AN INFORMAL INTRODUCTION TO HEAD DRIVEN PHRASE STRUCTURE GRAMMAR (HPSG)[1]

Wolf Paprotté
University of Münster

1. BACKGROUND

Head Driven Phrase Structure Grammar (HPSG) is a recent grammatical theory which incorporates Carl Pollard's work on Head Grammars (Pollard 1984) and eclectically draws on several grammatical theories such as GB, (Government and Binding), GPSG (Generalized Phrase Structure Grammar), CG and CUG (Categorial Unification Grammar) and kFG (Lexical Functional Grammar). It also builds on notions of situation semantics (Barwise; Perry (1983). Since its first presentation in Pollard/Sag (1987), HPSG has undergone a number of revisions (e.g. Pollard 1988; Pollard/Sag forthcoming) and is still a theory very much under development.

The linguistic objects which HPSG studies (i.e. words, phrases, sentences, texts) are associated with linguistic information represented in feature structures or attribute value matrices (AVMs). AVMs are used to describe (or model) linguistic information of all kinds: lexical, syntactic, semantic, pragmatic, phonological, even phrase-structural phenomena are described in AVM notation.

HPSG-PSG belongs into the group of unification grammars which use complex feature structures as partial descriptions of linguistic objects. The information contained in these descriptions is merged through a monotonic operation called unification. Examples of other unification grammar models are:

* Lexical Functional Grammar (LFG) (Bresnan 1982);
* Generalized Phrase Structure Grammar (GPSG) (Gazdar~Klein/Pullum/Sag 1985);
* Functional Unification Grammar (FUG) (KAY 1984; 1985);
* Categorial Unification Grammar (CUG) (Ades/Steedman 1982; Karttunen 1986; Uszkoreit 1986).
* Definite Clause Grammars (DCG) (Pereira/Warren 1980).

This collection of "spiritually and formally" related formalisms seems to point towards a change of paradigm in linguistics with a number of design choices which set these formalisms apart from the generative grammars that have evolved since Chomsky (1957). Recently, however, a move towards GB (Government and Binding; Chomsky 1981) is apparent. The shared basic convictions of the new paradigm require of a grammar that it is

* monostratal, surface based and directly describe the actual surface order of the string of verbal signs in a sentence. The grammar does not contain rules such as transformations which allow to derive or modify surface trees;
* informational and associate with the string information from specific informational domains;
* lexicalist and account for the valence or subcategorization potential of each linguistic sign by a lexicon containing complex information on the contextual features of each sign. Thus, the lexicon itself specifies well-formedness conditions on phrase structure trees. The notion of head2 of a phrase is a central notion in a lexicon which can be viewed as a word grammar;
* declarative and define the association between strings and informational elements in terms of what associations are permissible, not how they are computed; complex feature based; attribute value matrices (AVMs) are the means of linguistic description.

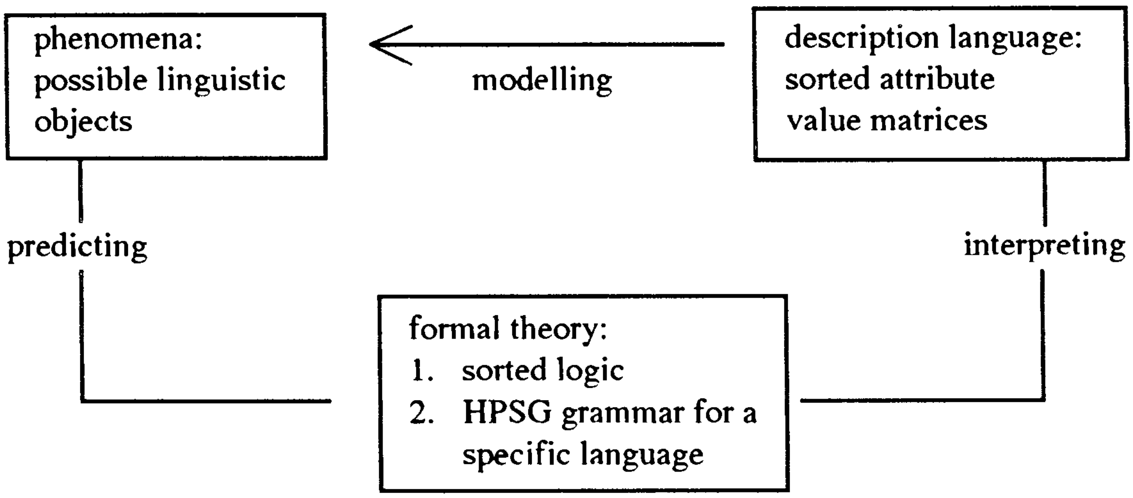
This spirit of non-derivational, lexicalist and feature oriented grammatical theory is also shared by Stanley Starosta's (1988) Lexicase Grammar, by Richard Hudson's Wordgrammar (1984) and Maurice Gross's work thus supporting the notion of an impending change of paradigm away from Chomsky's generative framework.

