Positions vs. precedences as primitives of constituent order

Lars Ahrenberg
Department of Computer and Information Science
Linköping university
S - 58183 Linköping

1. Introduction

This paper compares a number of different approaches to the formal description of constituent order regularities, in particular as they apply to Swedish. All approaches considered assume that it is meaningful and desirable to handle the sequential order of syntactic elements, the *topology* of the language, by rules and structures that are separate from those that express the combinatorial regularities, the *syntax* proper. Where the approaches differ is in the primitives assumed, and in the assumptions made about the nature of topological rules.

To begin with, a few remarks about our view on the nature of syntactic elements are in order. I will, following HPSG terminology refer to a syntactic object as a *sign*, and make the usual distinction between *lexical signs* (or *words*) and *phrasal signs* (Pollard&Sag, 1994: 31). The distinguishing property of a phrasal sign is that it is syntactically complex, i.e. has one or more syntactic objects as immediate constituents.

Three aspects of a phrasal sign will be distinguished: the type, the constituent signs and the sequential order of the constituents. The type includes categorial information as well as subcategorization properties, i.e. information that in HPSG is encoded under the CATEGORY attribute or as part of the sort. The constituents will primarily be identified with reference to their function as head, subject, complement or adjunct and be encoded by means of some subspecies of the DTRS ("daughters") attribute. I will make no attempt to define these functions, but just assume that they are valid.

In standard HPSG it is assumed that topological rules apply to the daughters of a phrasal sign, but this assumption is not made by all models that I will consider. I will use the term *constituent order domain* to refer to the domain of application of topological rules and represent it by the attribute DOM, when it is considered different from the set of daughter constituents.

The favoured kind of topological rule in many modern frameworks, including HPSG, is the LP-rule, or linear precedence rule. An LP-rule has the form 'D1 < D2', where D1 and D2 express syntactic constraints on constituents. Such rules will simply be referred to as *precedences* in the sequel.

The HPSG standard model turned out to be problematic when the theory was applied to other languages than English, *viz.* German. A case in point is the scrambling phenomenon, where complements appearing in the Mittelfeld are separated from their verbs by adjuncts or complements to other verbs. Scrambled constituent order is one of the phenomena that has been used to motivate the need for separate constituent order domains. The following clause, from Reape (1994), illustrates the problem:

(1) *dass es ihm jemand zu lesen versprochen hat.*
that it to-him someone to read promised has
"that someone has promised him to read it"

Here, the first complement (*es*) is the direct object of *lesen*, while the second complement (*ihm*) is a complement of *versprochen*, but separated from it both by the subject (*jemand*) and the governed verb *lesen*. While Swedish does not provide us with examples of this extreme nature, there are discontinuities in Swedish too, as will be illustrated shortly.

In Swedish, as in German, the order of constituents may also depend on the type. We have at least two cases where differences in constituent order seem constitutive for clause types: the distinction between main and subordinate clauses, as illustrated in (2) vs. (3), and the distinction between plain and inverted main clause, as illustrated in (2) vs. (4).

(2) *Han såg inte bilen.*
He saw not the_car.
"He didn't notice the car."

(3) *att han inte såg bilen*
that he not saw the_car
"that he didn't see the car."

(4) *Såg han inte bilen?.*
Saw he not the_car?
"Didn't he notice the car?"

This kind of dependency is seldom treated in depth, though Kathol (1995) is a noteworthy exception. However, it is not seen as a dependency between a type of phrasal sign and the application domain of a precedence; precedences are generally assumed to apply to domains without exceptions. Instead the type/order-dependencies are handled via features, or combinations thereof, which in turn are referred to in precedences.

Ahrenberg (1989; 1990; 1994) argued that notions from field grammar can be incorporated into formal frameworks such as LFG or HPSG to express topological constraints. In particular, I argued that rules that associate types of constituents with positions, together with a linear schema of positions, give us a simple and descriptively adequate account of the possible variations in the constituent order of Swedish clauses. The type/order-dependencies are handled by explicit category definitions that *inter alia* associate a type of phrase with specific type/position-associations. Thus, the difference between main and subordinate clauses will be grounded in the different positions associated with the finite verb for the two cases, and the inverted clause is characterized as a type of clause for which no constituent can be associated with the foundation (the first position of a clause).

The main purpose of this paper is to reconsider the usefulness of positions in a formal grammar for Swedish in light of some precedence-based models that have been proposed for German. In particular, I will be concerned with the nature of rules and domains: Are positions and schemata really required? Do we need larger constituent order domains than ones given by the immediate constituents of a clause?

I will first review some of the relevant work on German, primarily that of Reape (1994) and Kathol (1995), and contrast that with my own work for Swedish. This will provide the background for a comparison of the different approaches, including the HPSG standard model. Finally, I will conclude that a precedence-based account is indeed feasible, if we change our view of what is involved in phrase structure, and sketch an account based on this premise.

2. Three approaches to constituent order domains

When precedences were introduced into generalized context-free feature-based grammars (Gazdar et al., 1985), they were assumed to operate on the right-hand side of immediate dominance rules, or ID-rules. Together, ID-rules and precedences defined constraints on local phrase structure trees. As we noted in the previous section, standard HPSG adopts a similar assumption, although the role of ID-rules has been taken over by more general ID-schemata.

As Swedish is a V2-language the finite verb of a main clause is often separated from its complements. The constituents that may intervene are primarily the subject and various sentence adverbs. The latter was illustrated in (2) - repeated here for convenience - while the first is illustrated in (5). We note that the subject may appear before the adverbs, as in (6) or after them, as in (7).

(2) *Han såg inte bilen.*
He saw not the_car.
"He didn't notice the car."

(5) *Då först såg han bilen.*
Then first saw he the_car
"First then did he notice the car."

(6) *På grund av dimman såg han inte bilen.*
Because of the_fog saw he not the_car
"Because of the fog he didn't notice the car."

(7) *På grund av dimman såg inte fotgängaren bilen.*
Because of the fog saw not the_pedestrian the_car.
"Because of the fog the pedestrian didn't notice the car."

If we assume that a verb forms a phrasal sign with its complements, i.e. the sign or signs that it is subcategorized for, it will be difficult to develop a principled account for these sentences within the standard model. If we posit an ID-schema similar to Schema 3 of Pollard&Sag (1994), that combines the verb simultaneously with the subject and the complements, and the feature INV for verbs, (5) is accounted for. If we assume a rule that allows the verb to combine with a sentence adverb, as in Gambäck (1997), we can also handle (2) and (7), but then (6) still remains. Alternatively, we may follow the suggestions in Pollard&Sag (1987) and allow a schema where the verb combines with its complements and one or more adjuncts to account for (5), and another, similar schema that includes verb, subject, complements and at least one adjunct for (6) and (7).

While this latter approach seems the most promising for accounts of constituent order, it has not been given much support. Pollard&Sag (1987) states that it is beyond the expressive power of the formalism to express the necessary constraints between head and adjunct. Pollard&Sag (1994) argues that the adjunct selects its head rather than the other way around and assume head-adjunct structures to be binary, involving a single adjunct and a single head. If we accept this assumption, explicit constituent order domains seem inescapable. Kasper (1994), on the other hand, develops a model in which complements and adjuncts are allowed to be sisters, yet adjuncts behave semantically as functors. We

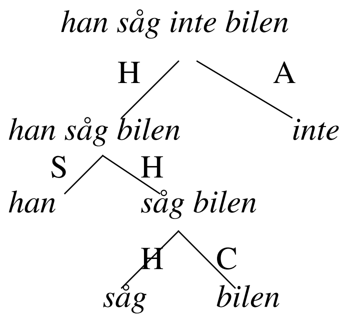

Fig. 1. Syntactic analysis of the sentence *Han såg inte bilen*.

will consider Kasper's model in a bit more detail in section 4. Now, let us turn to the models that actually use constituent order domains.

