The Immediate Dominance Schemata of HPSG

A Deconstruction and a Reconstruction

Frank Van Eynde

KU Leuven, Centrum voor Computerlinguïstiek

Abstract

For modelling constituency, HPSG-III employs seven Immediate Dominance schemata. In this paper they will be evaluated with respect to two parameters: their degree of cross-categorial generalization and their ability to prevent overgeneration. Since the evaluation will reveal a number of problems, I propose an alternative in which the ID schemata are replaced by constraints on phrase types, similar but not identical to a proposal in Sag (1997).

1   From Phrase Structure Rules to Immediate Dominance Schemata

In early transformational grammar constituency was modeled in terms of category specific phrase structure rules, such as

(1)   a. $VP \rightarrow V$ NP PP

         b. $PP \rightarrow P$ NP

This was changed in Chomsky (1970). In order to capture generalizations across categories, such as between a verb and its nominalized counterparts, Chomsky introduced a notation in which categories are replaced by variables with bar levels and in which the PS rules are replaced with X-bar schemata. In Chomsky (1986, p.3) they have the following form:

(2)   a. $X' \rightarrow X^{o}$ X''^{*} [Head–Complement]

         b. $X'' \rightarrow X''^{*}$ X' [Specifier–Head]

   X stands for any category, whether lexical (V, N, A, P) or nonlexical (Infl, Comp), and X''^{*} stands for zero, one or more maximal projections. (2a) covers among others the PS rules in (1).

   A similar development took place in monostratal grammar: while the immediate dominance rules of GPSG[1] are category specific and quite numerous, the immediate dominance schemata of HPSG generalize over categories and are few in number. The 47 lexical ID rules of Gazdar, Klein, Pullum and Sag (1985), for instance, are all replaced by the following general ID schema (Pollard and Sag 1994, 362):

(3) $XP \rightarrow \boxed{1}$, X^{o} [COMPS $\boxed{1}$]   [COMPLEMENTS–HEAD]

---

[1]The main difference between a PS rule and an ID rule is that the former imposes a linear order on the constituents, whereas the latter does not.

XP stands for a single or double bar projection (X' or X'') and $\boxed{1}$ ranges over lists of maximal projections; the latter have to match the subcategorization requirements of the lexical head (X°).

In spite of some obvious parallels, the ID schemata are not notational variants of the X-bar schemata, for first they apply to surface structure,[2] and second they are lexicalist. The latter implies a number of things. First, head daughters must be projections of lexical elements, and not of affixes or empty elements. Second, nonhead daughters must be projections of lexical elements as well. A specifier, for instance, is not a landing site for extracted phrases, which may be left empty, as the specifier of Comp in GB; instead it is the maximal projection of a lexical element (Y''), which is selected by an X' head.[3]

(4) $X'' \rightarrow \boxed{1}\, Y''\, ,\ X'[\text{SPR} <\boxed{1}>]$ [SPECIFIER–HEAD]

Third, the ID schemata explicitly refer to the selection requirements of the lexical elements (the values of COMPS and SPR) and use these requirements to prevent overgeneration.

Because of these differences, the ID schemata do not correspond one-on-one to the X-bar schemata of transformational grammar. Next to the usual two, HPSG-III[4] employs separate schemata for the introduction of adjuncts, subjects, markers and fillers, as well as a composite schema which simultaneously introduces a subject and a complement in inverted clauses.

The aim of this paper is to assess the seven ID schemata of HPSG-III with respect to two parameters: their potential to capture cross-categorial generalizations and their capacity to prevent overgeneration. These will be the topics of sections 2 and 3 respectively. As the discussion will reveal a number of problems, I propose an alternative in section 4.
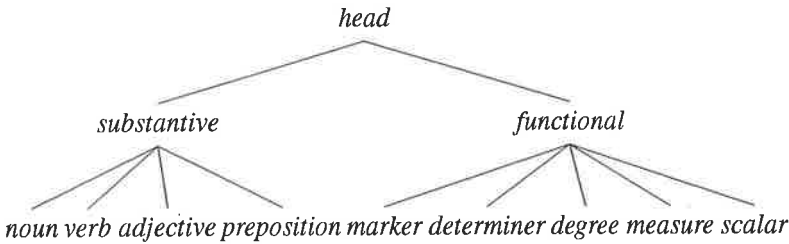
2   Cross-categorial generalization

An ID schema qualifies as cross-categorial if it does not impose any constraints on the parts of speech of its daughters. Since this can only be assessed with respect to some given set of categories, I first present the inventory of parts of speech which is employed in HPSG-III.[5]

---

[2]Chomsky (1986) explicitly states that "the X-bar constraints are satisfied at D-structure, but not at other levels of representation if adjunction has taken place in a derivation." (p.3)

[3]This is a simplified version of the schema. The full version will be given in the next section.