2. Pollard & Sag on the Nature of a Linguistic Theory

Pollard / Sag (forthcoming) assume that a linguistic theory is a scientific theory in which an empirical domain of phenomena (possible linguistic objects) is modelled by logical or mathematical structures: by attribute value matrices (AVMs). AVMs are a formal language to talk about or to describe such linguistic objects as words, phrases, sentences or texts which people produce or comprehend.

A linguistic theory is used to talk about or to interpret the modelling structures. An informal theory talks about the AVM descriptive model, using natural language and gives a more or less precise specification of which feature structures are considered permissible. Linguistic entities which correspond to the admissible feature structures are assumed to constitute the predictions of the theory.

A formal theory is a formal language, a mathematical model for specifying constraints on feature structures and for interpreting the descriptive devices (Moshier/Rounds 1987; Johnson 1987) or a computer implementation of a linguistic theory



The descriptive language of AVMs has its own well-formedness conditions; only certain descriptions are licensed. A linguistic string is grammatical iff it can be described by a description which is licensed by the theory. The language also admits partial descriptions, and descriptions may be ordered according to their degree of informativity (see below).

3. Linguistic Phenomena: the Objects of Descriptions

HPSG assumes that linguistics is not concerned with the individual linguistic events or utterance tokens but rather with the the system of signs, i.e. with a system of linguistic types corresponding to the knowledge shared among the members of a linguistic community. The ontological status of the system has been hotly debated between mentalists (conceptualists) and realists; there is no need for me here to attempt a solution to this foundational problem of linguistics. I will simply assume that language can be viewed as a mental system shared by members of a community in which an exchange of information is effected via the use of observable linguistic sign tokens.

The linguistic signs include not only words but also sentences, subsentential phrases and multi-sentence discourses. Signs can be conceived of as structured complexes of phonological, syntactic, semantic pragmatic and phrase structural information. That is, they have acoustic or visual form-characteristics specified by their phonological and graphemic properties; they mean something to the members of the community of sign users and therefore have semantic properties; they bear information as to part of speech membership and combinatory potential (valence) and thus have also syntactic properties.

Finally, signs have properties pertaining to their use by speakers on particular occasions, under specific conditions and for specific goals. These are the pragmatic properties of signs, corresponding to the CONTEXT attribute of HPSG.

4. The Language of Description: Attribute Value Matrices and their Formal Properties

Unification-based formalisms use a descriptive system of features and their values as their informational domain. The elements of this domain are called attribute value matrices. Intuitively, an AVM is just an information-bearing object that partially describes or represents another thing by specifying values for various properties of the described thing. Mathematically, an AVM is a connected, acyclic graph; a partial function from a domain of possible attributes/features to a domain of values. The values may be atomic or may themselves be structured and take other AVMs as their values. The common notation for such AVMs is (1):

(1)

	MANUFACTURER	Saab
	TYPE	9000
	CUBIC CAPACITY	1985 ccm
	POWER	175 hp
	UNLADEN WEIGHT	1345 kg
	MAX LADEN WEIGHT	1810 kg
	SPEED	215 km/h
car	COLOUR	blue

This AVM (1) is a partial description of an object of type car; it contains a number of attributes (MANUFACTURER; POWER; SPEED, ...) and gives a value for each of these attributes (Saab, 9000, 1985 ccm, ...). As can easily be seen, for various reasons (2) is no permissible AVM for an object of type car.