Reape can be said to have pioneered the notion of a constituent order domain within HPSG. His basic assumptions are the following (Reape, 1994):

- The mapping between the daughters of a phrasal sign and the phonology of that sign is mediated by its domain. Lexical signs do not have domains;
- A domain is an ordered sequence of signs;
- LP-rules define the possible orders of elements of a domain;
- The domain of a phrasal sign is composed compositionally from its lexical daughters and the domains of its phrasal daughters;
- A daughter may be an element of its mother's domain, or else its domain elements may be elements also of the mother's domain;[1]

Order domains can be combined into larger domains by an operation called domain union, or shuffling, which I will denote by the symbol '⊕'. The important property of this operation is that it preserves the relative order of elements. If S precedes T in some order domain, D, then S will precede T in any domain that is the result of unioning D with other domains. We note in particular that order domains are not assumed to have internal structure. They are simply containers that admit contents in the form of signs. The order between the signs is determined from properties and relations of the signs only.

Let us for the moment assume that sentence (2) can be analysed as the result of the application of three ID-schemata: a Head-Adjunct Schema, a Head-Subject Schema and a Head-Complement Schema. The details of these schemata need not concern us here (but see Pollard&Sag 1994: 402-3 for detailed versions); we simply note that *inte* (not) may be analysed as a clausal adjunct that combines with a sentence, that *bilen* (the car) is intended as a complement to the verb *såg* (saw), and that *han* (he) is the subject. This analysis is depicted in figure 1[2].

The position of the adverb *inte* is determined by precedences that in informal terms say the following:

Head < Complement

1. This assumption is further specified by Reape, but we omit the details as they are not important for our discussion.
2. If the negation *inte* is taken to combine with the predicate rather than with a full sentence, we will still need to shuffle it into the predicate, and sentences (5)-(7) would require shuffling of the subject into the predicate.

Head < ClausalAdjunct
Subject < Head
ClausalAdjunct < Complement

Each one of the phrasal verbal projections carries a domain attribute, DOM, the value of which is computed compositionally as follows:

DOM({såg, bilen}) = såg ⊕ bilen = <såg, bilen>
DOM({han, {såg, bilen}}) = <såg, bilen> ⊕ han = <han, såg, bilen>
DOM({inte, {han, {såg, bilen}}}) = inte ⊕ <han, såg, bilen> = <han, såg, inte, bilen>

Thus, the predicate, as well as the clause kernel, contribute their domain elements to the domain of the mother, in order to account for the discontinuity of the predicate.

Ahrenberg (1994) achieves the same effect by assuming a structured PATTERN attribute, the value of which is modelled on Diderichsen's main clause schema (Diderichsen, 1966). Unlike Reape's order domains, which contain signs, the patterns contain strings (corresponding to ordinary HPSG PHON values). Crucially, the pattern is taken to be shared by all projections of a lexical head. Each phrase structure schema provides a specific set of associations between daughters of different types and positions of the structured domain. An illustration is given in (8).

(8) vp (Verbal Head-Complement Structure);
PATTERN = S:<1, 2, 3, 4, 5>
DTRS:HEAD:STRING = 2
DTRS:CDTRS:STRING = 4

Using the same syntactic analysis as in figure 2, the constituent order emerges as a result of unifying strings into different positions of the pattern and structure sharing of the same pattern at the different head nodes.

Kathol&Pollard (1995) provides a number of arguments from German for the relevance of positions. In particular they follow den Besten (1983) in arguing for the view that the finite verb of a German main clause, and the complementizer of a subordinate clause, occupy the same position. They also argue that wh-expressions in German subordinate clauses occupy that position. I will not review the arguments here, but merely note that the first thesis has been argued for extensively also for Swedish (e.g. Platzack 1985; 1986), while the second requires some amendments when we look at the Scandinavian facts.

Kathol&Pollard (ibid.) also notes that finite verbs, complementizers and wh-expressions do not have categorial properties in common. Thus, they argue that we should not subsume them under some common expression type, but look for a proper encoding of positions. A detailed development of such an approach is provided by Kathol (1995).

While Kathol is inspired by field grammarians, he does not accept the notion of position as given. Positions are not proper linguistic objects, only signs are. We could say that Kathol reinterprets positions as different "linearization potentials" of constituents. These potentials are coded as values of the feature TOPO, for topology. For German, the values proposed are:

vf	Vorfeld
cf	Complementizer Field
mf	Mittelfeld
vc	Verbal Chain
nf	Nachfeld

The precedences that apply to constituents of a German clause are summarized below:

(9) [TOPO vf] < [TOPO cf] < [TOPO mf] < [TOPO vc] < [TOPO nf]

The propensity of the German complementizer *dass* to appear before all complements, verbs and adjuncts in a subordinate clause is thus accounted for in the lexicon through the feature specification [TOPO cf]. Also finite verbs are lexically specified as to their linearization potential, but by a disjunctive feature, [TOPO cf∨vc]. A rule ensures that only one constituent can carry the feature [TOPO cf], thus forcing the finite verb to carry the feature [TOPO vc] when it appears together with a complementizer.

Kathol also proposes a similar account for Swedish. The TOPO-values for Swedish are very much the same as those for German, apart from a small change of terminology and an extra value to handle initial complementizers in subordinate clauses with main clause constituent order. Using terms and abbreviations closer to the Scandinavian tradition, I will use the following list:

rc	Root Complementizer-initial
F	Foundation
v	Finite verb
Nf	Nexus field, apart from the first verb
Vf	Verbal Chain
Cf	Content Field

and the precedences:

(10) [TOPO rc] < [TOPO F] < [TOPO v] < [TOPO Nf] < [TOPO Vf] < [TOPO Cf]

Returning now to our sentence (2), *Han såg inte bilen*, we can show yet another way of getting the constituents in the proper order. Kathol's model employs a constituent order domain just as Reape's does, represented by the attribute DOM. Unlike Reape, Kathol holds that domain objects are not signs, but a "compaction" of signs, and give them a specific sort called *dom-obj*. As this distinction is not crucial for our discussion, I leave the details out. Another difference is that Kathol accounts for phrase structure in terms of composition relations rather than phrase structure schemata. Still, the essential traits of his treatment are that order domains are related compositionally to constituents and constituent domains, and that TOPO-values can be assigned either in the lexicon, or by the composition relations. The assignment of TOPO-values in rules is made through statements of the following kind:

> If the head is verbal and the compaction of a complement sign is a singleton list containing the domain object C, then C's TOPO-value is *Cf*;
>
> If the head is verbal and the compaction of an adjunct sign is a singleton list containing the domain object A, then A's TOPO-value is *Nf*;

We continue to assume the syntactic analysis for (2) according to figure 1. Going bottom-up in the tree of figure 2 and remembering that the finite verb is assigned [TOPO v], we then derive the constituent order as follows:

DOM({såg, bilen}) = DOM(såg) ⊕ bilen[TOPO Cf] = <såg, bilen>
DOM({han, {såg, bilen}}) = <såg, bilen> ⊕ han[TOPO F] = <han, såg, bilen>
DOM({inte, {han, {såg, bilen}}}) = <han, såg, bilen> ⊕ inte[TOPO Nf]
= <han, såg, inte, bilen>

Kathol's model can be said to provide a better integration of positions into a framework employing precedences as the basic means to account for constituent order constraints. In particular, it does not assume that order domains themselves have internal structure. On the other hand, it can express distributional equivalences in terms of TOPO values, something which Reape's model cannot.

3. A comparison

In this section, I will discuss the differences between the approaches in more detail and attempt to identify their advantages and drawbacks. The discussion will include conceptual and formal properties, as well as some new data. An overview of the most important differentiating properties of the three models under discussion is given in Table 1, where also the standard HPSG model is included for comparison.

If we follow Occam's razor, the burden of proof should fall on the side that makes more assumptions, and in this case that would seem to be the model assuming that domains are needed and required to have internal structure, i.e. Ahrenberg's model. Now, given Kathol's reinterpretation of positions as linearization potentials that can be referred to in precedence rules, it is not enough to argue that field schemata are essential to express distributional equivalences between different types of constituents, since that is adequately handled in his model. Neither can we argue that they are required to express cross-categorial similarities in the sequential ordering of constituents, since that, too, can be expressed by sets of rules referring to TOPO-values as in (9) or (10).