[4]HPSG-III is an acronym for the system described in chapter 9 of Pollard and Sag (1994). By contrast, HPSG-II stands for the system in the first eight chapters and the appendix of Pollard and Sag (1994), and HPSG-I for the version of Pollard and Sag (1987).

[5]*Measure* is assigned to words like *long, years* and *times*, and *Scalar* to the numerals. Since the parts of speech are subsorts of *head*, they are shared between mother and head daughter, as predicted by the Head Feature Principle.

```
                              head
                      _____|_____
                     /                 \
              substantive            functional
            ___/__|__\___          ___/__|__\___
           /   /  |   \   \       /   /  |   \   \
```

noun verb adjective preposition marker determiner degree measure scalar

A characteristic property of this hierarchy is the dichotomy between substantive and functional parts of speech. It resembles Chomsky's distinction between lexical and nonlexical categories, but its role in the grammar is quite different, as will be shown below.

2.1   The ID schemata

Surveying the ID schemata, there is one which clearly does not qualify as cross-categorial, i.e. the schema which models the combination of an extracted phrase (a filler) with the clause from which it has been extracted, as in

(5)   Who [do you think I saw yesterday] ?

The clause *do you think I saw yesterday* contains a gap, and *who* is the element which – conceptually – fills that gap. The general definition of this schema runs as follows (Pollard and Sag 1994, 381):

(6) $X'' \rightarrow Y''[\text{LOC}\ \boxed{1}]$, $V''\ [\text{SLASH}\ \{..., \boxed{1}, ...\}]$ [FILLER–HEAD]

The SLASH value of the head (V'') specifies the type of gap which it contains and this specification has to be matched by the LOC value of the filler (Y''). This schema does not qualify as cross-categorial, since the head is explicitly required to be a verbal projection.

Although they are less conspicuous, similar constraints hold in the schemata for adjuncts, specifiers and markers. Starting with the former, the definition of the schema looks as follows (Pollard and Sag 1994, 384).

(7) $XP \rightarrow Y''[\text{MOD}\ \boxed{1}]$, $\boxed{1}\ XP$ [ADJUNCT–HEAD]

The adjunct is a maximal projection (Y'') and has a MOD(IFIED) value which specifies the constraints which it imposes on its head; attributive APs, for instance, require a nominal head, whereas manner adverbials require a verbal one. The head is a single or double bar projection (X' or X'') and may belong to any part of speech, as long as it matches the MOD value of the adjunct. Also the adjunct does not seem to be subject to any part of speech restrictions, but this appearance is deceiving, for since it is required to have a MOD feature, and since MOD features are only declared for the substantive parts of speech (see section 2.2), it follows that the projections of the functional parts of speech cannot be used as adjuncts, and this implies that the schema does not qualify as cross-categorial.

A similar remark applies to the *head-specifier* schema (Pollard and Sag 1994, 362):

(8) $X'' \rightarrow \boxed{1} \, Y''[\text{SPEC} \, \boxed{2}] \, , \, \boxed{2} \, X'[\text{SPR} <\boxed{1}>]$ [SPECIFIER–HEAD]

The head is a single bar projection (X') which selects its specifier by means of the valence feature SPR, and the specifier is a maximal projection (Y'') which selects its head by means of the SPEC(IFIED) feature. Singular count nouns, for instance, select a determiner via SPR, and determiners select a nominal via SPEC. Also here, the head is not subject to any part of speech constraints, but since the specifier is required to have a SPEC feature, and since only the functional parts of speech have this feature (see section 2.2), it follows that the projections of the substantive parts of speech cannot be used as specifiers.

An even tighter constraint holds for the *head-marker* schema. This schema models e.g. the combination of a clause with its complementizer, as in

(9)   a. believe [that [the sun sets in the west]]

        b. hope [for [Bill to win the race]]

HPSG-III does not provide any explicit definition of this schema, but there is one in HPSG-II which –after transposition into the notation of HPSG-III– would look as follows:

(10) $X'' \rightarrow M^o [\text{SPEC} \, \boxed{1}] \, , \, \boxed{1} \, X''$ [MARKER–HEAD]

The SPEC feature mentions the constraints which the marker imposes on its head: the complementizer *that*, for instance, requires a finite clause, whereas *for* requires an infinitival one. There are no general constraints on the part of speech of the head, but the marker daughter is explicitly required to be a marker.[6] The defining characteristic of these elements is that they cannot have a phrasal projection. This is made explicit in the schema by the requirement that the marker daughter's bar value be zero; in all other schemata the nonhead daughter is required to be phrasal.

Taken together, the *head-adjunct*, *head-specifier* and *head-marker* schemata show a remarkable degree of complementarity, in the sense that they require the nonhead daughter to belong to mutually distinct subsets of the part of speech inventory. As a consequence, none of them qualifies as cross-categorial.

