structure trees for that matter. This implies the introduction of notions such as subcategorization, syntactic heads and domains, inheritance, feature percolation and adjacency. Furthermore, the proposal made here has a dynamic as opposed to a static nature. More concretely this means that expressions will be developed in a top down, left to right fashion under the assumption that this order is essential for the type of forms that we may, and may not find in languages. The standard model assumes scope as the leading (and only?) principle for the triggering of operators during expression, providing a centrifugal —i.e. inside out— application of the rules (cf. Dik 1997a: 362, 365f). This may easily give rise to right to left developments of form, which seems to go counter to evidence from psycholinguistics for the left to right approach, which is also advocated here (cf. Carroll 1994: 201f; Levelt 1989).

So far for a brief motivation for the further development of the expression rule component of the FG grammar model, and some of the major characteristics

get their model typologically adequate, quite an amount of syntactic sophistication is called for. Since semantic representations in RRG are not principally simpler than those in FG, it is not clear beforehand why FG could do with less syntax.

[2] In Dik (1997a) it is suggested that a "sandwich" approach to expression might turn out to be necessary. The proposal made in Bakker (1994) could be seen as a concrete elaboration of that idea.

of the extended version that I will propose below. The rest of this paper is organized as follows. In section 2 I will give a short survey of the standard model of the expression rules as presented in the literature. In section 3, some problematic aspects of this model will be discussed on the basis of concrete form and order phenomena from several languages. In section 4 a reformulation of the expression rules will be proposed that takes care of several of the problems observed in section 3. Section 5 gives a somewhat detailed example of the working of the model proposed in section 4 on the basis of an English sentence. Section 6 compares the results with the standard model and discusses some remaining problems.

2. EXPRESSION RULES: THE STANDARD MODEL

In their strictest form, the FG Expression Rules may be defined as follows. Given the (infinite) set of well-formed underlying representations S_U of language L, there is a function or algorithm E_L that maps S_U onto the set of expressions S_E of L. E_L embodies the set of expression rules of L. In the FG literature, the assumptions A1 and A2 are generally made:

A1. Every well-formed underlying representation is expressible.
A2. There is precisely one expression for each well-formed underlying representation.

In other words: there are no formal filters (A1), and there is no synonymy in language L (A2). Below, we will see that there are counterexamples to A1. However, these are rather marginal, and I will assume that A1 may nevertheless be seen as a sound principle in general. A2 would mean that there are, in fact, no pure, neutral options in grammar, and that any difference in form implies a difference in function, i.e. in meaning or pragmatics. Therefore, in the case that the expression rules generate more than one sentence for some underlying representation this would mean that this underlying representation is underspecified, and in fact describes a set of cohyponymous expressions rather than precisely one sentence. In such cases, the underlying representations should be made more specific. This leaves open the possibility that two underlying representations generate the same form.

Furthermore, I will assume that the following also holds, although I am not aware of its having been stated explicitly anywhere in the literature:

A3. Every well-formed expression of L has an underlying representation.

In other words: there are no expressions without meaning. This trivially holds if we assume that the ER component needs a UR for an input, and will not trigger without it. Autonomous productions of ER then fall outside the scope of the full grammar of L. A3 has implications for the reversibility of E_L and the parseability of the FG grammar model. If we indeed assume A1 and A2, then the output of ER is restricted by the constraints on its input —well-formed URs— and by constraints on the expression rules themselves. The latter should be taken in the sense of the way they work rather than that they were to have a filtering function (cf. A1).

As has already been observed above, the rules determining well-formed underlying representations have been given relatively much attention to. In general, we can say that the URs of language L are constrained by URC1 through URC5 below.

URC1: The set of functions and operators of L. This will be a subset of the universal set of functions and operators, further determined by universal and typological constraints, and possible idiosyncrasies of L.
URC2: The scope relations determined by the theory of underlying representations, i.e. the layered clause model.
URC3: Accessibility, i.e. the extent to which functions and operators may be assigned to elements of URs. Some well-known domains are relativization, raising and subject assignment (the Semantic Function Hierarchy).
URC4: The predicates in the lexicon of L, including predicate formation rules.
URC5: Compositional semantics, including selection restrictions on argument and satellite positions.

URs, thus constrained, are fed to the ER component. In this component they are processed in three steps:

Stage I: the computation of forms. Here, the definitive forms are produced. So, these rules represent inflectional morphology. They also take care of the generation of grammatical elements of L, such as articles, adpositions and auxiliaries. The result is a set of separate phonological strings ("words").
Stage II: linearization. After all forms have been computed, the strings of Stage I will be ordered according to the word order rules of L. This could be seen as the syntactic stage, though notions like "constituent" and "structure" do not seem to play a role whatsoever.
Stage III: phonological realization. To the ordered strings of stage II the set of phonological rules will be applied that make a well-formed linguistic expression in L out of them. This includes "sandhi" (i.e. the phonotactic processes at the word boundaries) and the application of sentential stress patterns and prosodic contour.

The rules for Stage I take the form of one or more µ (for morpho-syntactic) operators, to be applied to an operand O, leading to some value V, provided that a certain condition C is met. (1) below gives the general format of Stage I rules; condition C may be empty:[3]

[3] In fact, there were no conditions in the 1989 version of the expression rules. The + sign in (1) indicates that there may be one or more µ operators.

(1) μ+ [O] = V, IF C

μ operators may be of two types. The first are the primary operators. These are direct reflections of the π and ω operators and the respective functions in the corresponding UR. The second type are the auxiliary operators. These are introduced by the expression rules themselves. Examples are the cases, such as Ablative and Genitive, and verbal categories, such as Infinitive, Past Participle and Subjunctive. Typically, auxiliary operators have no direct and unique semantic interpretation themselves; they are formal categories which are the projection of one or more underlying semantic or pragmatic distinctions. In an earlier version of the theory there was a third type of μ operator: the contextual operators. These were mainly introduced to take care of agreement (cf. Dik 1989: 300f). In the version in Dik (1997a: 353f), they have been replaced by a second type of auxiliary operators. These are copies of primary operators, added to the UR after its completion. This is done by the expression rules, before expression proper starts. So, in the UR given in (2a) below, an expression rule adds those features of the subject which are relevant for agreement, here Number and Person, to the main predicate, giving (2b) as a result (example adapted from Dik 1997a: 351):[4]

(2) a. Pres e_i: talk [V] (d1 x_i: man [N])$_{AgSubj}$
b. Pres <1,p3> e_i: talk [V] (d1 x_i: man [N])$_{AgSubj}$

Such auxiliary operators are marked in URs by angled brackets. Apparently, also inherent features, such as animacy, gender and person, may be copied.[5]

In Stage I rules, as formalized in (1) above, there may be just one μ operator or there may be more than one, as in the case of portemanteau morphemes. E.g. the rule that computes the final form of the tensed verb in the expression of (2b) will have the following constellation of μ operators:

(3) Pres <1,p3>

As for the operand in rules of type (1), this is typically a predicate stemming from the UR. So, the rules that would give us the right form of the main verbal

[4] Here, and below, I will use the term “feature” for any type of function or operator, primary or auxiliary, both in underlying representations and in expression rules.

[5] As an alternative to this copying operation, Dik (1997a: 356) suggests contextual retrieval, i.e. the addition of the relevant features precisely at the moment when they are necessary for the calculation of the right form during expression. In the case of example (2a), number and person of the subject term would then be retrieved when the form of the tensed verb is to be calculated. The argument given for the preference of the copying operation is that “in the computational implementation as developed in ProfGlot (Dik 1992) it turned out that the copying method was easier to implement than the contextual retrieval method.” This seems to be a technical rather than a linguistic argument.

predicate of a clause would look like (4) below. Here, PRED is a variable ranging over the predicates in the lexicon.

(4) Pres <1,p3> [PRED [V]]

However, the output form cannot be calculated in a straightforward way since there are conditions as to its precise form. The English suffix for 3rd person singular present tense is *-iz*, *-s* or *-z*, depending on the last phoneme of the verb being a sibilant, a voiceless consonant or otherwise. A more or less complete version of this rule could look like (5) below (I have refrained from formalizing the condition part; this can not be done without a formalized morphophonemic subtheory):

(5) Pres <1,p3> [PRED [V]] =
[PRED-iz], IF (last phoneme of PRED is sibilant) ELSE
[PRED -s], IF (last phoneme of PRED is voiceless) ELSE
[PRED -z]

For the PRED part, the actual UR element will be substituted before rule application. Then the conditions can be applied. In the case of (2), this would result in (6):

(6) Pres <1,p3> [talk [V]] = *talk-s*

In the representations above and below, the italicized elements represent terminal forms, i.e. phonological strings that may not be operated upon anymore by some later Stage I rule, but that are ready for linearization. All other forms may be taken as an input by later rules, and changed in some way or other. In fact, apart from deletion, any transformation may be performed on an operand (cf. Dik 1997a: 352). However, empty operands are not acceptable. One of the possible transformations is the introduction of auxiliary operators. An example of this is found in (7) (adapted from Dik 1997a: 357):[6]

(7) Pres <1,p3> e_i: Perf talk [V] (d1 x_i: man [N])$_{AgSubj}$

One of the rules this constellation calls for is the application of the relevant operators on the main predicate. This is done in (8) below. Notice that first only the

[6] Note that, in (7), Perfect is coded as an aspect rather than a tense operator; Dik (1997a: 357) treats it as a tense operator. See the discussion in Comrie (1976: 52f) and Comrie (1985: 32f) on the English Perfect. In what follows I will leave out details on the ordering of rules and the communication between them, and fill in details where the theory does not explicitly say anything about it. This open endedness will precisely be one of the major problems to be discussed in the next section.