A putative advantage for the structured approach is that we can differentiate positions in terms of the number of elements they admit. In Kathol's model we can only make a distinction between linearization potentials that can be shared by several elements in a clause and those that can be carried by only one constituent. Moreover, the latter property is expressed, rather unintuitively, by a precedence of the form (11):

(11) [TOPO α] < [TOPO α]

	Standard model	**Reape**	**Kathol**	**Ahrenberg**
Type of domain	DTRS	No internal structure	No internal structure	Fixed internal structure
Rule types	Precedences	Precedences	Precedences with linearization potentials, and type definitions	Position associations, precedences, and type definitions
Type of derivation	Functional derivation	Rule-by-rule composition (shuffling)	Rule-by-rule composition (shuffling)	Unification

Table 1: Four models of topological rules.

Given the formal interpretation of precedences that Kathol employs , this will turn out to prohibit the occurrence of two constituents carrying the feature [TOPO α]. The reason is that a rule 'D1 < D2' is taken to mean that a constituent satisfying D2 must not precede a constituent satisfying D1 (cf. Kasper et al., 1995). But this is just what will be the case, if we have two constituents satisfying D1 with D1=D2.

The question then arises whether there are cases where we would like to make further distinctions. It is actually quite hard to find cases where two or three would constitute an upper limit on the number of elements in a position. However, there are cases where we find more than one constituent in the positions that are normally regarded as admitting only a single constituent. For example, there is a class of adverbs that may appear before the finite verb in Swedish main clauses:

(12) *Han bara skrattade åt oss.*
He just laughed at us
"He did nothing but laugh at us"

(13) *Det bara blev så.*
It just got so
"It just happened that way"

While *bara* is the most common adverb that can be placed in this way, there are a few others that allow this construction, e.g. *till och med* (even), *rentav* (simply, downright), *närapå* and *nästan* (both meaning 'almost', 'practically'). A common semantic feature of these adverbs is that they specifically relate to the meaning of the finite verb. When the relatum is a complement to the verb, the adverb obligatorily goes to the right of the verb:

(14) *Han är bara en liten pojke.*
He is just a little boy

(15) **Han bara är en liten pojke*
He just is a little boy

It should be observed that it is not obligatory to place the adverb before the verb, even in those cases where it semantically relates primarily to the verb. In both (12) and (13) the order of adverb and verb may be reversed.

If we would like to maintain that the v-position is a unary position, we have a few different options. We could assume the existence of an adverbial position between the foundation and the complementizer position, but that is quite ad hoc. We could also assume that the adverb forms a constituent with the verb, but then we would have to explain why this only happens when the adverb is placed before the verb.

Another word that may intervene between foundation and finite verb in main clauses is *så* (so), in sentences such as the following:

(16) *Om du vill så följer jag med dig.*
If you want_to so follow I with you
"If you want, I will keep you company."

(17) *Hur som helst, så är det en fin hund.*
In_any_case, so is it a nice dog
"In any case it is a nice dog."

The most salient constraint for the occurrence of this *så* is that the first constituent is a non-complement of the main verb, often a kind of meta-adverbial. Another important fact is that it is (almost always) optional. In fact, most normative works recommend you to avoid it in writing. The *så*-construction has been studied extensively from a functional perspective by Ekerot (1988), who analyses it as a kind of formal foundation, with the primary function of ensuring that the (declarative) sentence type is clearly indicated (*op.cit.* p. 273). Another view, that Ekerot also gives partial support, is that it functions as a marker of an important boundary, namely the end of an initial adverbial clause. Structurally he places *så* in the foundation, and argues that the preceding adverbial occurs outside of the clause proper. In my view, Ekerot's functional analysis is well motivated, but it seems awkward that the inclusion of an optional marker between the first constituent and the finite verb, necessitates a drastic re-analysis of the first constituent.

Another facet of the cardinality issue is whether positions may be empty. Both traditional descriptive analyses as well as Ahrenberg (1990; 1994) admit empty positions and sometimes even require it. A case in point is the inverted (yes-no interrogative) clause of the form illustrated in (4), repeated here for convenience, where Ahrenberg (1990) states a rule to the effect that the foundation must be empty.

(4) *Såg han inte bilen?*

However, this constraint can also be formulated in other ways, e.g. by saying that in this type of clause the first filled position must be the position of the finite verb. Kathol's (1995: 202) definition is:

(18) v1 ⟶ [DOM <[vb; [TOPO cf]]> o *l*]

where *l* is a list of domain objects and 'o' denotes the appending of two lists. It should be noted, however, that this definition introduces other means, apart from precedences, to account for a constituent order regularity. The impact of such a rule on the formalism is not clear and we are not told how it can be used in the reverse direction, i.e. to infer that a clause is v1 from the property expressed on the right hand side.

In many cases where descriptions based on field schemata assign empty positions, Kathol's model says nothing, which admittedly is simpler. This is an effect of the compositional nature of the model, and motivates his calling it a WYSIWYG model, since only constituents that are part of a clause are assigned any properties. A structured domain, on the other hand, always contributes its structure to the final analysis, which gives us analyses such as (20) for a simple sentence such as (19):

(19) *Kom!*
Come!

(20)

F	v	Nf	Vf	Cf
-	*Kom!*	-	-	-

Another problem, first noted by Togeby (1988), concerns spurious ambiguities, i.e. cases of unmotivated, alternative analyses at the domain level. Consider the following sentences:

(21) *Jag ser den inte.*
I see it not
"I don't see it."

(22) *Jag ser inte den.*
I see not it.
"I don't see it."

(23) *Jag har sett den*
I have seen it.

(24) **Jag har den sett*
I have it seen

(25) *Jag ser den.*
I see it.

Since the negation can only occur in the nexus field, (21) must be analysed as having the object pronoun there as well. The normal position for an object complement, however, is the content field, as shown by (23) and (24), and thus (25) can receive two different analyses.

(26)

F	v	Nf	Vf	Cf
Jag	*ser*	*(den)*	-	*(den)*

Evidently, the more fields we assume, the more spurious ambiguities we will obtain. If we divide the content field further into a nominal part and an adverbial part, as is traditionally done, we get trouble with the following sentences.

(27) *Jag insåg inte att det var sant förrän i går.*
I realized not that it was true until yesterday.
"I didn't realize that it was true until yesterday."

(28) *Jag insåg inte förrän i går att det var sant.*
I realized not until yesterday that it was true.
"I didn't realize until yesterday that it was true."

(29) *Jag insåg inte att det var sant.*
I realized not that it was true.
"I didn't realize that it was true."

The normal analysis for (27) would take the complement clause to appear in the nominal part and the temporal adverbial to appear in the adverbial field. But the order may be reversed, implying that the clause is now in the adverbial field (or yet another field following it). In either case, (29) can be analysed in two different ways.

It should be pointed out that these spurious ambiguities are not particularly dependent on schemata, but on the use of positions to encode alternative sequential placement. Thus, they arise in Kathol's model as well. The relevant generalization in the case of (21)-(25) seems to be that the verbal head precedes its complement, except when the latter is a filler placed in the foundation. Kathol (1995: 362) actually states a constraint to that effect to rule out (24). But as the rule refers only to verbs in the Vf-field, it is not as general as it could be.

Also other generalizations that can be made as regards sequential order are missed when positions are used. An example is the rule that the subject precedes all complements, again with the exception of a complement in the foundation. It is true that the subject can be placed either in the foundation or the mid field, so the fact as such can be inferred from the precedences in (10). It may be asked, however, whether it would not be better to state it directly as a single precedence, without making reference to positions or linearization potentials.

Yet another generalization that is not stated concerns the complementizer *att* (that). The truth of the matter seems to be that *att* is always the first element of the clause, whether the clause is subordinate as in (30) or main as in (31).

(30) (Jag är säker på) att han inte vill träffa henne.
(I am sure of) that he not want_to see her.
"(I'm sure) that he doesn't want to see her."