2.2   The part of speech inventory

The (partial) lack of cross-categorial generalization in the ID schemata is largely due to the fact that the part of speech inventory of HPSG-III is a rather awkward

---

[6]Because of its tautological appearance, this requirement does not seem to impose any real constraints, but this is not the case. In order to see its constraining effect, one has to realize that HPSG uses the term *marker* in two different senses. In one sense, it stands for a functional notion, and contrasts with terms like *complement, head* and *adjunct*; in another sense, it stands for a categorial notion, and contrasts with terms like *verb, noun* and *determiner*. So, what the constraint says is that a marker daughter (functionally speaking) has to be a marker (categorially speaking), and this is a rather strong restriction, comparable to the (hypothetical) constraint that a complement daughter must be nominal.

blend of categorial and functional notions: while notions like 'noun' and 'verb' are genuinely categorial, notions like 'measure' and 'degree' are more of a functional nature. Other more specific problems concern the dichotomy between substantive and functional parts of speech and the introduction of a special part of speech for the markers.

Starting with the former, the substantive parts of speech differ from the functional ones in two respects.

$$\textit{substantive}\begin{bmatrix} \text{MOD} & \textit{mod-synsem} \\ \text{PRD} & \textit{boolean} \end{bmatrix} \qquad \textit{functional}\begin{bmatrix} \text{SPEC} & \textit{synsem} \end{bmatrix}$$

One difference has already been mentioned: whereas the substantive parts of speech use the MOD feature to select their head, the functional parts of speech use the SPEC feature. This difference is not merely terminological, since these features take different values: while the one of SPEC is of type *synsem*, the one of MOD is of type *mod-synsem*, which is short for either *synsem* or *none*. The latter is assigned when the projection of a substantive part of speech (V'', N'', P'', A'') is not used as an adjunct, but rather as a complement or a subject. As a consequence, the fact that SPEC cannot have the value *none* implies that (the projection of) a functional part of speech can only be used as a specifier or as a marker.

The other difference concerns the fact that the substantive parts of speech have a PRD (PREDICATIVE) feature, whereas the functional ones have not. The purpose of this feature is to signal whether a word's phrasal projection must, can or cannot be used in predicative positions. Adjectives like *awake* and *alone*, for instance, are [+PRD], since they must be used predicatively, whereas *former* and *wooden* are [–PRD], since they cannot be used predicatively, and adjectives which can but need not be used predicatively, such as *red*, have the generic *boolean* value. The same distinction is made for the verbal, nominal and prepositional projections, but not for the projections of the functional parts of speech, which implies that the distinction has no relevance for them.

When put to the test, though, it turns out that neither of these distinctions corresponds to any real differences. First, the projections of determiners are not only used as specifiers, but also as complements or subjects, and the same holds for the numerals.

(11)　a.　I don't like *this*.

　　　b.　*This* is not worth *much*.

　　　c.　*Much* depends on how they will react.

(12)　a.　I ordered *six* but I only got *four*.

　　　b.　*Two* have left already.

Second, the [+/–PRD] distinction is not only relevant for the substantive parts of speech, but for the determiners and the degree words as well.

(13)　a.　His faults are *many* and his friends are *few*.

        b.   * This is *every/a/the.*

(14)   a.   This is not *enough.*

        b.   * This is not *too/very.*

Apparently, while *many, few* and *enough* are [+/–PRD], *every, a, the, too* and *very* are [–PRD].

This demonstrates that also the functional parts of speech should be allowed to have *none* as their SPEC value and to have a PRD value. But then as a result there is no longer any evidence for making a distinction between substantive and functional parts of speech in the first place.

Another problem with the inventory concerns the markers. Markers are words which cannot have a phrasal projection, such as the complementizers, the comparative words *than* and *as*, the coordinating conjunctions and – in certain languages – the case marking prepositions. Since it is not made clear, though, how the members of this rather heterogeneous class can be identified, there is little consensus about their treatment; the English complementizers, for instance, are presented as the prototypical examples of markers in Pollard and Sag (1994), but in Sag (1997) they are argued to be heads of complementizer phrases, as in GB.

In order to put the notion on somewhat firmer ground, Van Eynde (1997) provides a number of criteria for identifying elements which cannot have a phrasal projection or – using a shorter term – for identifying the minor signs. The main criterion is that minor signs cannot take any of the dependents which other elements of the same paradigm can take. The indefinite article, for instance, belongs to the paradigm of the numerals (cf. its quasi-homonymy with the numeral *one* in many languages) and qualifies as minor since it cannot take any of the dependents which numerals can take, cf. *exactly one car* vs. * *exactly a car*. The second criterion is that minor signs cannot be used as conjuncts, cf. *one or two cars* vs. * *a or two cars*, and the third that they cannot be the sole constituent of an elliptical clause, cf. *I'd rather have two cars than one* vs. * *I'd rather have two cars than a*.