Perf operator is applied (8b), which introduces the auxiliary operator PaP (for Past Participle). PaP is applied recursively, in (8d), before the rest of the operators may do their job in (8e), leading to a string that consists of terminal forms only.

(8) a. Pres <1,p3> Perf [talk [V]] =>
b. Perf [talk [V]] = [have [V] PaP [talk [V]]]
c. Pres <1,p3> [have [V] PaP [talk [V]]] =>
d. PaP [talk [V]] = *talk-ed*
e. Pres <1,p3> [have [V] *talk-ed*] = *has talk-ed*

An important principle in ER is that of so-called lexical priority. It has been applied in (8e), where a deviating form in a lexical paradigm is inserted instead of the form that would result from the application of a rule, as in (5) and (6).

A last feature of the Stage I rules I want to mention here is the relative ordering of the terminal forms. In principle, ordering at the template level is left to the Stage II rules. However, examples in the literature suggest that the local ordering of Stage I grammatical material, such as affixes, articles and auxiliaries vis à vis their lexical heads, i.e. nouns, adjectives and verbs, is determined by the corresponding Stage I rules themselves. These orderings are all based on one principle: the assumed iconicity between scope relations among operators in the UR on the one hand and the order of the elements resulting from the application of these operators on the other hand. More specifically, the operators are supposed to have an inside out way of expansion, resulting in an inside out ordering of the resulting forms. Interestingly, as shown in the literature, there are quite a few examples in morphosyntax where this centripetal expansion leads to grammatically right orders. The order of the auxiliaries resulting from the expansion of passive and several μ operators in English is a rather convincing example, as shown in (9) below (for the complete derivation see Dik 1997a: 383).

(9) Pres <1,p3> Perf Progr Pass [call [V]] (John)$_{Ag}$ (Bill)$_{GoSubj}$
=>
(Bill) has been be-ing call-ed (by John)

However, it remains an open question whether this sole principle could explain the vast majority of ordering instances of grammatical material in a typologically convincing way, and whether the counterexamples could be seen as language specific idiosyncrasies, leaving the principle as such unaffected. In fact, it is not hard to find counterexamples, as we will see in the next section.

On the assumption that all grammatical material has been generated and ordered relative to the corresponding lexical material in Stage I, these ordered chunks are now brought into line by the linearization rules of Stage II. This stage has two types of instruments: templates and placement rules. Templates represent the relative order of the functional positions at several levels of syntactic description, typically the sentence and the noun phrase. Example (10) gives a (simplified) template for the English main clause; I leave out extraclausal constituents such as Theme and Tail.

(10) P1, Subject, Vfin, Vinfin, Object, X

The slot labels are functional rather than formal. They serve to identify the UR constituent that should land there. Taken in its most straightforward sense, this implies that there is a one-to-one relationship between UR elements and template slots and that this relationship is still identifiable after the operations of Stage I. A less straightforward interpretation would open up the possibility that the Stage I rules alter existing functional categories and create new ones. As far as I can see, the latter position is never taken in the literature. Now, in none of the world's languages can there exist a strict correspondence between UR elements and any template for two reasons. Firstly, elements may optionally be absent from a specific sentence, especially when we take into consideration holophrastic sentences as discussed in Mackenzie (1998). Secondly, probably no language in the world has completely fixed orders at all levels of syntactic description.[7] This means that, apart from rigid templates, which give the basic positions, we need placement rules to specify the alternative orderings. The typical example of such a rule is the one going with the P1 position of template (10), that (obligatorily) hosts the first constituent of the sentence. Somewhat informally stated, this placement rule may take the following shape:

(11) Q-word to P1, IF NOT PRESENT THEN
Focal constituent to P1, IF NOT PRESENT THEN
Topic to P1, IF NOT PRESENT THEN
Subject to P1

The following should be observed about placement rules, among other things. Alternatives are put in the order of preference. UR elements may be referenced to in a multiple fashion (e.g. some element may be both Topic and Subject). And reference is made to a category, Subject, that has its own slot in template (10). The latter implies that, optionally, the functional position labeled "Subject" will remain empty. An alternative to placement rules is having several templates for the same level of analysis rather than one. E.g. for the Dutch sentence, a separate template has been suggested for subordinate as opposed to main clauses. A crucial point is that such elements be uniquely identifiable in URs on the basis of their function.

This completes our brief survey of the expression rules. In the next section we will have a critical look at them, and see whether they would work for some more or less complicated cases of morphosyntax.

[7] This is at least the case for the languages of Europe. For the 12 "Greenbergian" constituent pairs discussed in Bakker (1997) even the most inflexible languages, i.e. those belonging to the Altaic phylum, have some constituent order variation, while some languages, notably Latin and some Finno-Ugric languages, are almost completely flexible. Flexibility runs from 8% of the languages for the most conservative parameter to 73% for the most flexible one.

3. SOME FUNDAMENTAL PROBLEMS WITH THE STANDARD MODEL

First, let us look at the expression rules in a general way, restricting ourselves to Stage I and II. A point that may strike us is that, in comparison with the rules that generate URs, we seem to know much less about constraints on expression rules. Nothing has been established in FG theory that even remotely looks like URC1-URC5 in section 2 above. In fact, it seems that virtually anything is possible in expression. The only constraints I can see for Stage I is that these rules may not delete material, that operands should be identifiable elements of URs and that there may be no empty operands. The order of generated grammatical material is either iconic with underlying representations or it is idiosyncratic. Templates are constrained by the fact that the functional positions that they define should be the basic positions, and that deviating (marked) orders should be defined via placement rules.[8] Apart from these constraints, from those on URs, the potential constraints determined by Stage III, maybe some general optimality criteria, and the obvious fact that this all should lead to precisely the set of well-formed expressions of the language concerned, there seem to be no limits to:

a. the copying of primary operators in URs
b. the combination of μ operators
c. what could be an operand
d. the operations that may be performed on operands, and the resulting values
e. the type of conditions there may be, both on input and output
f. the order in which these rules should be applied
g. the shape and number of templates
h. the type and number of template slots
i. the relation between the respective templates
j. the way the slots are filled
k. the nature of placement rules

Even if we exclude deletion, the fact that there would be no constraints in all these areas would mean that we are dealing with an extremely powerful rule system here. This has crucial implications for its acquisition. We can, of course, interpret the points above as a research programme with regard to the further development of the ER component, taking the requirements of typological, pragmatic, cognitive and possibly also diachronic adequacy into consideration.[9] However, we should be con-

---

[8] The definitions in Dik (1997a: 358) seem to be suggestive of the fact that primary operators may not be recycled, i.e. applied more than once. However, they are applied without any restriction via the copying operations on the UR.

[9] Cognitive replaces psycholinguistic adequacy here. In Bakker (1998) I plead for diachronic adequacy as a fourth criterion. I think that this is especially called for with respect to the expression rules, since language change quite frequently involves change in form.

vinced, then, that the overall architecture of the ER component as sketched above can handle all kinds of expressions that we may come across in the languages of the world while by the same token it should not overgenerate. The following examples show that the current expression rules fail to produce several types of structures that occur quite frequently in the languages of the world.

The first two examples concern problems that are created by relying just on iconicity to get the right order for the forms that are generated by a complex set of operators. Both may be illustrated by the generation of auxiliary verbs. Above, in section 2 we have seen that these are produced by centripetal application of the μ operators on the verbal predicate. For English, this results in verbal strings like the one in (9) above, that may be inserted in a template slot in a more or less straightforward way. However, for Dutch this does not seem to work. Compare:

(12) Hij heeft dat horloge misschien willen stelen.
He have:3SG that watch perhaps want: INF steal: INF
'Perhaps he wanted to steal that watch'.

As in the English example of (10), the Dutch template for the main clause has a separate slot for the finite verbal element, which will be marked for Tense and Subject agreement (Vfin). In the case of (12) the tense expressing auxiliary *heeft* 'has' is the inflected form. The other two verb forms, the infinitives *willen* 'to want' and *stelen* 'to steal', are inserted towards the end of the template. This means that, unlike the English example in (9), these three verb forms cannot be generated as a well-ordered cluster: we must assume that the terminal forms are still marked for their respective functional categories, as to make them selectable for the ordering rules of Stage II. This would, in fact, be a movement transformation. If, on the other hand, we would want to generate the tensed form separately, in order to be in the position to process it independently, then we would have to admit an empty operand in this case.

A second problem arises when languages divert from the centrifugal scheme. English yes-no questions is a case in point. If we assume that, for the sentence in (13a) below we have an UR that contains an Interrogative operator at the illocutionary level, we will, at a certain stage, end up with the rule in (13b):

(13) a. Did your brother marry the girl next door?
b. Int Past <3 Sg> [marry [V]]

If we work out the sequence of operators in (13b) in an inside out fashion, we would end up generating the ungrammatical string in (14):

(14) *do married

So, the separate generation of the tensed auxiliary seems to be called for also in this case.

More in general, any type of discontinuity creates a problem for the standard ER model. This is the case above all for languages that are traditionally called non-configurational, such as many of the Australian languages. The example from Djaru in (15) may serve to illustrate this.

(15)	Jalu-ngu	lani-i	mawun-du	dadi	jambi-gu.
	that-ERG	spear-PAST	man-ERG	kangaroo	big-ERG
	'That big man speared a kangaroo'.				