(31) *(Han säger) att henne vill han absolut inte träffa.*
(He says) that her want_to he absolutely not see
"(He says) that he absolutely would not want to see HER."

The dependent clause of (31) is often said to be "asserted" and it seems that it is mostly occuring with verbs of saying and believing. The problem with it, in an account based on positions, is that the dependent clause has main clause constituent order with a finite verb in the v-position and even a foundation. Thus, *att* in (31) cannot be in that position, but instead has to be given a position of its own, the rc-position of (10). Again, this may give rise to different analyses of (30) unless we introduce some extra constraints, and introduces alternatives where apparently we have a general rule.

We may make a distinction between perceptually salient and perceptually abstract precedences. A precedence is abstract if it refers to features of a constituent that cannot be determined on the basis of its morphemes. Precedences referring to TOPO values are abstract in this sense and so are, of course, the field schemata. Now, it seems difficult to avoid abstract rules in grammars of natural languages, but if two models that differs in abstractness can account for the same range of data, the less abstract model should be preferred. Generally speaking, a precedence referring only to salient properties of two constituents, and some independently motivated relation between them, is less abstract than a rule that refers to arbitrary properties or relations that are posited for the purpose at hand.

If we turn for a moment to the models that make use of neither field schemata nor linearization potentials, they are not free of abstract features either, but in fact employ a number of them for topological relations. The feature +/-INV, that is used on verbs in standard HPSG, to account for inverted vs. non-inverted order, has some independent motivation in English, where only auxiliary verbs can invert, but is completely arbitrary in accounts of German or Swedish, where all verbs can invert. Thus, it could rightly be called a topological feature in a description of Swedish, since its sole motivation is topological. Reape employs the feature UNIONED with values '+' or '-' to distinguish a composition that allows discontinuities from one that does not. Similarly, features like +TOPIC, for constituents appearing in the foundation, or +EXTRA for extraposed constituents are basically topological features, although other properties are alluded to. As Kathol points out (1995: 129), these features are problematic. First, they are not independent as there can be no constituent being +TOPIC and +EXTRA at the same time. For this reason it seems preferable to treat topological properties by a common feature where different positions of a clause are represented by mutually exclusive values. Another objection is that these features, unlike most morphosyntactic features, do not divide the universe of constituents into natural classes; e.g. the feature -TOPIC does not pick out a natural class of constituent. This criticism may be directed also at the TOPO-feature, however. While Kathol&Pollard (1995) demonstrate the existence of topologically related phenomena at the left periphery of German clauses, which may be represented by the feature [TOPO cf] of the affected constituents, the other values are given no similar support. The feature [TOPO vc] is restricted to verbs and verbal particles, but since verbs have independent characteristics, the order constraints might just as well be expressed in terms of those characteristics.

So, what conclusion can we draw from these arguments? First of all, we have not been able to establish any convincing argument that structured domains are required in a formal account of the topology of Swedish clauses. The relevant aspects of a position schema, i.e. the order of the positions can be expressed by precedences referring to linearization potentials, as demonstrated by Kathol (1995). However, Kathol's notion of position has some drawbacks too; it is equally abstract and it does not support concise and unequivocal statements of some of the most salient general facts of constituent order in Swedish clauses. Moreover, there exist examples, such as (12) and (16), where two constituents appear before the finite verb, throwing doubts on the assumption that the first two positions are unary. On the other hand, an account that encodes idiosyncratic topological properties by a single feature, such as the TOPO feature, seems more adequate than one, such as Reape's, that employs a number of different features for the same purpose.

4. Towards a restricted precedence-based topology

As has been said already, it is difficult to develop a descriptive account of the constraints on constituent order in Germanic languages, without referring to arbitrary, and hence abstract properties and relations. What we can aim for is a formally restricted model that allows the right generalizations to be expressed with a minimum of arbitrary features. Here I will give a sketch of a model that goes some way to satisfy these desiderata.

An innovative grammatical concept of HPSG is the marker. According to Pollard&Sag (1994: 45) a marker is *"a word that is 'functional' or 'grammatical' as opposed to substantive, in the sense that its semantic content is purely logical in nature (perhaps even vacuous)."* Moreover it "marks" the phrase in which it occurs in a specific way. The prime example of a marker that Pollard&Sag (1994) provides is the complementizer *that*, but they argue that other conjunctions and subjunctions have a similar behaviour.

Underwood (1997), referring to Van Eynde (1994), argues that a broadened notion of marker is required. The distinction between functional and substantive cannot be made between parts of speech; instead it has to be recognized that this distinction cuts across the parts of speech. For example, auxiliary verbs are functional whereas most verbs are substantive.

Here I would like to argue for a broadening of the notion of marking in a different direction. If complementizers are markers, by marking clauses as subordinate (or subordinate of a specific type), it would be natural to speak of any constituent that contributes to the identification of a clause as being of a specific subtype, as a marker. Kathol (1995: 138) states that the complementizer of a subordinate clause, and the finite verb of the main clause, share the same position and pragmatic function, *viz.* to distinguish clause-types, but does not relate that to the notion of marking. Heltoft (1986), on the other hand, makes a distinction between the semantic, or logical function, of a finite verb, and its function as a marker of modality[3]. What he suggests is that the first two positions of the clause schema, the foundation and its immediate neighbor to the right, contributes to the marking of modality in Danish. A similar analysis will work also for Swedish, but since we have found good reason to abandon the positions and the schema, we will couch it differently.

Let us assume, then, that a constituent may perform two functions simultaneously in a Swedish clause. It may have a substantive function in contributing to the predicate/ argument-structure of a clause, and a marking function in contributing to the distinction between different clause types. Clause types are grammatically defined, but are of course correlated with differences in interpretation. To emphasize the formal nature of the clause types, I will avoid terms such as declarative and interrogative, and instead rely on terms

3. "Verbet er konstituerende fordi det angiver saetningens modalitet eller realitetsindhold (forestillingsindholdets "Forhold til Virkeligheden") (Heltoft, 1986: 109 referring to Diderichsen). "/at/ er en formel pladsutfyldning, der markerer det finitte verbs fravaer og dermed saetningens topologiske umarkerethed med hensyn till modalitet (*ibid.* p. 115).

such as plain, inverted, wh etc.

It seems necessary to distinguish two primary marking functions:

- the modality marking function, m, which is assigned to the finite verb of a main clause;
- the foundation, F, which can be assigned almost any element of a main clause apart from the finite verb.

In addition the subject may be treated as both a marker and as an argument provider. For simplicity, however, we will only employ one subject function.

There are in fact several ways in which we can move towards a formal account based on marking functions.[4] The one I will present here is like standard HPSG in that it makes use of precedences and has no DOM-attribute. It differs from the standard model in assuming a flat clause structure and in doing away with ID-schemata. Thus, all the different types of constituents that can be daughters of a clause are on the same level. In this respect the proposal is similar to models employing field schemata.

The proposal handles dependencies between clause types and constituent orders in terms of category definitions. These definitions fall within the borders of a system of well-typed feature structures, as is required in HPSG (Pollard&Sag, 1994: 18; Carpenter, 1992). This means that an appropriateness specification (Approp) exists for every combination of sorts and features that the grammar allows, where an appropriateness specification is a function from pairs of sorts and features to sorts, requiring

- that every feature used is introduced at a single, most general sort, and
- that the value of a feature f, at a sort s, is at least as specific as the value of f at any appropriate sort of which s is a subsort.

Sorts are defined explicitly in an inheritance hierarchy. There are no phrase structure schemata as such; the constituents that can go with a specific clause type are defined through the appropriateness specification. Constituent order is determined by the precedences. These, in turn, may refer the function as well as the features of a constituent.

The distinction between marking and ordinary constituents is formally made by splitting the DTRS attribute into two: M-DTRS for marking constituents, and P/A-DTRS for those contributing to predicate/argument-structure.