(2)

$$\text{car}\begin{bmatrix} \text{MANUFACTURER} & \text{Balmain} \\ \text{TYPE} & \text{suit} \\ \text{CUBIC CAPACITY} & \text{1985 hp} \\ \text{POWER} & \text{blue} \\ \text{UNLADEN WEIGHT} & \text{1345 sec} \\ \text{MAX LADEN WEIGHT} & \text{1810 m} \end{bmatrix}$$

A fundamental property of feature structures is their potential for hierarchiality: the value of some attribute in a feature structure may itself be specified by another feature structure rather than by atomic values: (3) is a hierarchically structured AVM of type *3rd-singular-nominal-phrase* one of whose attributes takes another AVM as a value.

(3)

$$\text{NP3rdsg}\begin{bmatrix} \text{CAT} & \text{NP} \\ \text{AGREEMENT} & \begin{bmatrix} \text{NUMBER SING} \\ \text{PERSON 3RD} \end{bmatrix} \end{bmatrix}$$

With embedded feature structures, it is convenient to have the notion *path of attributes.* A path is just a finite sequence of attributes, and the notion value of an attribute may be generalized to the notion of value corresponding to a path.

Feature structures can be re-entrant. A re-entrant AVM is one in which two feature structures *share* one common value. It is important to distinguish the case in which two features have similar values, i.e. they share a value type (4), from the case in which two feature structures share a value token (5):

(4)

$$\begin{bmatrix} \text{f [h: a]} \\ \text{g [H: A]} \end{bmatrix}$$

(5)

$$\begin{bmatrix} \text{f } \boxed{1} \text{ [h: a]} \\ \text{g } \boxed{1} \end{bmatrix}$$

A shared value is notated by using multiple coindexing boxes on the values; the feature that shares the coindexed value simply takes the coindexing box as its value.

5. Subsumption and Unification

Feature structures vary as to their informativity. Of two feature structures describing the same object, one may be more general and therefore contain less information. Informally speaking, the subsumption relation between feature structures is the following:

If a feature structure β contains at least all the information contained in a feature structure α (β is at least as informative as α), we say " α subsumes β" or "β extends α".

To give an illustration (vgl. Shieber 1986):

(6)	sign[ ]	subsumes	NP[CAT NP]
	NP[CAT NP]	subsumes	NPsg [CAT NP, AGREEMENT [NUMBER SG]]
	NPsg [CAT NP, AGREEMENT [NUMBER SG]]	subsumes	NP3rdsg [AGREEMENT [NUMBER SG, PERSON 3RD]]
	Np3rds [CAT NP, AGREEMENT [NUMBER SG, PERSONS 3RD]]	subsumes	NP3rdsgsu [CAT NP, AGREEMENT |1| [NUMBERS SG, PERSON 3RD], SUBJECT |1|]

Properties of subsumption: The subsumption relation is a weak partial ordering; therefore, the relation is

transitive: For all α, β, Γ [(α subsumes β & β subsumes Γ) =>
a subsumes Γ].

anti-symmetric: For all α, β [(α subsumes β & β subsumes α) => α = β].

reflexive: For all α, α subsumes α.

The most particular and informative feature structure that extends everything else is called Top (written T). Top contains too much and, therefore, conflicting information. The most general feature structure that subsumes everything else tells us nothing, and, therefore, contains no information; it is called Bottom (written ⊥). By adding to the descriptive language of AVMs these two elements Top and Bottom, such that for all α, *1* subsumes α, and a subsumes β, and for all β, β subsumes T, the domain of descriptions forms a lattice.

With respect to the above example (6), the following subsumption relations hold:

(7) ⊥ subsumes *sign* subsumes *NP* subsumes *NPsg* subsumes *NP3rdsg* subsumes *NP3rdsgsu* subsumes T.

Intuitively, unification is the operation that merges the information of two or more AVMs. Given two descriptions β, Γ of the same element, the most general description α is the unification of the descriptions β and Γ iff α is subsumed by both β and Γ. The unification of a set of AVMs is defined as the least upper bound of the subsumption relation on that set; i.e. the least AVM which subsumes each AVM in the set. (Notation: α U β).