If we analyze (15) such that, in its underlying representation, the three ergative elements are united in one term, as operator, nominal head and adjectival restrictor, respectively, then the placement rules will have the complicated task to disentangle the set of terminal forms that result from the expansion of this term. On the other hand, if we analyze the three elements as being in apposition, and represent them as independent entities of the UR on the basis of their detached expression, then we will need a major adaptation of the theory of well-formed underlying representations for such cases.[10]

That at least some sophistication is necessary with respect to the ordering of processes at Stage I may be shown by an example from Latin. In this language, as in others, auxiliary operators such as Case may not only be generated directly, e.g. triggered by a function in the UR, or by a verbal predication, but also indirectly, via markers that have themselves been generated by ER, such as adpositions. The latter I will call indirect auxiliary operators. So, while some cases may directly express a semantic function, like Genitive case for Possessive function, the Temporal function in (16) is expressed by the preposition *post* 'after' which, in its turn, introduces the Accusative.

(16) (sg x_1: memoria [N]: (pl x_2: homo [N])$_{Poss}$)$_{Temp}$

Post	hom-inum	memori-am.
After	people-PlGen	memory-SgAcc
'From time immemorial'.		

So, the "outer" function Temporal should be expanded first, giving the preposition *post*, with the corresponding Accusative Case operator, before any "inner" operators such as Number may be expanded.

Another problematic point, which is of a fundamental nature, are cases where there is a mutual interaction between form and order. The following examples, from Dutch (17), Breton (18) and Arabic (19) illustrate this.

(17)	a.	Jij	speel-t	vandaag	goed!
		you	play-2SG	today	well
	b.	Vandaag		speel	je goed!
		today	play-0	you	well
		'Today you play well!'			

[10] This would extend the notion "non-configurational" to the UR level. A much more in depth study of these word order phenomena is called for, of course, to decide on such matters.

(18)	a.	Ar	vugale	ne	lennont	ket	levrioù
		the	children	PCL	read:3PL	not	books
	b.	Ne	lenn	ket	ar	vugale	levrioù
		PCL	read:3SG	not	the	children	books
		'The children do not read books'.					
(19)	a.	'akawat-I		jalasna			
		sisters-my		sat:PL			
	b.	jalasat	'akawat-i				
		sat:SG	sisters-my				
		'My sisters sat'.					

In all cases, the agreement marking on the verb only appears when the subject has been expressed before the finite verb form appears. It is as if the features necessary for agreement are not fully accessible, or assumed to be relevant, when the corresponding source of agreement has not been processed yet by the expression component.[11] Interestingly, this is in line with the assumption made by von der Gabelentz (1901; as quoted by Plank 1998: 197) that noun-adjective agreement would be more likely to occur in languages with NAdj orders than those with AdjN orders. Note that, also in this case, the source of the agreement should precede the target.[12] Noun-adjective order in Spanish may also be a case in point. In (20), the adjective *bueno* 'good' gets the short form *buen* when it is in front of the noun it modifies, but not when it follows the noun.

(20)	a.	un	buen	hombre
		a	good	man
	b.	un	hombre	bueno
		a	man	good
		'a good man'		

Of course, from the perspective of FG only Gender agreement would be really crucial in this respect, since features like Number and Case originate from the term rather than the noun level, and are 'available' as soon as the term as a whole is processed for expression.

An example from the area of speech errors is highly suggestive of the fact that linguistic forms are not computed autonomously as in Stage I, independent of

[11] This is a very loose observation that will remain unsubstantiated here. On closer inspection there may well be different reasons for this phenomenon in different languages. E.g. the Breton example in (18) stems from Borsley and Stephens (1989: 411) and is discussed in depth from a syntactic perspective, in this case as an argument for Breton having VSO rather than SVO underlying structure.

[12] Greenberg's (1963: 95) universal number 40 points in the same direction, though it makes slightly different predictions: "When the adjective follows the noun, the adjective expresses all the inflectional categories of the noun."

their relative position in the syntactic structure. Consider the following examples from Dutch:

(21)	a.	een	klein	huis-je
		a	little	house-DIM
	b.	een	huis	klein-tje
		a	house	little-DIM
	c.	*een	huis	klein-je
		a	house	little-DIM

(21a) is the right form, with the diminutive suffix form phonologically adapted to the noun. The speech error in (21b) reverses the noun and the adjective. However, the Diminutive remains on the canonical noun position, but it is adapted to the form of the adjective. Errors like the one in (21c), where the suffix form is not adapted, and has the form it would have on the noun, are never observed. Apparently, the Diminutive operation is not applied to the noun in isolation, at an early stage in the production of the sentence, but retains a rather abstract status until the moment of expression. Also, it seems to be associated to the noun position rather than to the noun itself.[13]

The above examples are cases with respect to which the ER component undergenerates, i.e. it fails to produce grammatical forms (or to explain more or less systematically observed speech errors). These are empirical facts, and they need to be repaired in order for the FG model to be observationally adequate. However, given that there are hardly any constraints on most of its subprocedures, it is quite likely that ER will as well overgenerate, i.e. that it is too powerful and that it potentially produces forms that would never occur in any of the worlds' languages. This point is of an entirely different nature, given the rather dependent position of morphosyntax in FG theory, and in functionalist, as opposed to formalist approaches to grammar in general. It could be claimed that morphosyntactic overgeneration is of no real empirical interest since superfluous forms will never crop up in the first place as long as no underlying representations trigger them. They are just potentially there. There are reasons, however, to pursue this matter anyway. Even though the ER component may be highly subordinate and subservient to the functional components, and (diachronically) shaped by the functional component for that matter, it may well turn out that it is not completely arbitrary, and that it should be granted some autonomy, in the sense of Croft (1995) and (forthcoming). Motivation for this may be derived from the fact that, in the Greenbergian tradition in

[13] This example would also suggest that diminutive formation in Dutch is a term (or predicate) operator expanded by the ER rather than a predicate formation rule preceding UR formation, unless we assume that the final calculation of forms resulting from predicate formation is postponed until expression of the clause as well.

language typology, a large number of formal universals have been proposed that go counter to the idea that in the realm of expression "anything goes" (cf. Greenberg 1963; Hawkins 1983; Dryer 1992; Bakker 1997). The Chomskyan tradition in linguistics even takes the autonomy of form as a point of departure, by assuming syntactic structures to be innate in some more or less abstract way (Chomsky 1965; 1995). Of course, within the functional paradigm, proposals have been made to explain a great many of the formal patterns on the basis of higher, generally cognitive, categories and principles (cf. Bakker and Siewierska 1992 on word order explanation). Within FG proper, the Principle of Functional Explanation (Dik 1986) encourages the researcher to seek explanations as high as possible on the hierarchy of (22a) and the more detailed version of it in (22b) below:

(22) a. Social > Discourse > Functional > Formal Factors
b. Social >
Discourse >
Pragmatic > Semantic >
Syntactic > Morphological > Phonological Factors

However, this still leaves open the possibility that we are confronted with phenomena that can only be explained on the basis of formal assumptions, be they of a syntactic, morphological or phonological nature. This is evidently the case when we have to face expressions which have a fully interpretable underlying representation but which are ruled out on the basis of an unsolvable conflict between expression rules. Stated somewhat differently: in cases where the expression rules work as a filter. The Latin example in (23) below, taken from Dik (1997b: 208), is a case in point.

(23)	*Romani	in	et	ex	Asiam/Asia	transierunt.
	Romans	into	and	out-of	Asia-ACC/Asia-ABL	cross-over

'The Romans crossed over into and out of Asia'.

In (23) we find the coordination of two prepositions, *in* 'into' and *ex* 'out of', which share their prepositional object. The problem is that the former preposition governs the Accusative while the latter governs the Ablative Case. This leads to a conflict with respect to the form of their object, *Asiam* or *Asia*, respectively, which comes out only during expression.[14] But even though such "hard" examples of formal constraints, and autonomous operations by the ER component seem to be

[14] Of course, technically this conflict could be detected on the basis of the underlying representation, but this would imply that part of the expression rules have to be preprocessed. One wonders, however, whether such sentences would have been avoided or interrupted in their spoken form by a native speaker of Latin, and whether this is only a problem at the level of the written language. In this respect it is interesting to see that languages tend to solve such conflicts by selecting the neutral, or the least marked, or the nearest form. See for this Corbett (1991: 261f) who discusses

rather marginal, we would like to know what are the constraints on expression for their own sake, in other words: what are possible and impossible forms in language, since this is an aspect of a general theory of grammar as well.

Concluding this section, I make the following observations. The standard model of expression, which separates the construction of forms from their linearization falls short empirically: it gives rise to both undergeneration and overgeneration. The facts above suggest that we should look for a model of ER in which form and order of expression are interdependent, a model that directs the right functional and lexical information to the right place in a sufficiently abstract way, while postponing the computation of the eventual grammatical forms as long as is necessary and possible. Since left-to-right order of expression —in terms of actual speech production: relative moment of expression— seems to play a role, such a model should preferably be dynamic rather than static. In the next section I will propose a model for the expression rule component which attempts to meet these requirements.

4. THE EXPRESSION RULES REVISITED

For the construction of an alternative ER component, henceforth: the dynamic model, I start out from the original idea of templates with functional positions. However, instead of seeing them as more or less independent prescriptions for ordering phenomena at different levels of syntactic representation, I will combine them into a hierarchical structure. Although this is not ruled out as such by the theory, no explicit relationship has been postulated so far between the templates for the respective levels of description, at least not to my knowledge. Rather, they have been proposed as independent means for the linearization of the material contained in UR elements, such as clauses and terms. Here templates will be interpreted in a more interdependent way: as a structure. This structure of interdependent templates is not postulated as a static entity, but it comes into existence in a dynamic fashion on the basis of the UR material present.