The value of M-DTRS is a feature structure of one of the following sorts:

M, F, FM, m, Fm, Fzm, MFm, unmarked

Sorts including the letter *M* have a marker that is not relevant for predicate/argument structure, such as *att* (that). Sorts including the letter *F* have a foundation, while sorts including the letter *m* has the finite verb as marking constituent. A list of clause types and their sorts is given in Table 2. Of course, all differences between clause types are not covered by the topological marking system; other differences need also be encoded in the definitions.

Apart from these maximal sorts the description utilizes more general sorts that correspond to different ways of classifying clauses and a set of attributes for different kinds of markers.

4. The one presented here is not identical to the one given in my "Ledstillingsfestival" talk in Kolding (8th September 1997). The two proposals share the same underlying assumption, however, that constituents may have both a substantive grammatical function and a marking function. The major difference is that I use set-valued features here, instead of multi-valued roles.

Sorts	Examples
unmarked	Subordinate clause without marking: *(jag tror) han såg den* Relative clause without marking: *(bilen) han såg*
M	Subordinate att-clause: *att han inte såg bilen*; Subject Relative clause: *som såg bilen*;
F	Subordinate wh-clause: *vad han såg* Non-subject Relative clause: *vars bil han såg*
FM	Subordinate wh-clause: *vem som såg den*
m	Main inverted clause: *såg han bilen*
Fm	Various main clauses: *han såg bilen, då såg han bilen*
Fzm	så-construction: *Hur som helst så såg han den tydligt*
MFm	"Asserted" att-clause: *att han vill inte se den*

Tabell 2: M-DTRS sorts and clause types.

These attributes are:

M-DTR	Marker
F-DTR	Foundation
m-DTR	Modality marker
Z-DTR	Extra foundation

For instance, the sort __*m* will denote any clause that use the finite verb for marking, i.e. any main clause, and this sort will be the locus for the introduction of the attribute m-DTR. At the same time, it will be used to constrain M-DTRS in the definition of main clauses:

(32) Approp(__*m*, M-DTR) = [+VERB, +FIN]
Approp(*MainCl*, M-DTRS) = __*m*,

The hierarchy of sorts appropriate for M-DTRS is shown in figure 2, below.

The value of P/A-DTRS is a feature structure of one of the following sorts:

h, hS, hC, hA, hSC, hCA, hSCA

The labels for the sorts are meant to be mnemonic: the letter *h* indicates a clause with a head, the letter *S* a clause with a subject, *C* a clause with complements, and *A* a clause with adjuncts. The attributes to be used with these sorts are:

H-DTRS	A list of head daughters
S-DTR	Subject
C-DTRS	A list of complement daughters
A-DTRS	A list of adjunct daughters

The assumption that we may have more than one head is not standard, but is motivated by the fact that some auxiliaries do not contribute any valency constraints of their own. Rather, it is the valency constraints of the main verb that should be used. It is also a fact that head features of verbs may be distributed over an auxiliary and a main verb (cf. Underwood, 1997: 51). An example is (33), where we have a compound tense, and regard the auxiliary and the main verb as two different head daughters.

(33) *Jag har inte sett honom i dag.*
I have not seen him today.

The combinatorial properties of a head word is given by the value of the SUBCAT attribute. However, we distribute different types of argument on different attributes dominated by SUBCAT. The complements required by a head word are specified under SUBCAT|COMPS. The subject requirements are specified under SUBCAT|SUBJ. The main verb required by an auxiliary verb is specified under SUBCAT|VCOMP. To ensure that the subcategorization requirements can always be found on the same element, the first element of the H-DTRS list, we require this element to be the finite verb. Moreover, we require the subject and complement subcategorization properties of a non-finite main verb to be identical to those of an auxiliary verb that dominates it. The resulting structure is shown in (34).

(34)

$$
\text{H-DTRS}\left\langle
\begin{bmatrix}
\text{...|SUBCAT|SUBJ|[1]} & \\
\qquad\qquad\text{COMPS |[2]} & \\
\qquad\qquad\text{VCOMP |[3]} & \\
\text{... |HEAD} & \begin{bmatrix} \text{VERB} & + \\ \text{FIN} & + \\ \text{AUX} & + \end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\text{...|SUBCAT|SUBJ|[1]} & \\
\qquad\qquad\text{COMPS |[2]} & \\
\text{...|HEAD [3]} & \begin{bmatrix} \text{VERB} & + \\ \text{FIN} & - \end{bmatrix}
\end{bmatrix}
\right\rangle
$$

The complete analysis of (33) is illustrated in (35), where the local references are chosen so as to correspond with the constituent's sequential order in the clause.

(35)

$$
{}_{\textit{marked-headed-struc}}\begin{bmatrix}
\text{M-DTRS} & {}_{\textit{Fm}}\begin{bmatrix} \text{F-DTR} & \text{[1]} \\ \text{m-DTR} & \text{[2]} \end{bmatrix} \\
\text{P/A-DTRS} & {}_{\textit{hSCA}}\begin{bmatrix} \text{H-DTRS} & \{\text{ [2: har], [4: sett] }\} \\ \text{S-DTR} & \text{[1: Jag]} \\ \text{C-DTRS} & \{\text{ [5: honom] }\} \\ \text{A-DTRS} & \{\text{ [3: inte], [6: idag] }\} \end{bmatrix}
\end{bmatrix}
$$

To be well-formed a phrasal sign has to satisfy both combinatorial and topological constraints. The combinatorial wellformedness constraints are expressed in the general form of principles. In the manner of The Subcategorization Principle (Pollard&Sag, 1994: 399) we formulate three principles concerning a head and the constituents it subcategorizes for.

The Complement Principle: In a phrasal sign the value of P/A-DTRS|C-DTRS is token-identical with the value of SYNSEM|LOCAL|CATEGORY|SUBCAT|COMPS of the first element of the H-DTRS list.

The Subject Principle: In a phrasal sign the value of P/A-DTRS|S-DTR is token-identical with the value of SYNSEM|LOCAL| CATEGORY| SUBCAT|SUBJ of the first element of the H-DTRS list.

The Main Verb Principle: In a phrasal sign the value of SYNSEM|LOCAL| CATEGORY| HEAD for the n:th element of the H-DTRS list is token-identical with the value of SYNSEM|LOCAL|CATEGORY| SUBCAT|VCOMP for the (n-1):th element of the H-DTRS list.

As for adjuncts I follow Kasper (1994) in separating the syntactic and semantic constraints. I will assume a simple syntactic constraint that make adjuncts dependent only on the category of the head element:

The Syntactic Adjunct Principle: In a phrasal sign the MOD|SYN value of an adjunct is token-identical to the CATEGORY value of the first element of the H-DTRS list.

As for the semantics of adjuncts I assume that Kasper's principle can be transferred to the current model. The basic idea is that adjuncts apply in reverse order to the one of the A-DTRS list. The last adjunct on the list applies to the result of applying the head to the complements, while other adjuncts apply to the result of the previous application (Kasper, *op.cit.* p. 58). The result that pertains to the clause as a whole is given by the application of the first adjunct on the A-DTRS list. There are two complications in the current model, that are not present in Kasper's model: (i) the presence of a subject, and (ii) the possibility of having several heads. While I believe that a semantic account can be given along these lines, I will not pursue the matter here, but leave it as an assumption.

So, what about topology and precedences? Domains for precedences are given by the set of immediate daughters and this set is identical to the signs dominated by the attributes M-DTRS and P/A-DTRS. Formally, we may introduce a DOM-attribute to mediate between the signs and the string, but this seems quite unnecessary. As in standard HPSG, we may refer to a function 'Order-constituents' that apply to a set of daughters, and produce a string that is some permutation of the strings of the daughters.

A problem with a precedence of the kind 'Head < Complement', when the domain is a complete clause, is that it will be false for complements in the foundation. In the current model, however, all complements appearing in the foundation serve as markers, and a precedence of the kind 'Foundation < Head' is available. We still have a problem, though, in securing that the right precedence will apply. This can be achieved through the following principle:

The Preference for Markers Principle: If a constituent is a marking constituent, then only precedences referring to marking functions can apply to it.