The operation of unification which gives its name to a whole class of grammar formalisms is the single most important operation employed in HPSG. some examples:

(8)

$$\text{[CAT NP] U [AGREEMENT INUM SING]]} = \begin{bmatrix} \text{CAT NP} \\ \text{AGREEMENT [NUMSING]} \end{bmatrix}$$

(9)

$$\begin{bmatrix} \text{CAT NP} \\ \text{AGREEMENT [NUM SING]} \end{bmatrix} \text{ U } \begin{bmatrix} \text{CAT NP} \\ \text{[AGREEMENT [PERSON 3RD]} \end{bmatrix}$$

$$= \begin{bmatrix} \text{CAT NP} & \\ \text{AGREEMENT} & \begin{bmatrix} \text{NUM SING} \\ \text{PERS 3RD} \end{bmatrix} \end{bmatrix}$$

(10)

$$\begin{bmatrix} \text{SUBJECT [ ]} \\ \text{CATEGORY |1| [VFORM FIN]} \\ \text{PREDICATE [CATEGORY |1|]} \end{bmatrix} \text{ U } \begin{bmatrix} & \text{OBJECT [ ]} \\ & \text{CATEGORY |2|} \\ \text{PREDICATE} & \text{VERB [CATEGORY |1|] [AUX]} \end{bmatrix}$$

$$= \begin{bmatrix} \text{SUBJECT [ ]} & \\ \text{CATEGORY |1|} & \begin{bmatrix} \text{VFORM FIN} \\ \text{AUX +} \end{bmatrix} \\ \text{PREDICATE} & \begin{bmatrix} \text{OBJECT [ ]} \\ \text{CATEGORY |2|} \\ \text{VERB [CATEGORY |1|]} \end{bmatrix} \end{bmatrix}$$

If two descriptions are inconsistent, as in (11),

(11) [AGR [PERS 3RD]] U [AGR [PERS 2ND]] = T
(failure),

e. g. if there is no object satisfying them both, then the unification fails and is T (Top).

Some properties of unification:

(12)	idempotency	A U A = A
	commutativity	A U B = B U A
	associativity	(A U B) U C = A U (B U C)
	Top	T U A = T
	Bottom	⊥ U A = A
	atoms (only valid for atomic AVMs)	
		a1 U a2 = fails iff a1 =/= a2.

6. Augmentations of the Descriptive Language of Feature Structures

The above system turns out to be insufficient for modelling certain aspects of linguistic information. For that reason, some elaborations are necessary. The first augmentation concerns the notion of feature structure sort. Feature structures are geared to the description of specific objects for which only specific attributes and values make sense. Intuitively, a sort tells us what type of object the feature structure in question models and what kinds of attributes make sense for this type of object (Pollard & Sag 1987). For example, appropriate attributes for the sort sign are PHON, SYnSEM and QSTORE, for the sort *syntactic category* HEAD and SUBCAT, and appropriate attributes for the sort *head* are the head features MAJ, VFORM, NFORM, PFORM, AUX etc. In general, the value of a given attribute is required to be of a certain sort.

Lexical sign (word) and *phrasal sign* are subtypes of *sign* such that *sign* is the supertype of *lexical sign* and *phrase*. As Pollard notes, by using a sorted feature structure, the subtype relation on types of objects can be modelled by imposing a partial subsumption ordering on the set of feature structure sorts, a meet-semi lattice. Any attribute that is appropriate for a given sort should also be appropriate for any subsort of that sort; and, if a sort imposes restrictions on its attributes, any subsort will impose at least those restrictions on its attributes. Thus, a feature structure inherits all the attributes and corresponding type restrictions on its values from all of its supertypes. However, any sort or subsort may introduce attributes of its own, which it does not inherit from its supersort but which it imposes. Thus the sort *sign* has the attributes PHON, SYNSEM, QSTORE which are inherited by its subsorts *lexical sign* and *phrasal sign* but *phrase* has the additional attribute DAUGHTER which is typed to yet another sort *constituent structure*. Notation: the sort name is typically subscripted at the right lower corner of an AVM.

As there is no restriction on sorts having two or more incompatible supersorts, multiple inheritance is possible for subsorts. (Multiple inheritance is used in the analysis of the organisation of the lexicon.)