Depending on the specific features of some underlying representation U_i that is about to be expressed —typically a clause— the right template T_i is selected. This template is labeled by a formal category, e.g. 'Sentence' in case a main clause is to be expressed. This is the top category of the expression. The template has slots for all elements that may be contained in U_i, and that have a function directly related to U_i, much like standard templates. However, the slots available may also be dynamically tuned to the concrete contents of U_i. In the standard model templates are maximally specified out of necessity, which typically leads to a certain amount of empty slots. In

---

the resolution of conflicts in gender agreement under coordination. And probably all languages with subject or object agreement are bound to run into this kind of problem under coordination. Cf. example (i):

(i) ?Either his parents or his sister have/has helped him financially.

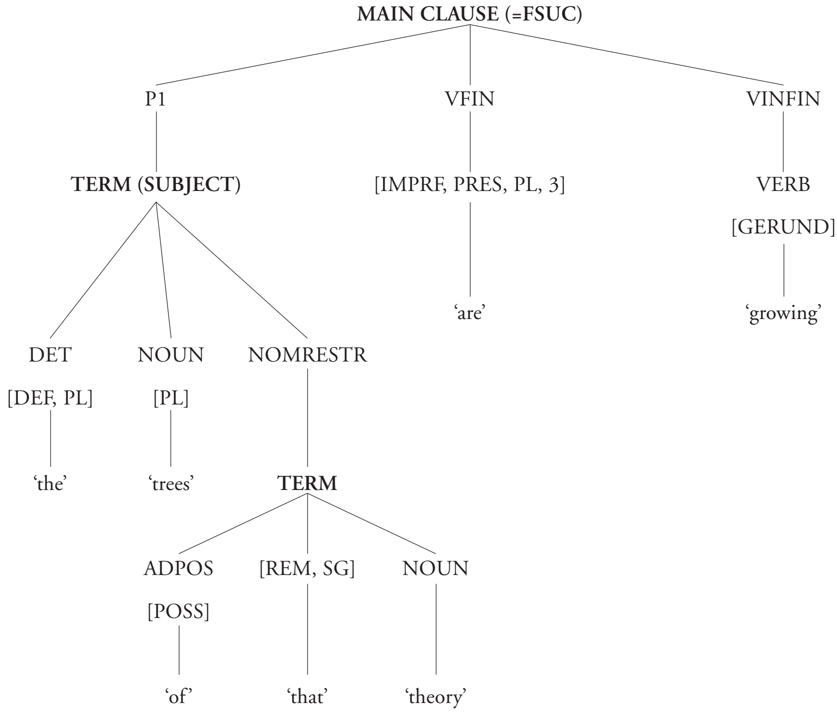

Figure 1. Constituent structure tree of nested templates

the dynamic model, the U_i material is directed to the corresponding slots in a left-to-right order. Depending on its qualifications, element $U_{i,j}$ of U_i in slot $T_{i,k}$ of template T_i in its turn selects a template fit for its own expression. So, if $U_{i,j}$ happens to be a term, then it might select term template T_j with category 'Noun Phrase'. If term $U_{i,j}$ has a nominal head, the latter will be directed to the Nominal Head position of T_j, together with the relevant operators that determine its final shape, such as Number and Case. Auxiliary operators are created in the process. So, in this example, the constellation of μ operators (i.e. the term operators and functions) of term $U_{i,j}$ will generate the right Case after term template T_j has been invoked, adding it to the ω operators of the term. This value will then be available by inheritance for all templates that will sprout from T_j, and for which it is relevant. In this top down process, lexical material is directed to the relative location of expression. At the same time, auxiliary operators are created at the right moment, and grammatical elements, such as affixes, adpositions and auxiliaries, are inserted at the right position on the basis of the locally available and relevant set of μ operators, which have been directed to that position. This recursive process terminates when, on all paths down from the top category, all μ operators have had their effect, and only separate lexical elements are available for insertion into template slots. The result is a tree-like structure such as the one depicted in figure 1 in a simplified form.

This is a first and rough impression of the dynamic ER model. We start out with purely functional information, i.e. the underlying representation that is to be expressed. As we go down the tree while selecting the right at the respective levels, semantic elements will be distributed in an increasingly more fine-grained fashion over the template slots, and will gradually mix up with grammatical elements, such as syntactic constituents, auxiliary operators, articles, adpositions and pronouns. Still further down the tree morphosyntactic and purely morphological elements will start to appear, such as particles, clitics and affixes. In a complete model, this could be extended to Stage III, i.e. morphophonology and the purely phonological level. The overall idea is that semantics, syntax, morphology and phonology merge at their boundaries, while blurring these boundaries much in the way in which they are blurred in diachronic processes.

This sketch will now be refined in three ways. Firstly, we will have to say more about the dynamic aspects, i.e. how structures as exemplified in figure 1 come about. Next, we will discuss the notion of templates in the dynamic model, in other words: what is the status of the nodes in these trees and what is the nature of branching in such structures. Thirdly, we will have a closer look at the types of information that may be available at each template slot, or, rather, each node of the tree.

4.1. The dynamics of tree construction

Five formal principles direct the construction of the constituent structure tree, of which three have already been introduced above. Although I will give short motivations for each of them, the main reason for introducing them is that they give a rationale for certain order phenomena, and that they potentially put constraints to the information flow in the morphosyntactic structures that will be developed. It is then an empirical matter whether they might turn out to be the right constraints, whether we will have to relax them, or whether we will have to rehash the overall organization in the sense that we assign elements and features to the UR level that would formerly be assigned to ER and the other way around. There may also be implications for the status of certain lexical forms, i.e. whether they are mere grammatical elements or should be considered as full predicates, with semantics of their own.[15]

As a point of departure for the design of a dynamic ER component I propose the following 5 principles to hold for the constituent structures that will be produced by it.

ERP1: Constituent structures are developed *top down*. In this way, and other than in the standard model, the generation of linguistic form follows (the linguistic part of) the explanatory hierarchy given in (22b) above.

[15] Compare the discussion on the status of English prepositions in Mackenzie (1992).

ERP2: Development takes place from *left to right*. This is the "natural" order in which linguistic forms are uttered in the first place. Left to right ordering may be expected to shape and have shaped language over time, at least to some extent.

ERP3: Development works *depth first*. This implies that of any two contiguous slots in a template, the leftmost one will be completely expanded up to its terminal forms before the rightmost one will be considered, and this recursively. This means that of the complete information that is necessary for the production of the whole utterance, including copies of UR material, only a fraction will be available at any one time. This considerably reduces the burden for short term memory, inherent to breadth first development. In combination with ERP2, depth first processing may help explain certain order phenomena, such as the ones illustrated in examples (17)-(19) above.

ERP4: For any node N in the tree, all features found on a direct path from that node to the top node are in principle available to be *inherited* by N. Trivially, this implies that overt primary operators would be available for every node in the tree. However, a distinction will be made between "raw" UR material (functions, π and ω operators, inherent features of predicates such as Animacy or Gender) on the one hand, and μ operators on the other hand. In order to be accessible for inheritance, I will assume that any UR feature has to be transformed into a μ operator, either in a one-to-one or a more-to-one fashion, as in the case of portemanteau operators. An implication of this is that only those features of URs qualify as such if they are turned into a μ operator at some stage during expression. Apart from this there may be universal or language (type) specific *downward barriers* that make features inaccessible to lower nodes. Barriers may be both of a functional and a formal nature, i.e. determined by a layer or substructure in UR or by some syntactic or morphological boundary in ER.

ERP5: Operators may *percolate*, i.e. move upwards to higher nodes. Also for percolation there may be universal and language (type) specific *upward barriers*, that make them inaccessible to higher nodes. Barriers to percolation are only of a formal nature.

These principles should contribute to the (cognitive and typological) adequacy of the dynamic model. Externally, this organization means that the linguistic forms are produced precisely in the order in which they are uttered by the speaker (and reach the hearer) in a life setting, thus giving the model a procedural flavor. Furthermore, several constraints follow from these principles which restrict the formal power of the expression component, and determine what information necessarily has to come in from the UR, including its level of representation, or from an earlier stage of expression. I will mention the more obvious constraints here, and leave others for the discussion of the concrete example in section 5 below.

Given the notions of inheritance and percolation introduced under ERP4 and ERP5 above, top down development makes clear at which depth in the morphosyntactic tree μ operators, both primary and auxiliary, should be available for agreement purposes. These mechanisms replace the pre-expressional copy operations on the UR in the standard model. In the latter model, it is not always clear how operators can be copied to other locations of the UR in the first place. Copying seems to be complicated for inherent features such as Gender, and virtually impossible for indirect auxiliary features, since the latter are triggered by forms that themselves are generated during expression, such as Case assigned by an adposition (see example (16) above).[16]

Left to right combined with depth first development of constituents puts constraints on the material that is available in a "vertical" sense. This is especially relevant for inherent features and, even more so, for auxiliary operators, direct as well as indirect. For instance, in order for Gender information to be available for agreement purposes on nodes to the left of its source term, the nominal head in question will have to be inspected at a relatively "early" moment of expression, i.e. high in the tree, often before the term of which that nominal head is part is processed for expression itself. Indirect auxiliary operators can only become available to other nodes by percolation, and are therefore inaccessible for nodes to the left, since these have already been expressed.