This principle is tantamount to saying that the precedences are divided into two sets, one set that applies to markers, and another set that applies to pairs of non-markers. For Swedish, the most important precedences that refer to markers are:

(36) F-DTR < m-DTR
F-DTR << Z-DTR
M-DTR < m-DTR$_i$
M-DTR < P/A-DTR$_i$
M-DTR|*att* < F-DTR
F-DTR < M-DTR|*som*
m-DTR < P/A-DTR$_i$

Here, if an attribute X-DTRS has a list value, the reference X-DTR$_1$ is used as a reference

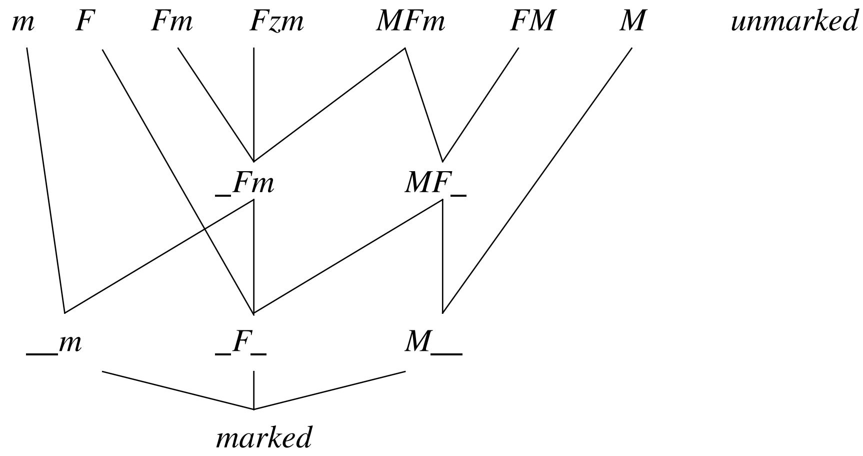

Fig. 2. A sort hierarchy for combinations of marking constituents.

to the first element, X-DTR$_2$ as a reference to the second element, and so on. The reference X-DTR$_i$ is used for an arbitrary element on the list or set X-DTRS. The symbol '<<' means "immediately precedes".

For non-markers we can state the following precedences:

(37) H-DTR$_i$ < C-DTR$_j$
S-DTR < H-DTR$_i$
S-DTR < C-DTR$_i$

The first precedence is actually too strict, as there are cases where a complement can precede a head. This holds in particular for negative complements, as in (38).

(38) *Hon har ingenting sett.*
She has nothing seen
"She has seen nothing"

Also the third is too strict when we consider reflexives, but my intention is not to give a complete set of precedences that are valid for Swedish, so we will leave this matter as it is. Some further remarks are in order, however, to show what would be required for the construction of such a list.

First, let us consider the distinction between main and subordinate clauses in more detail. The possible sorts for M-DTRS with a main clause are *m, Fm, Fzm*, while the possible sorts for a subordinateclause are *M, F, FM* and *unmarked.* The value *MFm* is appropriate for an asserted dependent clause and can be regarded as a third alternative: both main and subordinate. A possibility worth some consideration is to regard *unmarked* as a separate class (neither main nor subordinate), but since an independent occurrence of an unmarked clause can always be analysed as main clause of sort *Fm*, we regard *unmarked* as a sort for dependent, subordinate clauses.

We introduce the more general sorts *M__* for clauses with a non-head, non-argument marker, *_F_* for clauses with a foundation, and *__m* for clauses with a marking finite verb. The underscore in a sort label is used to indicate the possible, but not necessary presence of another marker, where the position of the underscore reveals which marker is intended. This gives us the sort hierarchy in figure 2 as appropriate for clauses and the attribute M-DTRS.

Similarly, we introduce a hierarchy of sorts that are appropriate for H-DTRS, as in figure 3. As before the underscore indicates the optional presence of a certain type of constituent.

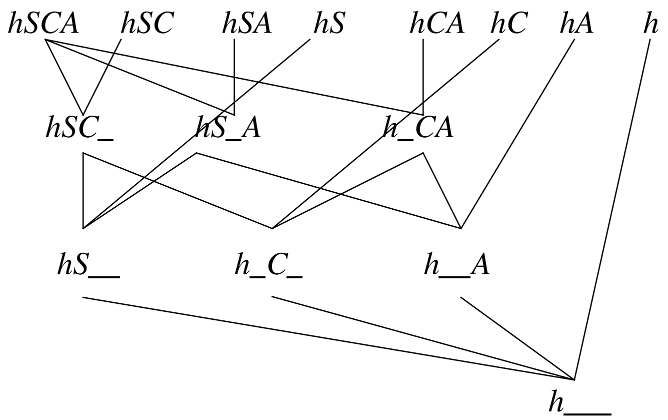

Fig. 3. A sort hierarchy for combinations of heads and arguments.

The order of the different types of constituents in a sort name is head, subject, complement, adjunct.

A main clause (*MainCl*) can now be defined as a clause satisfying the following conditions, where the sort *Clause* has already been defined as a subsort of phrase requiring the attribute P/A-DTRS with a value of sort *h___*. The symbol '⊂' is used to denote the subsort (or is-a) relation:

(39) *MainCl ⊂ Clause;*
Approp(*MainCl*, M-DTRS) = *__m*,
M-DTRS|m-DTR = P/A-DTRS|H-DTR$_1$

Thus, we take the prime characteristic of a Swedish main clause to be the use of the finite verb as a modality marker. We then obtain the different subtypes of main clauses by adding further conditions. For instance, inverted clauses as in (40), imperative clauses as in (41) and a plain subject-initial declarative clause as in (42).

(40) *InvCl ⊂ MainCl;*
Approp(*InvCl*, M-DTRS) = *m,*
Approp(*InvCl*, P/A-DTRS) = *hS__*,
P/A-DTRS|H-DTR$_1$ |...|FIN = +;

Example: *Såg han bilen?*

(41) *ImpCl ⊂ MainCl;*
Approp(*ImpCl*, M-DTRS) = *m*,
P/A-DTRS|H-DTR$_1$ |...|VFORM = *imp*;

Example: *Se på mig!*

(42) *PlainCl ⊂ MainCl;*
Approp(*PlainCl*, M-DTRS) = *Fm,*
Approp(*PlainCl*, P/A-DTRS) = *hS__*,
M-DTRS|F-DTR = P/A-DTRS|S-DTR,
P/A-DTRS|H-DTR$_1$ |...|FIN = +;

Example: *Han såg bilen.*

To distinguish a plain clause from a wh-clause with initial subject, we need to put further conditions on the subject foundation, but we omit the details of this distinction.

A subordinate clause can now be characterized negatively as a clause which is not main.

For short I refer to the relevant sorts by the label *__0*, implying their disjunction. Thus, the zero is used to indicate the absence of some function. There are several different varieties of subordinate clauses, however. Clauses that contain the independent marker *att* (that) are defined as follows:

(43) *attCl* ⊂ *Clause;*
Approp(*attCl*, M-DTRS) = *M__,*
Approp(*attCl*, P/A-DTRS) = *hS__,*
M-DTRS|M-DTR|STRING = att,
P/A-DTRS|H-DTR$_1$ |...|FIN = +;

In the normal case there will be no other marking apart from the conjunction itself. Thus, we obtain the definition of the "unmarked" *att*-clause, *attSubCl*.

(44 *attSubCl* ⊂ *attCl;*
Approp(*attSubCl*, M-DTRS) = *M;*

Example: ... *att han såg bilen.*

The subordinate, "asserted" clause can be defined as a clause that is simultaneously an *att*-clause and a main clause, something which seems desirable, given its distribution and interpretation.

(45) *attMainCl* ⊂ *MainCl;*
attMainCl ⊂ *attCl;*
Approp(*attMainCl*, H-DTRS) = *MFm,*

Example: ... *att han såg inte bilen.*

This definition is not entirely satisfactory, as it allows a simple *att*-clause to be analysed either as *attSubCl* or as *attMainCl*. We could remedy this by introducing further constraints, but it is not certain that this is the best way to go about it. These clauses have a restricted distribution as complements of verbs of saying and believing, and in such cases it seems that we have a real ambiguity, or at least a non-specific interpretation.