A second extension of the description language concerns the use of the logical connectives ¬ (negation), v (disjunction), and => for implication; unification U amounts to the logical conjunction of feature structures. The syntax of the sorted feature structure logic (Rounds Kasper 1986) is then as given in (12), (cf. Pollard 1988)

(13)
1. α *is* a description (a ε Sorts);
2. Bottom and Top are descriptions;
3. if α and β are descriptions, then $\alpha \vee \beta$ *is* a description
4. if α and β are descriptions, then $\alpha \Rightarrow \beta$ is a description
5. $\neg \alpha$ is a description.

The semantics boil down to this: An object satisfies $\alpha \vee \beta$ if it satisfies either α or β, or both. An object satisfies $\alpha \Rightarrow \beta$ if it satisfies α, then it also satisfies β. An object satisfies $\neg \alpha$ if it does not satisfy α. An object satisfies Bottom always and Top never.

Pollard & Sag (1987) give a third extension of their descriptive language of feature structures:

6. if $\alpha_1 ..., \alpha n$ are descriptions, the set $\{\alpha_1 ..., \alpha_n\}$ and the list $<\alpha_1 ..., \alpha_n>$ are also descriptions.

A set description describes a set of objects. It is assumed that every object in the set of linguistic objects under description is described by at least one description of the set of descriptions; and each description gives information about only one member of the set under description.

A list description describes a list (a totally ordered set) of objects. The SUBCAT feature, which gives information about the subcategori~ation potential or valence of a sign, is typed to the sort *list description* in HPSG; each member of the list will be a feature structure of sort sign corresponding to one of the unfilled grammatical functions required by the category in question. Thus the verbforms *walks* and *sees* have the following SUBCAT lists (14)

$$
\begin{array}{llll}
(14) & \textit{walks} < \mathrm{NP} \begin{bmatrix} \mathrm{CASE\ NOM} > \\ \mathrm{PERS\ 3RD} \end{bmatrix} & & \\
 & \textit{sees} < \mathrm{NP} \ [\mathrm{CASE\ ACC}], & \mathrm{NP} \begin{bmatrix} \mathrm{CASE\ NOM} > \\ \mathrm{PERS\ 3RD} \end{bmatrix} &
\end{array}
$$

The order of elements on the SUBCAT list is significant; in HPSG it corresponds to the notion of obliqueness in traditional grammar; the following ordering of grammatical functions is assumed (cf. the accessibility hierarchy presented in Keenan & Comrie 1977) (15):

(15) SUBJECT DIRECT OBJECT INDIRECT OBJECT OBLIQUES GENITIVES OBJECTS OF COMPARISON.

The last extension to the description language concerns the use of functionally dependent values. These values depend on other values and the dependency is constrained by some function. Thus the value of some feature may be a list that is obtained by a function *append* which joins together two other lists (16)

(16) *append* (<[NP NOM], PP>, < [NP ACC]>)
= <[NP NOM], PP, [NP ACC]>

or by the set function *union* which takes as arguments two set descriptions and returns as value a set which is the union of the two arguments (Pollard & Sag 1987).

7. The Organisation of an HPSG Grammar

The descriptive language for everything contained in the grammar is the language of AVMs; rules, principles, categories etc. are all specifiable in terms of AVMs.

Pollard & Sag postulate that a grammar consists of

- UG (universal grammar) = P_1 & ... & P_n; a (small) number of universal principles such as the subcategorization principle, the head feature principle, the binding inheritance principle, the gap introduction principle and the control agreement principle (all expressed as AVMs);

- LgsP = P_n, P_{n1} & . . . & P_{n+m} ; a (small) number of language specific principles;
- RC = R_1, R_2 & ... & R_n; a rule component consisting of a (small) number of grammar rules (ID-rules (immediate dominance) and LP-statements (linear precedence);
- LEX = L_1 & ... & L_n & LR_1 & ... LR_n; a (complex) lexicon consisting of lexical entries of different sorts and lexical redundancy rules.