So, by postulating a dynamic development for our formal framework, we may make several predictions as to the relative frequency, or possibility, of certain phenomena in the languages of the world. On the basis of this, I tentatively propose the following agreement hierarchy, for which typological relevance is predicted here:

(24) *Agreement Hierarchy*

Overt primary operator	(Number, Person)	>
Inherent primary operator	(Gender, Animacy)	>
Direct auxiliary operator	(Case via function)	>
Indirect auxiliary operator	(Case via grammatical marker)	

So, I expect there to be more languages that have agreement on Number than on Gender. Also, if a language has agreement on Gender, it is likely to also have agreement on Number. The same may be expected to hold for e.g. agreement on Number and Gender on the one hand and Case on the other hand. Finally, although the direction of agreement will be completely free for Overt primary operators, and relatively free for Inherent ones, Case agreement left of the source will be relatively rare if it is direct, i.e. derived from some function in the UR, and is even ruled out when it is indirect, e.g. when derived from an adposition.[17]

[16] The only solution also in these cases would be to preprocess part of the expression rules.

[17] The first prediction is reflected by Greenberg's (1963) universals 32 and 36. U32 predicts that when there is subject and object agreement on the verb in Gender, there will also be Number agreement. U36 paradigmatically predicts Number distinctions whenever there are Gender distinctions. My second prediction may get some support from universal 39. This universal holds

Of course, these predictions are all very tentative. They serve as a first attempt to restrict the formal power of the expression component, and they have to be verified on the basis of a representative language sample. However, they are clearly more open to empirical test than the standard model, and an improvement at least in this respect.[18]

4.2. Branching and the nature of nodes

In this section we will have a closer look at the global organization of the tree structures dynamically created by the expression component. The nodes from which the tree is constructed originate from slots in functional templates. Just like slots they represent the relative order of UR elements in expressions since no movement is allowed. Although no complete set of criteria has been developed within the standard model for what types of elements may be united in one specific template, I will assume that the UR elements which are linearized by the same template will stem from one specific functional level. In other words: they should have a direct functional relationship to one and the same higher category, i.e. the category of the node that is expressed by the template. In line with this, as a general point of departure I will assume that there exists some type of iconicity between UR and ER in terms of direct constituency and of scope. Thus, branching in trees will be restricted as far as possible to acknowledged hierarchies in underlying representations, i.e. to what in the standard theory would lead to the postulation of a specific template. This means that, for the overall processing of structures such as (15) above, we should first consider a UR solution before resorting to discontinuity. However, I will not rule out altogether that there may be compelling independent syntactic, morphological or phonological arguments for introducing syntactic constituency that is not completely iconically motivated. In principle, this leaves us with a rather "flat" tree structure, where the number of daughter nodes is, in fact, unrestricted and where there are both semantic and formal motivations and restrictions on the introduction of subconstituents.[19]

---

that, where both Number and Case of the subject are reflected on the noun, and the affixes are on the same side of the stem, the Case marker is almost always on the outer side. This may be interpreted as diachronic evidence for the primacy of Number.

[18] For falsification, at least as important a procedure for testing the acceptability of proposals like this one, negative predictions have to be formulated to the extent that certain phenomena are predicted not to occur, or with extremely low frequencies. For research into relative frequencies, rather large samples are required, say several hundreds of languages. See Rijkhoff and Bakker (1998) for a method that helps to establish a representative sample of the languages of the world.

[19] Such flat tree structures are in stark contrast with the strictly binary character of trees in the more recent versions of Generative Grammar (but see the discussion in Culicover 1997: 162f). In syntactic theories, specific subtheories and principles such as c-command and trace theory help

A second point is that the original templates, and by implication the constituent trees that are derived from them, are maximally specified in the sense that there is a slot for any functional element that may potentially be expressed by it. This means that at least part of them may have no filler —i.e. functional material that is expressed via that node— in actual cases, since the corresponding UR material is simply not present. The well-known X slots for satellites in sentence templates are a case in point. In the dynamic model, such optional nodes are dealt with in one of two ways. Either they are left out of the resulting structure altogether —they are deleted, or possibly: they are not generated in the first place— or they remain "empty" but nevertheless crucially present in the final structure. The assumption here will be that both types of nodes are in fact needed: certain syntactic and morphological phenomena may be explained precisely if we assume that the corresponding constellations are sensitive to nodes that are structurally present but not filled in a concrete case of expression. An example of such a "trace" phenomenon is soft mutation in Welsh and other Celtic languages (cf. Ball and Müller 1992).

A third, and last point I want to make here concerns slots that may have multiple fillers. We can think of series of adjectival term restrictors. Two solutions present themselves in this case. Either we linearize the respective elements, an operation which is indicated with an asterisk in string algebra, as in example (25) below. Or we embed the elements, an operation which is performed by a recursive rule in formal syntax, as in example (26). Note that the grammars of (25) and (26) are weakly equivalent, i.e. they generate the same language fragment but assign different structures to it: (25) produces strings of coordinated adjectives; (26) produces a right branching tree with embedded adjectives.

(25) Det Adj* Noun
(26) a. NP -> Det Noun
b. NP -> Det AP Noun
c. AP -> Adj
d. AP -> Adj AP

In the days of early transformational grammar there were strong objections to the intuitively unnatural embeddings that grammars like (26) tend to generate, especially for those cases where the semantic analysis would have them coordinated rather than in the scope of each other. I will take the position that such syntactic

---

decide on branching and other constituency matters, and thus solve certain problems such as anaphor binding. In functional theories such tree structuring principles are not part of the basic toolkit. The shape of structures is preferably iconically based on structures from the functional level. However, there may be independent reasons to introduce syntactic structure, e.g. the fact that adpositions appear to be the syntactic heads of adpositional phrases while they do not appear in URs and therefore can not be semantic heads of terms.

embeddings are only acceptable if they serve an independent purpose, i.e. not only should they reflect the semantic scope relations iconically but they should also get formal expression in some way or other, e.g. in terms of stress patterns or prosody, in order to distinguish them from the corresponding non-embedded structures. In that case distinct representations such as (25) and (26) provide a syntactic interface between function and (phonological) form. Arguably, this distinction should be made in the case of (27) and (28) below. (27), with coordinated adjectives, could get a flat representation; (28), with stacked restrictors, and contrastive focus on the outer one, leading to a specific stress pattern, could get an embedded structure that reflects the scope relations.

(27) The big, bad wolf
UR: (def sg x_1 : wolf [N] : (big [A] & bad [A]))
ER: [$[\text{the}]_{Det}$ [$[\text{big}]_A$ $[\text{bad}]_A$ $]_{AP}$ $[\text{wolf}]_N$ $]_{NP}$
(28) A bad big wolf
UR: (indef sg x_1 : wolf [N] : big [A] : bad $[\text{A}]_{FOCUS}$)
ER: [$[\text{a}]_{Det}$ [$[\text{bad}]_A$ [$[\text{big}]_A$] $]_{AP}$ $[\text{wolf}]_N$ $]_{NP}$

4.3. What is in a node?

In traditional phrase structure trees, which are typically presented as static, after the fact constructions without a 'history', the information represented per node is often restricted to a category label and, possibly, a small number of features, relevant for that node. Only at the terminal nodes, which represent lexical items, are more extended clusters of semantic, syntactic, morphological and phonological information to be found. The information flow as such is generally not made explicit.[20] However, the trees created by the modified expression rules do bear witness of their coming into existence by way of a complete set of features on all nodes of the resulting structure.

The first distinction that will be made here is between functional and formal information. Functional information originates exclusively from the UR level. Formal information is developed during tree expansion. Figure 2 below gives a

[20] See, for instance, Head-Driven Phrase Structure Grammar (Pollard and Sag 1994). In the respective versions of Generative Grammar part of the transformational history of phrase structure trees may, of course, be reconstructed on the basis of traces. Furthermore, there are a number of subtheories and principles which control the well-formedness of phrase structures in a more static fashion. Dynamism is part of the trees in Incremental Procedural Grammar (Kempen and Hoenkamp 1987). In FG, underlying representations of utterances are also presented in an ahistoric way, the well-formedness criteria being applied on hindsight rather than dynamically, as an integrated aspect of UR construction.

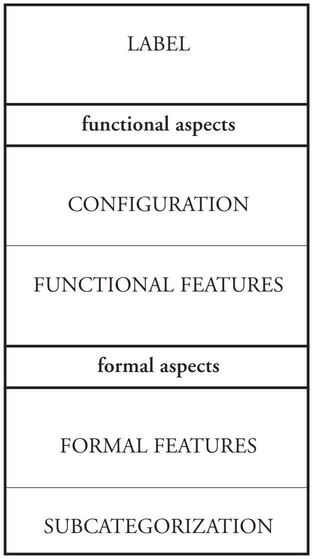

Figure 2. Structure of a node

complete picture of the types of information available per node, distributed over five fields.

Each node has a label, functional aspects and formal aspects. These are defined as follows.

a. LABEL (abbreviation: Lab): an indication much like the functional labels of template slots.
Examples: P1, Subject, Vfin.

The functional part of a node consists of the following types of information:

b. CONFIGURATION (Config): a specification describing that part of the underlying representation in the CONFIGURATION of the mother node of the current node that will be expressed via the current node. The top node of the expression has the complete UR for its Config.
Examples: the specification of the subject term, a temporal satellite, the head of a term, the main verbal predicate.
Provided that they belong to the same expression category, such as a noun phrase or an adjectival complex, Config may consist of an ordered set of alternatives for the expression of the node concerned.
Config may also contain the description of a purely grammatical entity that is the expression of a set of operators.
Examples: an auxiliary, a demonstrative.

c. Functional features (FncFtrs): primary μ operators relevant for this slot. These may be derived from operators, functions and lexical elements of the Config, or they are directly inherited from the FncFtrs field of the mother node. Examples: Number, Tense, Animacy.