When applied to a broader range of elements, it turns out that the minor signs not only include members of typically functional categories, such as complementizers and determiners, but also of categories which are normally treated as lexical or substantive. The Dutch personal pronouns, for instance, come in two variants, of which one can have a phrasal projection and the other cannot (Van Eynde 1997). Similarly, while most prepositions are heads of PPs, there are some which are better treated as nonhead sisters of verbal projections (Van Eynde 1998). In sum, the distinction between words which can and words which cannot have a phrasal projection turns out to be cross-categorial, and can therefore not be used to justify the postulation of a separate part of speech.

2.3   Summing up

Of the seven ID schemata of HPSG-III there are four which do not qualify as cross-categorial, and in three of those four cases, this is due to the introduction of distinctions in the part of speech inventory which turn out to lack empirical motivation.

3   Preventing overgeneration

Because of their highly abstract nature, the cross-categorial schemata easily lead to overgeneration. In order to avoid this, transformational grammar disposes of a number of devices, such as the context-sensitivity of lexical insertion and the filtering effect of transformations, but neither of these are available in monostratal frameworks. As an alternative, HPSG makes ample use of selection features and, on a more modest scale, of bar-level constraints. The former is typically lexicalist, since selection requirements originate in lexically specified types; the latter is non-lexicalist, since it concerns a property of the schemata themselves. Both methods will be reviewed in this section.

3.1   Selection

HPSG-III employs two kinds of selecting features: those which specify the constraints which a head imposes on its nonhead sisters (the valence features COMPS, SPR and SUBJ, and the nonlocal SLASH), and those which specify the constraints which a nonhead daughter imposes on its head sister (MOD and SPEC).

Because of this multiplicity, it is possible that head and nonhead daughter select and constrain each other. In most cases this is prevented, since the value of MOD may be *none*, but there is one exception: in the *head-specifier* schema the head selects its specifier via SPR, whereas the specifier selects its head via SPEC. From a computational point of view, this mutual selection is not attractive, since it leads to cyclicity in the feature structures, which in turn leads to complications in the definition of the unification operation, see e.g. Pollard and Sag 1987 (p.37). In order to avoid such complications it would be preferable to eliminate either SPEC or SPR. Formally, it does not matter which, but empirically there are two considerations which favour the latter.

First, while the use of the SPEC feature is argued for extensively in Pollard and Sag (1994, pp. 50-1;371-3), both on syntactic and semantic grounds, the use of SPR is simply based on the intuition that projections without specifier are in some sense incomplete. This sense of incompleteness, though, is difficult to square with the fact that specifiers are almost always optional. Adjectival and prepositional projections, for instance, can but never need be specified, verbal projections do not seem to take any specifiers (in HPSG), and in nominal projections, the presence of a specifier is not required for mass nouns, plural nouns and proper nouns; the only exceptions are the singular count ones, and even they do not always need a specifier.

(15)   a. What sort/kind of *book* are you looking for?

    b. He read *book* after *book*, but did not find an answer.

    c. Everybody likes *pilot* George.

Second, for modelling the agreement between head and specifier SPEC is more appropriate than SPR. To show this, let us take the number agreement in [Det–Nominal] combinations.

(16)   a. one/a/two hundred cars

      b. a few/dozen cars

      c. a great/good many cars

If the agreement in the first NP is modeled in terms of SPR, we have to say that the nominal requires a plural specifier, but if *one/a/two hundred* is plural, and if we make the reasonable assumption that *hundred* is the head of this phrase, then *hundred* must be plural too, and since plural heads select plural specifiers (via SPR), this would imply that *one/a hundred* cannot be grammatical, but it is.

By contrast, if the agreement is modeled in terms of SPEC, we have to say that *one/a/two hundred* and its head *hundred* require a plural nominal, but we are not committed to the claim that the specifier itself is plural. Instead, its NUMBER value can be left unspecified, so that *hundred* is not only compatible with specifiers which require a plural head, such as *two*, but also with specifiers which require a singular head, such as *one* and *a*. The same reasoning applies to the other NPs: if it is assumed that the nominal selects a plural specifier, one needs ad-hoc measures to avoid the exclusion of *a few*, *a dozen* and *a great/good many*, but if it is assumed that *few*, *dozen* and *many* require a plural nominal, without being plural themselves, no such problems arise.

In sum, lexical selection is a powerful device for preventing overgeneration, but if one wants to avoid the complications which arise from mutual selection, as in the *head-specifier* schema, one should eliminate either SPR or SPEC, and of these two the former option is preferable for empirical reasons.