8. Information Types / Levels in HPSG

In HPSG it is assumed that linguistic signs normally possess three attributes PHON, SYNSEM, QSTORE; phrasal signs (i.e. all signs which are not lexical signs) introduce a fourth attribute of their own, DAUGHTERS (DTRS) which encodes constituent structure information (17):

(17)

[PHON <a list of phonemes, graphemes>
SYNSEM [an object of type synsem]
QSTORE {a set of objects of type quantifier}]
sign

[PHON <a list of phonemes>
SYNSEM [an object of type synsem]
QSTORE [a set of objects of type quantifierl
DAUGHTERS [an object of type constituentstructure]]
phrasal sign

The **PHON** attribute takes as its value a string of phonemes or graphemes as some kind of feature representation of the sign's phonological (orthographic) structure and ultimately of its acoustic properties.

The SYNSEM attribute represents a complex of information which was formerly (Pollard & Sag 1987) distributed among the two attributes SYNTAX and SEMANTICS but has been reorganised on the assumption that this attribute forms a natural class. SYNSEM is a natural class because other signs utilize this bundle of information in their SUBCAT attribute, and because SYNSEM information is shared by the complement subject and the controller in equi and raising constructions (Pollard & Sag, forthcoming).

An AVM of type SYNSEM has the values LOCAL (LOC) and NONLOCAL (NONLOC); NONLOC attributes are needed for the description of of unbounded dependency phenomena (topicalization, Wh-questions, relative clauses, it -clefts, pseudoclefts etc.). The NONLOC values are SLASH, REL, QUE.

The attribute LOC has the values CATEGORY (CAT), CONTENT (CONT) and CONTEXT (CONX).

8.1. *The syntactic categories, the CAT value*

The path SYNSEM|LOC|CAT includes a specification of the syntactic category/ part of speech specification of the sign in question, a specification of what grammatical arguments it takes, and whether or not it is a lexical sign; it takes the values **HEAD, SUBCAT** and **LEX.**

The value of the HEAD attribute is a feature structure containing information that is structure-shared with its phrasal projections. (A principle of universal grammar, the Head Feature Principle, requires that the HEAD value of any sign is always structure shared with that of its phrasal projections.) The HEAD information is of the sorts *parts of speech;* roughly speaking, the HEAD value of a sign is its part of speech (18); appropriate values for HEAD are given below (19).

SUBCAT information is essentially the sign's valence, given as a list of SYNSEM values of those signs that correspond to a complement of the described sign. SUBCAT thus specifies with which signs a particular sign must combine in order to become *saturated.* (The Subcategorization Principle of Universal Grammar requires that the SYNSEM value of a complement is unified with the corresponding specification on the SUBCAT list of a sign which selects for the complement.) A constituent whose SUBCAT attribute has the empty list as its value is saturated. The SUBCAT feature thus also encodes the degree of saturation of a constituent. HEAD and SUBCAT values taken together permit an efficient cross categorial formulation of grammar rules and function as an HPSG specific substitute for conventional phrase structure rules.

It may be in order to note here that HPSG has taken a step towards becoming a minimalist syntactic theory. Thus, although it still adheres to the fundamental assumption that natural languages should be described in terms of constituent structure, at least where possible, the redistribution of information on matters of constituency among the rules for immediate dominance (ID-Rules) and the rules for linear precedence (LP-Rules), points to a the fact that the primary data of syntax are the temporally or linearly ordered string of words - (rather than constituent structure trees); and that everything else in syntactic description is part of the theoretical enterprise of the linguist. Constituent structure in HPSG comes in via the valency characteristics of head signs which in a way set bounding conditions for phrases, but it has also a procedural aspect to it as the result of unification.

LEX is a binary valued feature, i.e. [LEX +] or [LEX -] which specifies whether the sign in question is a lexical sign (+) or
a phrasal sign (LEX -).