The formal part of a node consists of the following fields:

d. Formal Features (FrmFtrs): auxiliary μ operators relevant for this slot. They may be derived from the Config and the FncFtrs, be inherited from the mother node, or percolate upwards from a daughter node.
Examples: Case, Finiteness.

e. Subcategorization (SubCat): template consisting of nodes for the functional categories into which Config is to be split up. This template is selected on the basis of the functional part of the node.
Example: for a standard template see (10) above. A specific template may be inserted on the basis of lexical information (an instance of lexical priority), typically from the head, for cases where the default node specifications would lead to the selection of the wrong syntactic environment.

In fact, the above characterization of a node is a maximum one: Lab is the only fixed element. For the rest, nodes may be found in four states: uninstantiated, partially instantiated, semi-instantiated and fully instantiated. In the uninstantiated state, they are theoretical constructs in the grammar, which are waiting, as it were, in a minimally prespecified form, to be used in the expansion of a concrete UR. In this form, in which they represent static grammatical knowledge, Config will be specified to the extent that it is necessary to select the right type or types of substructures of URs in concrete instances of expression. Both FncFtrs and FrmFtrs are underspecified in the sense that the relevant features are given (e.g. Tense, Number, Gender), but the corresponding values (e.g. past, dual, neuter) are left open. These values will eventually be determined on the basis of the Config information available once a concrete UR has been selected to be expressed. Subcat is left open, also to be selected on the basis of a specified Config.

Tree expansion starts with the specific UR_i that is about to be expressed. On the basis of its characteristics, a node is selected from the grammar of which one of the descriptors in the Config fits UR_i. Typically, this will be a node with a label like 'Sentence'. However, since I assume a FG-like grammar to reflect full conversational capacity rather than some abstract notion of 'grammatical well-formedness' it may also be a node that expresses a single term, or any other category that could represent an utterance. This will be the top node of the expression. In the case of a fit UR_i will replace the contents of Config of the top node. At the same time, the unspecified features for FncFtrs and FrmFtrs of the node are assigned values on the basis of the UR in Config, i.e. they get the values of the corresponding features in UR_i. At this stage we have a partially instantiated node. Finally, on the basis of Config the right subcategorizing template will be selected, with its own uninstantiated slots. The latter are the (potential) daughter nodes of the top node. This gives us a semi-instantiated node. Now, going from left to right, for any node of SubCat elements will be selected from the Config that fulfill the configurational specifications of the respective daughter nodes. This is done in the order in which possible alternatives are specified in the Config of the daughter node, and to the extent that the relevant information is available in the Config of the mother node. If a filler is indeed found, then it replaces the Config information of the uninstantiated daughter node. Its feature fields are specified on the basis of its own Config. Features that remain unspecified after the Config has been filled in are inherited from the respective fields of the mother node. They have to be explicitly specified there at an earlier stage. Thus, in principle, local values get priority over global ones. In this way, nodes are recursively created and instantiated. After their completion nodes are semi-instantiated. Values which remain open also after the Config has been replaced by a concrete UR element are supposed to percolate from lower nodes at a later stage of the tree expansion process. As a result, nodes may only be fully instantiated when all terminal nodes that are dependent on them have been created. On the other hand, no node may be semi-instantiated after all its dependents have been established.[21]

Every downward expansion of the tree will end in terminal nodes. A terminal node is one which contains only lexical material in its Config field. Such a node will get no further expansion, but will be handed to the phonological component for expression. In the first place the lexical material of terminal nodes stems from the UR level, in which case we are dealing with some type of predicate. However, lexical material may also be inserted during expression. For this purpose, there are nodes which have a lexical specification for their Config rather than a specification that selects a constituent in the Config of the mother node. Such a lexical specification, typically a grammatical category such as Aux or Adpos plus a set of features, will be interpreted as an instruction to select an appropriate element from the lexicon. It will be assumed that, within the lexicon, there is a separate subset of 'grammaticalized' entries, which are not accessible to the mechanisms that construct underlying representations, only to the expression rules.[22]

[21] One could think of a default mechanism that fills in values for features that remain uninstantiated after the top down creation of a node. This and other technical and procedural details will not be discussed here.

[22] This state of affairs has some evident implications for the discussion around the lexical or grammatical status of elements such as adpositions. (cf. Mackenzie 1992). If we take not only the lexical vs. grammatical meaning criterion but also the morphosyntactic impact of such elements into consideration —e.g. adpositions may assign case; auxiliaries may have their own subcategorization aspects— then they get a more or less intermediate status, halfway full predicate and purely grammatical element. The amount of 'lexicality' that such elements have in a language, both semantically and formally —arguably an indication of the length of the diachronic process they have gone through— may well be reflected in the relative height in the tree of the node that introduces them during expression.

So far for a first approximation to the dynamic Expression Rule model for FG-like grammars. Some of the details will be clarified on the basis of a concrete example in the next section.

5. DERIVING AN EXPRESSION

In this section we will derive the (prephonological) expression of a simple English sentence on the basis of the principles introduced in section 4. The main goal will be to demonstrate the dynamics of the approach and to show that certain problems with the standard model have been solved by it. For a fully worked out implementation of the dynamic model I will assume a feature-value notation with an accompanying set of logical operators and a unification mechanism to be available. Such a formal system is necessary to make the representations explicit and unambiguous, and to define the operations on them in a straightforward way.[23] However, in order not to burden the discussion here with unnecessary formal complexities I will use a simplified, pseudo-formal notational system, assuming that the labels that are used for features and values are more or less self-explanatory. The notational system uses the following conventions.

a. Unbound features are written in upper case: TENSE.
b. After binding features will be replaced by their value in lower case: past, rather than by a complete feature-value pair: TENSE=past.
c. Prephonological lexical elements are written in quotes: 'bicycle'; 'under'.
d. Grammatical elements are represented as sets of features or values between brackets without any lexical material:
[TENSE, NUMBER, PERSON, suffix] and after value assignment [pres, sg, [-sp,-hr], suffix] specifies the English verbal suffix 's'.
e. For Config, the usual representation for underlying clauses will be employed rather than a feature-value notation that makes the elements more explicit. Only the relevant parts of URs will be shown.
f. In case of possible ambiguity, a feature or value will be extended with a dot plus a disambiguing label: NUMBER.subject, sg.subject.

In none of the representations below, an attempt will be made to give full credit to the richness of the syntactic and morphological patterns of English, or other languages for that matter. The only goal envisaged here is to clarify the major mechanisms of tree expansion.

---

[23] A proposal for such a system may be found in Bakker (1994: 263-296). It is meant to be used for all levels of representation in FG, including underlying representations and the lexicon. This gives a unified flavour to all representations, and facilitates the construction of interfaces between the respective components of a grammar.

The sentence in (29) with its corresponding underlying representation will be taken as the point of departure here. It is relatively straightforward, and contains only a fraction of the morphosyntactic complexities the language in question may provide.

(29) Those girls are smart
[decl E_1: [X_1: [pres e_1: [‘smart’ [A] (rem pl x_1: ‘girl’ [N])$_{Zero,Subject}$]]]]

The first step in the derivation is the matching of the UR in (29) with a node in the grammar that expresses full underlying clauses. This will become the top node of the current expression. In its uninstantiated form, and leaving out some aspects not relevant for this discussion, that node might look as follows:

NODE N (uninstantiated)
Lab: sentence
Config: [ILLOC [[TENSE [PRED CAT ARG1]]]]
FncFtrs: ILLOC, TENSE, CAT, NUMBER.SUBJECT, PERSON.SUBJECT
FrmFtrs:
SubCat:

After unifying this node with the UC of (29), and thereby replacing the respective variables with their corresponding values, we get the partially instantiated node below:

NODE 1
Lab: sentence
Config: [decl E_1: [X_1: [pres e_1: [‘smart’ [A] (rem pl x_1: ‘girl’ [N])$_{Zero,Subject}$]]]]
FncFtrs: decl, pres, [A], pl, [-sp,-hr]
FrmFtrs:
SubCat:

The value for third person is assigned by default for all cases that are not first or second person. In order to make this node into a semi-instantiated one, the grammar has to provide it with a frame for the SubCat field. Under the assumption that any well-formed UR is expressible (cf. Dik 1997a: 151; Dik 1997b: 357f) there should be at least one such a ‘Sentence’ frame in the grammar.[24] In its semi-instantiated state, and again leaving out aspects that are irrelevant for this example the top node might look as follows.

[24] Under the (stricter) assumption that there are no fully synonymous expressions in a language there is precisely one such frame.

NODE 1 (semi-instantiated)
Lab: sentence
Config: [decl E_1: [X_1: [pres e_1: ['smart' [A] (rem pl x_1: 'girl' [N])$_{Zero,\ Subject}$]]]]
FncFtrs: decl, pres, [A], pl, [-sp,-hr]
FrmFtrs:
SubCat: p1, subject, vfin, mainpred

The next step now will be the expansion of the leftmost slot of the SubCat, p1. In the grammar, the corresponding node may look as follows:

NODE N (uninstantiated)
Lab: p1
Config: (TERM)$_{Focus}$ OR (TERM)$_{Topic}$ OR (TERM)$_{Subject}$
FncFtrs: DEF, DEM, NUMBER
FrmFtrs:
SubCat:

The logical expression in the Config field states that the preferred filler for this node is a term with Focus function. If such a term is not available in the Config of the mother node (here: NODE 1), then a term with Topic function should be selected, and finally the Subject term. In our case the latter will happen. So we get (empty and irrelevant fields will be left out from now on):[25]

NODE 2 (partially instantiated)
Lab: p1
Config: (rem pl x_1: 'girl' [N])$_{Zero,Subject}$
FncFtrs: def, rem, pl

On the basis of the contents of Config the node will be extended with a SubCat with label NounPhrase as follows:

NODE 2 (semi-instantiated)
Lab: p1
Config: (rem pl x_1: 'girl' [N])$_{Zero,Subject}$
FncFtrs: def, rem, pl
SubCat: det, num, adj, nomhead, relcl

[25] Note that, unlike examples in the FG literature, NODE 2 only expands termlike fillers of p1. For possible verbal fillers of p1 there should be an alternative version since, in general, the set of functional and formal features will be dependent on the type of filler. Note further that the value for definiteness is not formally present in the configuration. I will assume a redundancy rule to be active here, like rem → def.