When there is no complementizer in a subordinate clause the sort of the M-DTRS value is *unmarked* and there are no daughters with a marking function. Apart from these differences, the definition contains the same clauses as (43).

We have so far not discussed the variation in order between clausal adverbs and finite verbs, that is usually taken as a discriminatory property of main and subordinate clauses. First it should be noted that the precedences in (36) guarantee that the finite verb will be placed before any adverb which is not a foundation. The precedences given for P/A-DTRS sofar, however, do not say anything about the possible order between heads and adjuncts.

First of all it is necessary to introduce a distinction between different kinds of adverbials. At least we need to distinguish three types of adverbials: (i) those that must precede the head verbs of a subordinate clause, (ii) those that must follow, and (iii) those that can appear on either side. It is evident that these differences are correlated with size and with semantic impact. The adverbs that are most common in mid position are words and lexicalized phrases that express modality or speech act, while circumstantial adverbials of time, location or cause are rarer there. Nevertheless, I will only offer an arbitrary solution, by introducing a feature CLAUSAL with values '+' or '-'. Adverbials that can appear in either position will be unmarked for this feature.

With reference to this feature, we may formulate the following precedences:

(46) [CLAUSAL +] < H-DTR$_j$
H-DTR$_j$ < [CLAUSAL -]

Sentences such as the following would then be ruled out, given appropriate lexical specification of the adverbs *inte* (not) and *i dag* (today):

(47) **Jag har inte i dag sett honom.*
I have not today seen him.

(48) **Jag har sett honom inte i dag.*
I have seen him not today.

The precedences stated so far also account for object shift. Both (49) and (50) are tolerated, while (51) is ruled out because of the precedence that require a complement to follow a head:

(49) *Jag ser den inte.*
I see it not.
"I don't see it."

(50) *Jag ser inte den.*
I see not it.
"I don't see it."

(51) **Jag har den (inte) sett.*
I have it (not) seen.

What we don't account for is the preference of (49) over (50), but this is something that cannot be accounted for within a model relying on strict formal rules. Also, to rule out clauses such as **Jag ser bilen inte* (I don't see the car), where a full np precedes the adverb, requires further precedences.

The main, indicative clauses that have other constituents than the subject as foundation can be divided into three types: (i) those for which the foundation is an adjunct of the matrix verb, (ii) those for which it is a complement of the matrix verb, and (iii) those for which it is a complement of some constituent. What they have in common is that there is a foundation, and this common property may be attributed to a sort *FInvCl* (Inverted clause with foundation):

(52) *FInvCl* $\subset$ *MainCl;*
Approp(*FInvCl*, M-DTRS) = *Fm,*
Approp(*FInvCl*, P/A-DTRS) = *hS__,*
P/A-DTRS|H-DTR$_1$ |...|FIN = +;

The first two subtypes (complement foundation or adjunct foundation) have straightforward definitions:

(53) *CFInvCl* $\subset$ *FInvCl;*
Approp(*CFInvCl*, P/A-DTRS) = *hSC_,*
M-DTRS| F-DTR = P/A-DTRS|C-DTR$_i$;

Example: *Den ser jag inte.*

(54) *AFInvCl* $\subset$ *FInvCl;*
Approp(*AFInvCl*, P/A-DTRS) = *hS_A,*
M-DTRS| F-DTR = P/A-DTRS|A-DTR$_i$;

Example: *Nu ser jag den.*

A special variant of adjunct foundation is the one where a formal foundation, *så*, intervenes between the adjunct and the finite verb (cf. examples (16), repeated here for convenience):

(16) *Om du vill så följer jag med dig.*
If you want_to so follow I with you

"If you want, I will keep you company."

It is natural to regard the *så* as simply an extra marker and the whole construction as an alternative to (55). The construction is defined thus:

(55) *AFInvCl2* ⊂ *FInvCl;*
Approp(*AFInvC2*, M-DTRS) = *Fzm,*
Approp(*AFInvCl2*, P/A-DTRS) = *hS_A,*
M-DTRS| F-DTR = P/A-DTRS|A-DTR$_i$;
M-DTRS| Z-DTR = så;

The third subtype involves an unbounded dependency. Here, I would rather not take a stand on all the complexities involved in unbounded dependencies and all the different theories that have been developed for this phenomenon. We can observe, however, that in an approach such as this one that employs a "layer" of marking daughters, we get a nice modularisation of constituent order phenomena and the unbounded dependencies, where the marking functions, in particular the F-DTR, constitutes an interface. Where we find an unbounded dependency in Swedish of the filler-gap variety, whether in a main clause or a subordinate clause, the filler is invariably a foundation.

In Pollard&Sag (1994) the theory of unbounded dependencies crucially involves a Head-Filler Schema. In our model there is no room for the Head-Filler Schema, but I take from it the idea that filler-gap constructions involve a daughter that is identified as a FILLER-DTR. The role of the HEAD-DTR in the Head-Filler Schema, i.e. the place where we find information about the gap, is taken over by the one of the P/A-DTRS where the gap actually occurs. Instead of the schema itself, we introduce a clause type, *FillerCl*, where the filler is token-identical with the foundation, and there is some other daughter, which we for simplicity will take to be a complement, that contains a gap appropriate for the filler:

(56) *FillerCl* ⊂ *Clause;*
Approp(*FillerCl*, M-DTRS) = *_F_,*
Approp(*FillerCl*, FILLER-DTR) = *sign,*
M-DTRS|F-DTR = FILLER-DTR,
FILLER-DTR|SYNSEM|LOCAL =
P/A-DTRS|C-DTR$_i$||SYNSEM| NONLOCAL| TOBIND| SLASH;

Apart from this change of a rule for a clause type, the model of Pollard&Sag (1994) can be taken over as it is. Gaps can be introduced and information about them be inherited according to the same principles. The major difference is that there are fewer nodes in the structure, since every node involved dominate more constituents.

Topicalization where the topicalized constituent is an extracted complement can then be analysed as clauses which are simultaneously filler clauses and inverted clauses with a foundation:

(57) *XFInvCl* ⊂ *FInvCl;*
XFInvCl ⊂ *FillerCl;*
Example: *Den tror jag inte att jag har sett förut.*

Subordinate wh-clauses can be analysed similarly. We make a four way distinction of types with respect to the function of the foundation: it may be subject, complement or adjunct to the matrix verb, or be a filler of a filler-gap construction. The common properties of them are expressed for a sort we term *WhSubCl*:

(58) *WhSubCl* ⊂ *Clause;*
Approp(*WhSubCl*, M-DTRS) = *_F,*

Approp(*WhSubCl*, P/A-DTRS) = *hS__,*
M-DTRS|F-DTR|...|WH = +,
P/A-DTRS|H-DTR$_1$ |...|FIN = +;

For the case when the foundation is the subject, we obtain (59) and when it is a filler, we get (60). The other two cases can be defined similarly.

(59) *WhSubjSubCl* $\subset$ *WhSubCl;*
Approp(*WhSubjSubCl*, M-DTRS) = *FM,*
M-DTRS| F-DTR = P/A-DTRS|S-DTR,
M-DTRS| M-DTR = som;

Example: *(Jag vet) vem som kommer*.

(60) *WhXSubCl* $\subset$ *WhSubCl;*
WhXSubCl $\subset$ *FillerCl;*

Example: *(Han vet) vad vi vill att han ska titta på.*

Relative clauses come in three different varieties. Relative clauses that employ wh-pronouns can be treated similarly to subordinate wh-clauses, so I omit discussion of them. The other two types are the one with the complementizer/pronoun *som*, illustrated in (61), and the bare relative clause in which some other constituent than the subject has been relativized, as illustrated in (62). In relative clauses of the latter kind, a *som* may be optionally inserted.