(18)

$$\begin{bmatrix} \text{PHON} & \\ \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix}\text{HEAD } [\,]\\ \text{SUBCAT } \langle [\,] \rangle\\ \text{LEX}\end{bmatrix}\\ \text{CONTENT} & \begin{bmatrix} \text{PARAMETER} & \begin{bmatrix}\text{INDEX} & \begin{bmatrix}\text{PERS}\\ \text{NUM}\\ \text{GEND}\end{bmatrix}\end{bmatrix}\\ \text{RESTR} & \begin{bmatrix}\text{RELATION}\\ \text{INSTANCE}\end{bmatrix}\end{bmatrix}\\ \text{CONTEXT} & \begin{bmatrix}\text{BACKGROUND} & \begin{bmatrix}\text{RELATION}\\ \text{INSTANCE}\end{bmatrix}\end{bmatrix}\end{bmatrix}\\ \text{NONLOC} & \begin{bmatrix}\text{SLASH}\\ \text{REL}\\ \text{QUE}\end{bmatrix}\end{bmatrix}\\ \text{QSTORE } \{\ \} & \end{bmatrix}$$

(19)

HEAD attributes	values
MAJOR	N, V, A, PP, DET, ADV
CASE	NOM, GEN, ACC
VFORM	FIN, BSE, PSP, PRP, PAS, INF(=to), GER
NFORM	THERE, IT, NORM
PFORM	OF, ON, IN, UNDER,
AUX	+, -
INV	+, -
PRED	+, -

8.2. *The semantic categories: the CONTENT attribute*

The path SYNSEMILOCICONTENT specifies the signs semantic contribution to the context-independent semantic interpretation of a phrase containing it, especially with respect to referential meaning.

The CONTENT attribute of nominals is typed to a feature structure of sort *nomilIal object* that takes an attribute called PARAMETER (**PARA**). PARA serves as a reference marker that anchors an NP to a situation or to a discourse representation (cf. Kamp 1981). Semantic (or thematic) roles are assigned to parameters. Parameters fall into the subsorts *expletive object* (with dummy or pleonastic pronouns) and *referential object* according to their mode of referring. Referential parameters are used for contentful nouns and nonpredicative PPs and are sorted into at least three subsorts, *personal pronoun, (ppro) anaphor(ana) and non-pronoun (npro)*[3].

It is important for the analysis of agreement phenomena to keep in mind that PARAMETER allows an attribute, INDEX with the values PERSON (PERS), NUMBER (NUM), and GENDER (GEND). Two nominals are said to agree, that is to be coinde~ed, if their indices are token-identical.

The second attribute of CONTENT besides PARA is typed to objects of the sort *referential object;* it is called RESTRICTION (**RESTR**) and introduces a semantic restriction on the parameter. RESTR takes a set of parametrized states-of-affairs (cf. situation semantics, Barwise & Perry 1983) represented by the feature structure values RELATION and INSTANCE (19). By way of an example, compare the nominal entry for book in (20) and the different verbal signs in (21) - (23).

8.3. *The pragmatic information in HPSG*

In the forthcoming revised version of HPSG, Pollard & Sag acknowledge the long standing discussion on a linguistic sign's pragmatic properties. Thus, the **CONTEXT** attribute contains context dependent information that relates to indexicality, presupposition and conventional implicature. The single CONTEXT attribute mentioned, **BACKGROUND,** also takes as its values a set of parametrized states-of affairs *(psoas)* (cf. situation semantics, Barwise & Perry 1983). Pollard & Sag note that these background psoas should be considered rather like felicity conditions on the context of utterance. A second CONX attribute called CONTEXTUAL-INDICES (C-INDS) represents information about various indexical coordinates such as SPEAKER, ADDRESSEE, indices of spatiotemporal location, etc.

The CONTENT attribute of verbs takes values typed to the sort *basic circumstance* (Pollard & Sag 1987) or *(pJsoa (parametrized) states-of-affairs)* (Pollard & Sag forthcoming) . A (p) soa is represented by a feature structure which specifies a relation (RELN) together with the argument roles of that relation. In the forthcoming volume, Pollard & Sag apparently envisage an additional **LOCATION** role in the verbal CONTENT value together with certain CONTEXT attributes concerning temporal location (Reichenbach's analysis of time and the three notions of utterance time, reference time and event time underlie this treatment). A present tense verb will be marked with an EVENT TIME value in CONTEXT having the index of the LOCATION value in the CONTENT attribute and adding a BACKGROUND psoa restricting the EVENT- TIME value to temporally overlapping the UTTERANCETIME value. In the HPSG analysis, grammatical features of the verbal system thus appear modelled by both syntactico-semantic and pragmatic values.