Following the depth first strategy, the leftmost slot of the current SubCat will now be expanded, before the second slot of the SubCat of the mother node 1, i.e. subject, is taken into consideration. The grammar provides us with the following filler node for category 'det':

NODE N
Lab: det
Config: [FORM, [DET], DEF, DEM, NUMBER]
FncFtrs: DEF, DEM, NUMBER

After instantiation we are left with a terminal node of which the SubCat has precisely the necessary information to retrieve the lexical element 'those' from the lexicon. The form that was retrieved will fill the Subcat.

NODE 3 (instantiated)
Lab: det
Config: ['those', [DET], def, rem, pl]
FncFtrs: def, rem, pl
SubCat: 'those'

Since we reached a terminal node, the process pops back to the SubCat of the mother node, Node 2. After skipping the num and adj slots because there are no fillers for them, we reach the nomhead slot, which will expand the node for the nominal head. This will trigger the following instantiation:

NODE N (uninstantiated)
Lab: nomhead
Config: FORM, [N]
FncFtrs: NUMBER

NODE 4 (semi-instantiated)
Lab: nomhead
Config: 'girl', [N]
FncFtrs: pl
SubCat: noun, numsfx

Further expansion gives us the next two terminal nodes.

NODE 5 (instantiated)
Lab: noun
Config: 'girl'
SubCat: 'girl'

NODE 6 (instantiated)
Lab: numsfx
Config: ['s', [NUMSFX], pl]
FncFtrs: pl
SubCat: 's'

Since expansion stops at the prephonological level and no "lower" features will percolate up to them, nodes 3, 4 and 5 are by implication fully instantiated. In the SubCats we find the items that are to be processed further by the phonological component. For now, I will assume that the forms which are given in quotes are mere labels that give access to the precise phonological and lexical information, to which to apply intonation, sandhi, etcetera.[26]

Now we pop up via node 4 to the SubCat of node 2 where it is established that there is no filler for the last slot, relcl.[27] After this node 2 will be fully instantiated, i.e. its contents may not be altered anymore. Finally, we pop back to top node 1 again.

The next step leads us to the second slot in the SubCat of node 1, i.e. subject. This slot will retrieve a node from the grammar that has a Config prespecified for a UC element with Subject function. Since a general rule will mark all lexical UC material that has already been expressed, the instantiation of this slot with the Subject constituent will block.

The next slot to be explored is vfin, the slot for the finite verb form. There will be alternative nodes in the grammar for this node, at least the following two:

NODE N (uninstantiated)
Lab: vfin
Config: FORM, [V]
FncFtrs: TENSE, NUMBER.SUBJECT, PERSON.SUBJECT
SubCat:

NODE N (uninstantiated)
Lab: vfin
Config: [FORM, [COP]]
FncFtrs: [A], TENSE, NUMBER.SUBJECT, PERSON.SUBJECT
SubCat:

[26] There is evidence from psycholinguistics, however, that also at earlier stages some phonological information is available, at least for the predicates. We will come back to this briefly in section 6.

[27] Arguably, this slot could have been left out in advance by inspecting the Config, which does not contain a clausal restrictor. We need a theory on the pruning of templates on the basis of Config contents.

Note that in the second alternative, the Config specifies a grammatical item from the lexicon, a copula, rather than a substructure of the Config of the mother node. In the FncFtrs field, the value for the category of the main predicate has been preset to adjective. This works as a condition on the contents of the FncFtrs of the mother node. The second node is the one that will be selected in our case. After instantiation we get:

NODE 7 (semi-instantiated)
Lab: vfin
Config: [FORM, [COP]]
FncFtrs: [A], pres, pl, [-sp,-hr]

Lexical priority will select a template for this node that does not have the regular suffix slot for expressing a tense/number/person suffix. The whole contents will be expressed by one verbal form:

NODE 7 (semi-instantiated)
Lab: vfin
Config: ['be', [COP]]
FncFtrs: [A], pres, pl, [-sp,-hr]
SubCat: verbform

NODE 8 (uninstantiated)
Lab: verbform
Config: [FORM, EXPR, CAT, TENSE, NUMBER, PERSON]
FncFtrs: TENSE, NUMBER, PERSON

NODE 8 (instantiated)
Lab: verbform
Config: ['be', 'are', [Cop], pres, pl, [-sp,-hr]]
FncFtrs: pres, pl, [-sp,-hr]
SubCat: 'are'

Again we pop up to the top node and then fill the last slot for which there is material left in the Config, the main predicate slot:

NODE 9 (uninstantiated)
Lab: mainpred
Config: FORM, [A]

NODE 9 (semi-instantiated)
Lab: mainpred
Config: 'smart', [A]
SubCat: adjhead

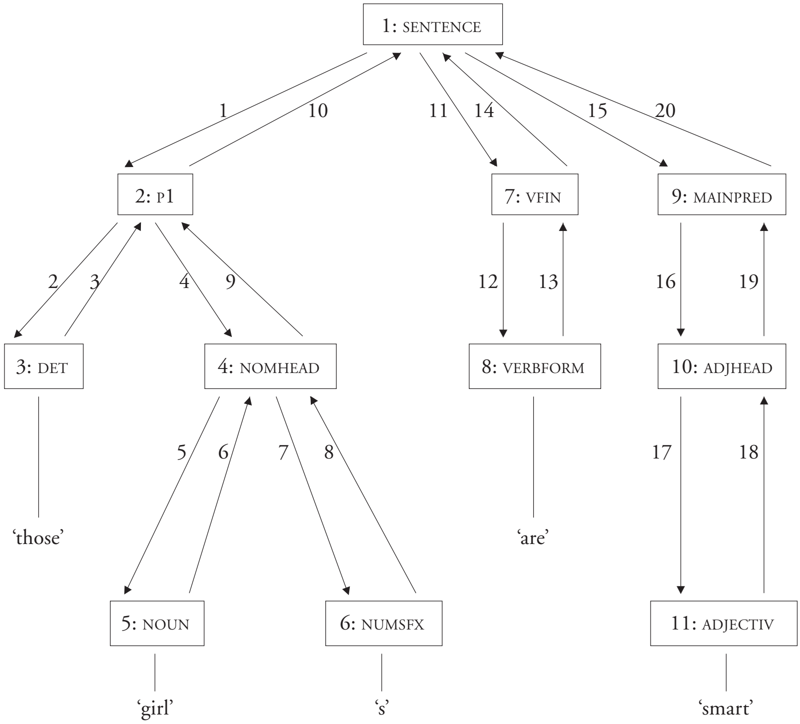

Figure 3. Fully expanded constituent tree

NODE 10 (semi-instantiated)
Lab: adjhead
Config: 'smart', [A]
SubCat: adjective

NODE 11 (instantiated)
Lab: adjective
Config: 'smart'
SubCat: 'smart'

After this control goes back to the top node, where all the material has been used for expression, and which is now fully instantiated, so the process comes to a halt. The trajectory that we have passed through could be rendered graphically as in figure 3.

This, of course, is not more than a first demonstration of only the most general aspects of the dynamic ER model on the basis of a very simple example, with many details still left out. A more complex example of tree expansion, based on the phenomenon of Relative Attraction (cf. Pinkster 1984 and Rijksbaron 1981), and which shows the working of the percolation mechanism may be found in Bakker (1999).

6. CONCLUSIONS

In this section we will see to what extent the dynamic ER model deals with the problems of the standard model as discussed in section 3. Firstly, the several examples of undergeneration will be discussed; secondly, we will have a look at the overgeneration aspect.

As shown in examples (13) and (14), summarized in (30) below, the standard model runs into errors when we apply the centrifugal principle too strictly.

(30) int past <3 sg> [marry [V]] -> *do married

The dynamic model has no problems with this construction. When we come to filling the Vfin slot the Int operator will trigger a node specifying the interrogative auxiliary. This node will also inherit the tense, number and person operators, eventually leading to the selection of the form 'did', as in nodes 7 and 8 in section 5. It will also introduce an auxiliary operator infinitive in its FrmFtrs field, which will percolate upwards to the top node and then downwards to co-determine the form of the main predicate 'marry'.

A second point that is not problematic for the dynamic model is the one demonstrated in example (15), more in general: any form of discontinuity in the expression of elements stemming from one underlying constituent. There does not seem to be any principled problem with expressing the respective elements of some underlying entity, say: a term, at different intervals during the expression process. The only condition is that they can be specified in an explicit manner, with the right set of operators, which is a technical rather than a linguistic matter.[28]

A third problem that we will look at is the one demonstrated in example (16), repeated partially in (31) below, where case is assigned by a grammatical element, here a preposition.

(31) Post hom-inum memori-am.
After people-PlGen memory-SgAcc
'From time immemorial'.

[28] Intuitively, it seems to be more complicated to extract an element such as a restrictor from a term while keeping track of which parts have been expressed rather than expressing the whole term at once. This intuition has been formalized in FG by Rijkhoff's (1990) Domain Integrity Principle.