3.2 The use of bar-levels

The bar-level device is a typically nonlexicalist way of preventing overgeneration, since the constraints which it imposes on the application of the schemata are wired into the definition of the schemata themselves. For instance, if the bar level of the head is changed, as in the *head-complement* schema, then it is not possible to apply the schema to its own output; conversely, if the bar level of the head is kept constant, as in the *head-adjunct* schema, iterative application is allowed.

Because of this nonlexicalist vein, the bar-level device could be seen as a welcome complement to the otherwise strongly lexicalist HPSG framework, or – conversely – it could be seen as an intruder, as a germ picked up at a GB conference. However, no matter which stand one takes on this issue, the way in which the bar-level device is employed in HPSG-III is rather unsatisfactory, for, on the one hand, most of the constraints which it imposes turn out to be too strong and, on the other hand, the constraints which it could impose are already captured by other means.

The constraining power of the bar-level device is most noticeable in the two schemata which change the bar level of the head, i.e. the *head-complement* schema which raises it from 0 to 1 or 2, and the *head-specifier* schema which raises it from 1 to 2. The former implies that all complements of some given word have to be sisters of that word ($X°$), and not of one of its projections (X' or X''). The latter

implies that a head cannot have more than one specifier. In practice, both of these predictions are problematic.

The requirement that all complements be sisters of their lexical head, for instance, may be plausible for English, but for languages with a relatively free word order, such as Dutch or German, it is often felt to be too strong. See, for instance, Hinrichs and Nakazawa (1994) and Kiss (1995).

The limitation on the number of specifiers leads to problems in the handling of nominal projections with two determiners, such as *his many dreams* and *those few pages*. A possible way out is to claim that one of the determiners is not a specifier: *many* and *few*, for instance, could be argued to be N′-adjuncts. However, since these same words have to be treated as specifiers when they are not preceded by another determiner, as in *many dreams* and *few pages*, this introduces a rather artificial kind of ambiguity. Besides, the problem gets worse, if one also includes the specifiers which belong to other parts of speech, such as *Measure* and *Scalar*.

(17)	a. more than twenty years her senior	[Measure [Det N′]]
	b. once, twice, three times a lady	[Measure* [Det N′]]
	c. his one last hope	[Det [Scal N′]]

To save the single specifier constraint, one could stipulate that the measure phrases are N″-adjuncts and the numeral an N′-adjunct, but since the same measure phrases are clearly specifiers in adjectival projections, such as *more than twenty years older* and *three times deeper*, and since the numeral is clearly a specifier in *one child*, this again amounts to the postulation of an artificial ambiguity.
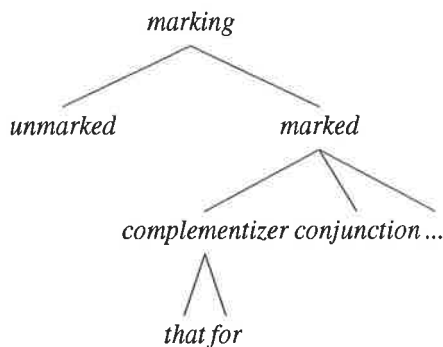
Another infelicitous restriction which the bar-level device imposes concerns the status of the nonhead daughters. With the sole exception of the markers, these are explicitly required to be maximal projections (X″). In other words, complements, subjects, specifiers, fillers and adjuncts have to be double-bar categories. However, since they may consist of a single word, one needs a vacuous application of the *head-complement* schema just in order to get a phrasal projection. Now, since the *head-complement* schema cannot be applied to its own output, there is no danger of iterative vacuous expansion, but from a computational point of view a grammar which does not resort to vacuous expansion is certainly more attractive than one which requires it. Furthermore, the limitation to phrasal signs in nonhead positions might also be empirically inadequate, for if it is accepted that some languages, such as Dutch, have pairs of personal pronouns, of which one may have a phrasal projection, whereas the other one may not, as argued in Van Eynde (1997), then the constraint erroneously prevents the latter from being used as complements.

Let us now turn to some cases in which the bar-level device could have been used to prevent overgeneration, but in which HPSG-III has chosen other means to achieve the same effect. Predictably, these concern combinations which are modeled by ID schemata in which the bar level of the head is kept constant, but which should nonetheless be prevented from applying to their own output. A good example is the *head-subject* schema (Pollard and Sag 1994, 381).

(18) $X'' \rightarrow \boxed{1}\, Y''\ ,\ X''[\text{SUBJ} <\boxed{1}>]$ [SUBJECT–HEAD]

This schema should not be allowed to iterate, since a phrase cannot have more than one subject, but the way in which this is achieved is not by a bar-level change, but rather by the principle which states that valence requirements are discharged as soon as they are satisfied, i.e. the Valence Principle.