(61) *(Bilen) som körde om oss.*
(The_car) that overtook us

(62) *(Bilen) han såg.*
(The_car) he saw.

(63) *(Bilen) som han såg.*
(The_car) that he saw.

For relative clauses with a relativized subject we use the sort *RelSubjCl*:

(64) *RelSubjCl* $\subset$ *Clause;*
Approp(*RelSubjCl*, M-DTRS) = *M,*
Approp(*RelSubjCl*, P/A-DTRS) = *h0__,*
M-DTRS| M-DTR = som;

For unmarked relative clauses we use the sort *UnmRelCl*:

(65) *UnmRelCl* $\subset$ *Clause;*
Approp(*UnmRelCl*, M-DTRS) = *unmarked,*
Approp(*UnmRelCl*, P/A-DTRS) = *hS__,*
...| NONLOCAL| INHER|SLASH = ...| NONLOCAL| TOBIND| SLASH;

The last identity of this definition ensures that the unmarked relative clause contains a gap, in the manner proposed by Pollard&Sag (1994: 210). The unmarked relative clauses may then be further specified according to the kind of constituent that has been relativized.

While *som* is optional in a relative clause such as (63) it is nevertheless a marking constituent and requires a special sort, *MRelCl*:

(66) *MRelCl* $\subset$ *Clause;*
Approp(*MRelCl*, M-DTRS) = *M,*
Approp(*MRelCl*, P/A-DTRS) = *hS__,*
M-DTRS| M-DTR = som,
...| NONLOCAL| INHER|SLASH = ...| NONLOCAL| TOBIND| SLASH;

Let us finally consider cases where there is an adverb intervening between the foundation and the marking finite verb, as in (12):

(12) *Han bara skrattade åt oss.*
He just laughed at us
"He did nothing but laugh at us"

There seem to be no reason to regard these adverbs as markers, so the following precedence, repeated from (36) above, threatens to be false:

m-DTR < P/A-DTR$_i$

As indicated above we could see them as forming a constituent with the verb and regard this constituent as the bearer of the marking function. We first observe that some phrases actually can occur in the v-position, e.g. coordinations as in the following examples:

(67) *Hon varken kan eller vill komma.*
She neither can nor wants come.
"She is neither able to, nor wants to, come."

(68) *Han både röker och dricker.*
He both smokes and drinks.
"He is both a smoker and a drinker."

Since the *bara*-class of adverbs can appear in front of almost any type of constituent, it becomes more plausible to analyse them as a sort of adjunct that can stand adjacent to whatever is in their scope semantically and form a constituent with this argument. Still the verb would be the head and the head features of the constituent would be inherited from there.

5. Conclusions

Two questions were posed at the outset: Are positions and schemata really required in a formal description of topological regularities in Swedish? Do we need larger constituent order domains than ones given by the immediate constituents of a clause? In the first part of the paper I pointed out some general drawbacks with positional notions, and in the final part I have sketched an account that relies solely on precedences. One characteristic of the field grammar approach remains, however, the idea that all constituents of a clause belong to the same topological domain. This trait definitely sets it apart from standard HPSG approaches.

Many questions have not been answered in the current sketch. For instance, can the model account for all the word order variation in the nexus field? Can it be applied equally well to Danish, German and other Germanic languages? And are there really no vp:s in Swedish? For now, I'll just leave all of them open. However, I believe the proposal has enough merit to warrant further investigations. In particular, it allows the succinct expression of a number of topological generalizations for Swedish, that other accounts have failed to express. And it does so within the restricted framework of well-typed feature structures.

Another question concerns the usefulness of field schemata. The hypothesis that we can do without them for formal, descriptive purposes, if true, does not really entail that they are useless for analytic, pedagogical or even computational purposes. The point of this paper is merely that precedences are to be preferred as topological primitives. In fact, it is not difficult to synthesize a linear schema from the precedences that have been stated for Swedish in the text, although we have to apply some sound judgement as regards the grouping of constituents into common positions/fields. The schema below is very similar to the one argued for in Heltoft (1986), although less detailed. The only major difference is that I have grouped complementizers with foundations, rather than with the finite verb, since that fits better with the Swedish facts. The nexus field (Nf) as usual is the field for the subject and the adverbs categorized above as +clausal, while the intended contents of the other fields should be obvious.

Marking field		Kernel field			
F/M	m	Nf	Vf	Cf	Af

Referenser

Ahrenberg, L. 1989. A Formal Field Grammar. Research Report LiTH-IDA-R-89-46.

Ahrenberg, L. 1990. A grammar formalism combining field structure and phrase structure. Proceedings of COLING'90, Vol. 2: 1-6.

Ahrenberg, L. 1994. Topological frames in sign-based grammars. In R. Eklundh (ed.) *Proceedings of NODALIDA'93*. Dept of linguistics, Stockholm University.

den Besten, H. 1983. On the Interaction of Root Transformations and Lexical Deletive Rules. In W. Abraham (ed.): *On the formal nature of Westgermania*. Amsterdam, John Benjamins.

Carpenter, B. 1992. *The Logic of Typed Feature Structures*. Cambridge Tracts in Theoretical Computer Science, CUP.

Diderichsen, P. 1966. *Elementær Dansk Grammatik*, 3. Udgave. Copenhagen, Gyldendal.

Gambäck, B. 1997. Processing Swedish Sentences: A Unification-Based Grammar and Some Applications. SICS Dissertation Series 21, Stockholm, Royal Institute of Technology.

Heltoft, L. 1986. Topologi og syntaks. I *NyS* (Nydanske studier & almen kommunikationsteori) 16-17: 105-130, Copenhagen, Akademisk Forlag.

Ekerot, L-J. 1988. Så-konstruktionen i svenskan. Lundastudier i nordisk språkvetenskap A:42. Lund, Lund University Press.

Kasper, R. 1994. Adjuncts in the Mittelfeld. In J. Nerbonne, K. Netter and C. Pollard (eds.): *German in Head-Driven Phrase Structure Grammar.* CSLI Lecture Notes, No. 46: 39-69.

Kasper, R., A. Kathol and C. Pollard, 1995. Linear Precedence Constraints and Reentrancy. In J. Kilbury and R. Wiese (eds.): Integrative Ansätze in der Computerlinguistik. University of Düsseldorf.

Kathol, A. 1995. Linearization-Based German Syntax. Ph.D. thesis, Dept. of Linguistics, University of Ohio.

Kathol, A. and C. Pollard, 1995. On the Left Periphery of German Subordinate Clauses. In Proceedings of the Fourteenth West Coast Conference on Formal Linguistics.

Platzack, C. 1985. A Survey of Generative Analyses of the Verb Second Phenomenon in

Germanic. *Nordic Journal of Linguistics*, Vol. **8** 49-73.

Platzack, C. 1986. Diderichsens positionsschema och generativ transformations-grammatik. I *NyS* (Nydanske studier & almen kommunikationsteori) 16-17: 161-170, Copenhagen, Akademisk Forlag.

Pollard, C. and I. Sag, 1987. *Information-Based Syntax and Semantics, Volume 1: Fundamentals.* CSLI Lecture Notes 13.

Pollard, C. and I. A. Sag, 1994. *Head-Driven Phrase Structure Grammar*. Chicago, The University of Chicago Press.

Reape, M. 1994. Domain Union and Word Order Variation in German. In J. Nerbonne, K. Netter and C. Pollard (eds.): *German in Head-Driven Phrase Structure Grammar.* CSLI Lecture Notes, No. 46: 151-197.

Togeby, O. 1988. Parsing Danish Text in Eurotra. *Nordic Journal of Linguistics*, Vol. **11** (1-2) 175-191.

Underwood, N. L. 1997. A Typed Feature-based Grammar of Danish. *Nordic Journal of Linguistics*, Vol. **20** (1) 31-62.

Van Eynde, F. 1994. Auxiliaries and Verbal Affixes. A Monostratal Cross-Linguistic Analysis. Proefschrift Aggregatie Hoger Onderwijs, Universiteit Leuven, Departement Linguistiek.