Without some type of sandwiching there is no way in which the standard model could get the determination of the case affix of the head noun (typically a stage I operation) before the insertion of the adposition (typically a stage II operation). We have seen that in the dynamic model, the adposition could be inserted at any stage of the expression process, just like the demonstrative or the copula above. I suggest that this insertion takes place as soon as the corresponding term is selected for expression. On the basis of the functional information of the term itself, or lexical information from the main predicate, the right preposition may be retrieved from the lexicon and with it the case it assigns to its (formal) dependent. We might see this as an example of an element that has diachronically lost its status as a semantic head (it might in fact be a grammaticalized verb that had the corresponding term for its argument) but still retains its status as a syntactic head by being the "highest" element in the expression of the term.

A fourth category of problems for the standard model are all cases where form and order interact, as demonstrated in examples (17) through (19). In a general treatment of such cases the crucial features will become available only when certain elements have been expressed. Rather than selecting them in advance at a higher node, in the languages concerned they are captured when the corresponding node is expressed, and reach the rest of the expression via percolation. This will solve cases such as (18) and (19). Arguably, we may have an unspecified number feature before the subject term has been made available for expression, leading to a (default) singular verb form. The Dutch case of (16) is somewhat more complicated, since it also involves the person feature.

So it seems that at least part of the undergeneration in the standard model has been catered for. As for the overgeneration, and the lack of constraints in the standard model, we have already argued in section 4.1 that the dynamic model does away with several of the objections from the list (a)-(k) in section 3. The model allows for empirically testing constraints and for typologically experimenting with them, and also to derive predictions such as the agreement hierarchy in (24).

I will conclude this section, and this article by briefly mentioning some other aspects and problems of the dynamic ER model proposed here. This is not to say that the list is complete: only a large scale research programme based on a wide variety of morphosyntactic phenomena from the languages of the world may reveal the strength and the weaknesses of the current proposal.

A major point is the dynamics of the ER component. It seems that its left-to-right character, as opposed to the inside-out character of the standard model, apart from being intuitively attractive, both facilitates expression and might provide an explanation for a number of phenomena in speech behaviour. However, psychological models of speech suggest that a purely left-to-right model is too simplistic. There is ample evidence, not in the least from the analysis of speech errors, that language processing in the speaker takes place in a parallel fashion, roughly: while a first part of the sentence is being expressed, a second part may still be in the process of being semantically conceptualized, while a third part is still at a prelinguistic stage (cf. Levelt 1989). Such a cognitively still more adequate but principally more complex model for FG can only be worked out on the basis of an incremental model for the construction of underlying representations.

About the bottom side of the model —the determination of the phonological form— hardly anything has been said at all. It may well turn out that phonology should be integrated much in the way syntax and morphology have been integrated here. In general, this would imply that at least some phonological information is available at a much higher level than only at the terminal nodes. One might think of the predicates, which are encoded in URs anyway. Regularly observed speech errors as the one in example (32) from Dutch are highly suggestive of this. In (32b) the onset of the verb *kraait* 'crows' has been erroneously swapped with the onset of the noun *haan* 'cock' which is still far from being formally processed in most incremental models.

(32) a. Er kraait 'smorgens altijd een haan
b. *Er haait 'smorgens altijd een kraan
'Every morning a cock crows'.

Also other phenomena, such as backward vowel harmony and the "binyanim" of the Semitic languages (cf. Junger 1987), suggest a more integrated approach to phonology in ER.

Finally, several points that have been observed above only in passing are fundamental for a complete version of the dynamic model, and a complete theory of expression, if not already for the standard model. I mention the determination of the set of nodes in the grammar; a theory of subcategories and the existence of empty slots and their role. Also the status of grammatical morphemes vis à vis the lexicon should get the amount of attention it deserves, but never got within FG theory. However, having asked already enough patience of the reader, I will leave these points for another occasion.

Grammatical labels and abbreviations

A	adjective
ARG	argument
CAT	lexical category
cop	copula
decl	declarative
DEF	definiteness operator
def	definite
DEM	demonstrative operator
ERG	ergative marker
fin	finite
hr	hearer
ILLOC	illocutive operator
indef	indefinite

infin	infinite
INF	infinitive marker
int	interrogative
N	noun
PAST	past tense marker
pl	plural
Poss	possessive semantic function
PRED	predicate
pres	present
rem	remote (demonstrative)
sg	singular
sp	speaker
Temp	temporal semantic function
V	verb
Zero	empty semantic function

WORKS CITED

BAKKER, Dik. *Formal and Computational Aspects of Functional Grammar and Language Typology*. Amsterdam: IFOTT, 1994.

— "Flexibility and Consistency in Word Order Patterns in the Languages of Europe." *Constituent Order Patterns in the Languages of Europe*. Ed. Anna Siewierska. Berlin: Mouton de Gruyter, 1997. 383-420.

— "Language Change: Some Implications for the Theory of Functional Grammar." *Working Papers in Functional Grammar* 66 (1998).

— "FG Expression Rules: From Templates to Constituent Structure." *Working Papers in Functional Grammar* 67 (1999).

BAKKER, Dik, and Anna SIEWIERSKA. "A Contribution to the Problem of Constituent Order Explanation." *ESF Working Papers* II-5 (1992).

BALL and MÜLLER. *Mutation in Welsh*. London: Routledge, 1992.

BOERSMA, Paul. *Functional Phonology*. Amsterdam: IFOTT, 1998.

BORSLEY, Robert D., and Janig STEPHENS. "Agreement and the Position of Subjects in Breton." *Natural Language and Linguistic Theory* 7: 3 (1989): 407-428.

CARROLL, David W. *Psychology of Language*. Pacific Grove: Brooks/Cole, 1994.

CHOMSKY, Noam. *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT P, 1965.

— *The Minimalist Program*. Cambridge, Mass.: MIT P, 1995.

COMRIE, Bernard. *Aspect*. Cambridge: Cambridge UP, 1976.

— *Tense*. Cambridge: Cambridge UP, 1985.

CORBETT, Greville. *Gender*. Cambridge: Cambridge UP, 1991.

CROFT, William. "Autonomy and Functionalist Linguistics." *Language* 71.3 (1995): 490-532.

— *Radical Construction Grammar*, forthcoming.

CULICOVER, Peter W. *Principles and Parameters: An Introduction to Syntactic Theory*. Oxford: Oxford UP, 1997.

DIK, Simon C. *Functional Grammar*. Amsterdam: North-Holland, 1978.

— "On the Notion 'Functional Explanation'." *Belgian Journal of Linguistics* 1 (1986): 11-52.

— *Functional Grammar in Prolog: An Integrated Implementation for English, French, and Dutch*. Berlin: Mouton de Gruyter, 1992.

— *The Theory of Functional Grammar. Part I*. 1989. Berlin: Mouton de Gruyter, 1997a.

— *The Theory of Functional Grammar. Part II*. Berlin: Mouton de Gruyter, 1997b.

Dryer, Matthew S. "The Greenbergian Word Order Correlations." *Language* 68.1 (1992): 81-138.

Greenberg, Joseph H. "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements." *Universals in Language*. Ed. J.H. Greenberg. Cambridge, Mass.: MIT P, 1963. 73-113.

Hannay, Mike. "Pragmatic Function Assignment and Word Order Variation in a Functional Grammar of English." *Journal of Pragmatics* 16 (1991): 131-155.

Hawkins, John A. *Word Order Universals*. New York: Academic Press, 1983.

Hengeveld, Kees. "Layers and Operators in Functional Grammar." *Journal of Linguistics* 25.1 (1989): 127-157.

— "Cohesion in Functional Grammar." *Discourse and Pragmatics in Functional Grammar*. Ed. John Connolly, Roel Vismans, Chris Butler, and Richard Gatward. Berlin: Mouton de Gruyter, 1997. 1-16.

Junger, J. "Predicate Formation in the Verbal System of Modern Hebrew." Diss. University of Amsterdam, 1987.

Kempen, Gerard, and Eduard Hoenkamp. "An Incremental Procedural Grammar for Sentence Formulation." *Cognitive Science II* (1987): 201-258.

Kroon, Caroline. "Discourse Markers, Discourse Structure and Functional Grammar." *Discourse and Pragmatics in Functional Grammar*. Ed. John Connolly, Roel Vismans, Chris Butler, and Richard Gatward. Berlin: Mouton de Gruyter, 1997. 17-32.

Levelt, Willem J.M. *Speaking: From Intention to Articulation*. Cambridge, Mass.: MIT P, 1989.

Mackenzie, Lachlan. "English Spatial Prepositions in Functional Grammar." *Working Papers in Functional Grammar* 46 (1992).

— "The Basis of Syntax in the Holophrase." *Functional Grammar and Verbal Interaction*. Ed. M. Hannay and A.M. Bolkestein. Amsterdam: John Benjamins, 1998. 267-295.

Pinkster, Harm. *Latijnse Syntaxis en Semantiek*. Amsterdam: Grüner, 1984.

Plank, Frans. "The Co-Variation of Phonology with Morphology and Syntax." *Linguistic Typology* 2.2 (1998): 195-230.

Pollard, Carl, and Ivan A. Sag. *Head-Driven Phrase Structure Grammar*. Chicago: U of Chicago P, 1994.

Rijkhoff, Jan. "Explaining Word Order in the Noun Phrase." *Linguistics* 28 (1990): 5-42.

Rijkhoff, Jan, and Dik Bakker. "Language Sampling." *Linguistic Typology* 2.3 (1998): 263-314.

Rijksbaron, Albert. "Relative Clause Formation in Ancient Greek." *Predication and Expression in Functional Grammar*. Ed. A.M. Bolkestein, et al. London: Academic Press, 1981. 235-59.

Van Valin, Robert D., and Randy LaPolla. *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge UP, 1997.