In the *head-marker* schema also, iteration should be prevented, since markers cannot be stacked: a clause, for instance, cannot have more than one complementizer. To model this HPSG introduces an ancillary mechanism, i.e. the Marking Device. It involves the addition of a MARKING feature to the AVMs of all signs, both lexical and phrasal ones, and its value is any of

```
                          marking
                         /       \
                        /         \
               unmarked            marked
                                   /    \
                                  /      \
                      complementizer conjunction ...
                            /  \
                           /    \
                        that for
```

In the lexicon, all words are assigned the value *unmarked*, except the markers, which receive a *marked* MARKING value. This value is shared with the mother, as stipulated by the Marking Principle. As a consequence, if the marker requires its head to be unmarked, it cannot be added to a clause which already has a marker, so that stacking is prevented.

In sum, there are at least two cases in which the bar-level device could have been used to prevent unwanted iteration, but in which HPSG-III prefers other devices to do this job.

At this point, one could wonder whether the theory would not be better off without the bar-level mechanism, for the constraints which it imposes turn out to be too strong and the constraints which it could – rightfully – impose are already captured by other means. Interestingly, this scepticism confirms one of the main points of Kornai and Pullum (1990, p. 25): "the interest of the standard X-bar restrictions resides mainly in the notion of 'headedness', bar-levels as such being epiphenomenal and even eliminable.".

3.3 Summing up

HPSG-III employs two devices to control the application of the ID schemata, i.e. selection and bar-level manipulation. The former is in line with the lexicalist vein of the framework, but its ubiquitous use occasionally leads to mutual selection, which had better be avoided, since it introduces cyclicity in the feature structures. The latter is a typically nonlexicalist method and could be complementary to the selection

devices, but in the form in which it is used in HPSG-III it causes more problems than it solves.

4   An alternative

In the two previous sections it has been demonstrated that the seven ID schemata of HPSG-III do not fully exploit the cross-categorial potential of the X-bar notation, that they over-use the method of lexical selection and that they make an awkward use of the bar-level device. Another problem, not mentioned thus far, concerns the arbitrariness of the partitioning. Notice, for instance, that HPSG-I distinguished only 3 schemata (*head-complement*, *head-adjunct* and *head-adjunct-complement*). HPSG-II then dropped the latter, and added *head-filler* and *head-marker*, whereupon HPSG-III further added *head-subject*, *head-specifier* and *head-subject-complement*. Since then, various authors have proposed to eliminate the latter, as well as the *head-marker* schema, and in Sag and Wasow (n.d.) the same fate is reserved for the *head-subject* schema, which is reanalysed as an instance of *head-specifier*. In sum, the inventory of schemata is anything but stable.

What I will do in this section is develop the outline of an alternative, which stays close to the notation and the substance of HPSG, but which solves the problems noted above.

As a first step, I will simplify the part of speech inventory. More specifically, I will only employ the standard set of substantive parts of speech, augmenting it with the adverbs.

head

verb noun adjective adverb preposition

The members of HPSG's functional categories are assigned to one of these five. The numerals and the determiners, for instance, are classified as adjectives, the degree words and some of the measure words (*long, right, just*) as adverbs, and the other measure words (*day, times, years*) as nouns. The words without phrasal projection (HPSG's markers) are assigned to one of the above as well. The Dutch minor pronouns, for instance, are classified as nouns, the articles as minor adjectives and the complementizers as either prepositions or adverbs; see Van Eynde (1997) for a treatment which is compatible with the present proposal.

As there is no longer a distinction between substantive and functional parts of speech, there is no need to distinguish between SPEC and MOD. Instead, I will assume that all words have a SELECT feature, whose value is either *none* or *synsem*.

$$_{head}\left[\text{SELECT}\quad \textit{synsem-or-none}\right]$$

Next, I return to the ID schemata themselves. For a start, in order to integrate them better in the formalism, I will reformulate them as constraints on headed

phrases, following a proposal of Sag (1997). Then, instead of introducing them more or less as the need arises, I will base the partitioning of the headed phrases on a single principle, i.e. the one of syntactic selection. Since a headed phrase consists of a head and a nonhead daughter, there are four logical possibilities:

1. The head selects the non-head and the non-head selects the head.

2. The head selects the non-head and the non-head does not select the head.

3. The head does not select the non-head and the non-head selects the head.

4. The head does not select the non-head and the non-head does not select the head.

The first of these is the one of mutual selection and will not be employed for the reasons mentioned in the previous section. The second and the third are taken on board and will be called respectively *head-dependent* and *head-functor*. The fourth does not receive any attention in HPSG-III but since it may be useful for modelling certain types of adjunction and apposition, I will reserve a special type for it, i.e. *head-independent*. The resulting inventory looks as follows.
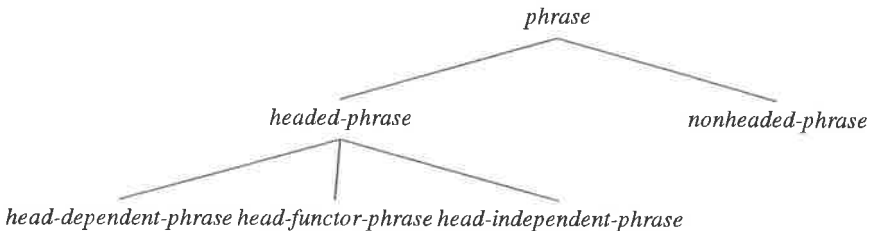
All headed phrases have one and only one head daughter plus a (possibly empty) list of dependent daughters:

$$\begin{bmatrix} \text{HEAD-DTR} & sign \\ \text{DEP-DTRS} & list\left(sign\right) \end{bmatrix}_{headed\text{-}phrase}$$

As usual, they are constrained by the Head Feature Principle and the Valence Principle.

Nothing specific has to be added for the *head-dependent* phrases, except for the requirement that the list of DEP-DTRS be non-empty. This phrase type covers among others the combinations which HPSG-III models in terms of the *head-complement*, *head-subject* and *head-subject-complement* schemata.[7] Since the dependent daughters are not required to be phrasal, there is no need for vacuous expansion in the case of one-word complements or subjects, and since the head daughter can be a phrase as well as a word, there is no requirement that complements only

---

[7] If it is found convenient, the phrases of type *head-dependent* can be partitioned along similar lines.

combine with lexical heads, nor that subjects only combine with phrasal heads; the former facilitates the treatment of Dutch and German VPs, the latter avoids vacuous expansion in the case of one-word VPs, as in *this bucket leaks* and *I agree*.

The *head-functor* phrases have a functor daughter and are required to have an empty list of dependent daughters.

$$\text{head-functor-phrase}\begin{bmatrix} \text{DEP-DTRS} & \langle \ \rangle \\ \text{FUN-DTR} & sign \end{bmatrix}$$

Since the functor is not required to be phrasal, there is no need for vacuous expansion in the case of one-word functors. The principle which models the selection in *head-functor* phrases is the Functor Principle.[8]

$$\text{head-functor-phrase} \ \rightarrow$$
$$\begin{bmatrix} \text{HEAD-DTR} \,|\, \text{SYNSEM} & \boxed{1} \\ \text{FUN-DTR} \,|\, \text{SYNSEM} \,|\, \text{LOC} \,|\, \text{CAT} \,|\, \text{HEAD} \,|\, \text{SELECT} & \boxed{1} \end{bmatrix}$$

In contrast to the *head-dependent* combination, which can only apply as many times as there are elements on the valence lists of the head (maximally), the *head-functor* combination can be applied iteratively. To handle cases in which this should be blocked, I make use of a generalized version of the Marking Principle.[9]

$$\text{head-functor-phrase} \ \rightarrow$$
$$\begin{bmatrix} \text{SYNSEM} \,|\, \text{LOC} \,|\, \text{CAT} \,|\, \text{MARKING} & \boxed{1} \\ \text{FUN-DTR} \,|\, \text{SYNSEM} \,|\, \text{LOC} \,|\, \text{CAT} \,|\, \text{MARKING} & \boxed{1} \ marked \end{bmatrix}$$

In this way, the number and the kinds of functors which a projection is allowed to take can be monitored by the interplay of the SELECT and MARKING values of the functors. Nominal projections, for instance, can be allowed to take both a quantifying determiner and a demonstrative without this leading to overgeneration. For a full-fledged treatment of the Italian NP specifiers along these lines, see Allegranza (1998).

As for the coverage of this type of phrase, I assume that it covers the *head-marker* and *head-specifier* combinations of HPSG-III, as well as some – but not all – of the *head-adjunct* combinations. Adjuncts which can be extracted, for instance, do not qualify as functors, since functors share the property of heads that they cannot be extracted. By this criterion, the prenominal adjectives and the negation marker *not* qualify as functors, whereas the temporal adverbials do not.

(19)   a.   * Red, I saw the ball.

       b.   * Not he said he has seen us.

---

[8]This principle resembles the Spec Principle of Pollard and Sag (1994) but has a wider application, since it covers all cases of selection by non-heads.

[9]This version is generalized, because it applies to all functors and not only to the markers. In the other types of headed phrases it is the head daughter which shares its MARKING value with the mother.

(20)   a.   How long ago do you believe that Bill left ?

         b.   Yesterday I believe Kim left.

Another phenomenon which may shed some light on the status of some given adjunct is the one of word order. In languages with a relatively rigid word order, such as English, there is a strong tendency for the functor to precede its head.[10] As a consequence, if an adjunct has to follow its head, it is less likely to be a functor. Interestingly, this confirms the extraction data, for while *not* and the prenominal adjectives precede their head, the temporal adjuncts follow their head (when they are not extracted).

The *head-independent* phrases are the nonlexicalist residue of the headed combinations. Since there is no lexically based selection mechanism which controls the application of this combination, it will have to include more specific information in the general phrase type, or – more realistically – in its subtypes, probably along the lines suggested in Sag (1997), which provides a detailed example of how multiple inheritance techniques can be applied to the typology of phrases. What kinds of subtypes ought to be distinguished and what they cover in terms of the ID schemata of HPSG-III is left for future work. Let it just be said that it provides a natural way to handle apposition and – perhaps – those types of adjuncts which do not qualify as functors.[11]

5   Conclusion

In order to model the constituency of phrases, HPSG makes use of ID schemata. In their current formulation (HPSG-III), though, the schemata do not fully exploit the potential for cross-categorial generalization, and partly rely on some unattractive devices to prevent overgeneration: the multiplicity of selection devices, for instance, sometimes leads to mutual selection and hence to cyclicity in the feature structures, and the use of the bar-level mechanism causes more problems than it solves (vacuous expansion, spurious constraints). Furthermore, the haphazard way in which the inventory of schemata grows or shrinks reveals the lack of a unifying principle for their differentiation.

An alternative has been proposed in which these drawbacks are avoided. First, the partitioning into different types of headed phrases is based on a single principle and therefore stable. Second, the three types of combination each qualify as cross-categorial, since none of the phrase types imposes any constraints on the parts of speech of the daughters. Third, there is no need for vacuous projection since dependents, heads and functors are all allowed to be words. Fourth, overgeneration is

---

[10] As usual with LP constraints, there are exceptions. The adverb *enough*, for instance, is a degree word, just like *very, so, as, too, more* and *less*, but while the latter all precede their adjectival or adverbial head, *enough* has to follow it.

[11] Another topic which is left for future work concerns the integration of the *head-filler* combination in the hierarchy. Since fillers are selected by their head in HPSG-III, it would seem logical to treat it as a subtype of *head-dependent*, but before taking this move, one should at least consider the alternative of treating them as functors which select a clause with a non-empty SLASH value, for especially in the case of *wh*-phrases this treatment has some initial plausibility.

prevented by lexicalist means in both the *head-dependent* and *head-functor* combinations, but there are no cases of mutual selection, and while some nonlexicalist means will be needed to constrain the *head-independent* combination, no use has been made of the rather rigid bar-level mechanism.

References

Allegranza, V.(1998), Determiners as Functors: NP Structure in Italian, *in* L. Dini and S. Balari (eds), *Romance in HPSG*, CSLI/Cambridge University Press, Stanford/Cambridge, 55–108.

Chomsky, N.(1970), Remarks on nominalization, *in* R. Jacobs and P. Rosenbaum (eds), *Readings in English transformational grammar*, Blaisdell, Waltham (Mass.), 184–221.

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

Gazdar, G., Klein, E., Pullum, G. and Sag, I.(1985), *Generalized Phrase Structure Grammar*, Basil Blackwell, Oxford.

Hinrichs, E. and Nakazawa, T.(1994), Linearizing AUXs in German verbal complexes, *in* J. Nerbonne, K. Netter and C. Pollard (eds), *German in Head-Driven Phrase Structure Grammar*, CSLI, 11–37.

Kiss, T.(1995), *Infinitive Komplementation: Neue Studien zum deutschen Verbum infinitum*, Linguistische Arbeiten, Niemeyer, Tübingen.

Kornai, A. and Pullum, G.(1990), The X-bar theory of phrase structure, *Language* 66(1), 24–50.

Pollard, C. and Sag, I.(1987), *Information-based Syntax and Semantics*, Vol. 13 of *CSLI Lecture Notes*, CSLI, Stanford.

Pollard, C. and Sag, I.(1994), *Head-driven Phrase Structure Grammar*, CSLI and University of Chicago Press, Stanford/Chicago.

Sag, I.(1997), English Relative Clause Constructions, *Journal of Linguistics* 33, 431-434.

Sag, I. and Wasow, T.(n.d.), Syntactic Theory: A Formal Introduction, Partial draft of September, 1997.

Van Eynde, F.(1997), On the notion 'Minor category', *in* J. Landsbergen, J. Odijk, K. van Deemter and G. Veldhuijzen van Zanten (eds), *CLIN VII. Papers from the Seventh CLIN meeting*, IPO, University of Eindhoven, 191–206.

Van Eynde, F.(1998), Functional projections and Dutch complementizers, *in* T. Kiss and D. Meurers (eds), *Current Topics in Constraint-based Theories of Germanic Syntax*, Saarbrücken. Workshop Notes ESSLLI 1998, 149-164.