8.4. *Some examples of lexical signs:*

(20)

```
PHON /dog/
SYNSEM  LOC  CAT      HEAD          MAJ NOUN
                                    NFORM NORM
                                    CASE...
                      SUBCAT < DET >
                      LEX +
             CONTENT  PARA | 1 |    INDEX  PER 3RD
                                           NUM SNG
                                           GEND NEUT
                      RESTRIC  {  RELATION DOG
                                  INSTANCE | 1 |  }
             CONTEXT  BACKGR   {  RELN female
                                  INSTANCE | 1 |  }
        NON
        LOC        { }
QSTORE
```

Verbs differ from nouns not only with respect to their SUBCAT values but also with respect to their CONTENT attribute which in the case of a verb directly encodes the verb's underlying predicate argument structure (cf. above). In (21), only the PHON and SYNSEM/LOC values are given:

As an intransitive 3rd-person-singular-verb, *walks* requires that the index on the subject be [PERS 3RD, NUM SNG]; this inde~ is given on the single element on the verbs SUBCAT list which is an NP [NOM]. Being a 3rd-person-singular-verb is thus simply a matter of selecting a 3rd-singular-subject. The CONTENT value expresses the verbs relation *walk* and ensures that the AGENT or WALKER role (cf. Volume I) of the relation *walk* is filled by the referential parameter of the subject (Pollard & Sag forthcoming).

In (22), the transitive verb *sees* requires two arguments on its subcat list; the first, the subject notated as NP[NOM] with the referential parameter 01', fills the EXPERIENCER role while the second argument's parameter is assigned to the THEME role. *Gives*, as a ditransitive verb requiring three arguments, cf. (23), accordingly has one more element on its SUBCAT list. However, the analysis here does not say anything on how to determine the difference in obliqueness of the two arguments NP[ACC]. This is a problem in sofar as the use of a list as value of the SUBCAT attribute enforces a total ordering: no two elements on the list can be equally oblique. Another problem has to do with the case marking which is used. There is little evidence for an accusative case in English, therefore, the AVMs given should be taken cum grano salis as they apparently introduce an entity not overtly marked in the morpho-syntactic surface strings.

(21)

```
PHON /walks/
SYNSEM  [ LOC  [ CAT  [ HEAD    [ MAJ V      ]
                      [         [ VFORM FIN  ]
                      [         [ AUX  -     ]
                      [         [ INV  -     ]
                      [ SUBCAT  <NP[NOM]            >
                      [                |1|[3RD,SNG]
                      [ LEX +                       ]
                 CONTENT  [ RELN wal  ]
                          [ AGENT |1| ]             ]]]
```

(22)

```
PHON /sees/
SYNSEM   LOC   CAT   HEAD     MAJ V
                              VFORM FIN
                              AUX    -
                              INV    -
                     SUBCAT  <NP[NOM],        NP [ACC] >
                              |1| [3RD,SNG],  |2|
                     LEX +
               CONTENT        RELN see
                              EXPER |1|
                              THEME |1|
```

(23 a)

```
PHON /gives/
SYNSEM   LOC   CAT   HEAD     MAJ V
                              VFORM FIN
                              AUX    -
                              INV    -
                     SUBCAT  <NP[NOM],        NP [ACC], NP [ACC]>
                              |1| [3RD,SNG],  |2|, |3|
                     LEX +
               CONTENT        RELN give
                              AGENT |1|
                              GOAL |2|
                              THEME |3|
```

(23 b)

```
PHON /gives/
SYNSEM   LOC   CAT   HEAD     MAJ V
                              VFORM FIN
                              AUX    -
                              INV    -
                     SUBCAT  <NP[NOM],        NP [ACC], NP [to]>
                              |1| [3RD,SNG],  |2|, |3|
                     LEX +
               CONTENT        RELN give
                              AGENT |1|
                              GOAL |2|
                              THEME |3|
```

It should be noted that the use of NP[NOM], NP[ACC] on the verbs' SUBCAT list is in fact an abbreviation for feature structures of sort *synsem;* the abbreviations are intended to increase legibility. Fully spelled out, feature structures for NP, VP, and S would look like (24) - (26):

(24)
$$\begin{matrix}\text{NP}\\ |i|\end{matrix}\;\begin{bmatrix}\text{LOC} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{HEAD noun}\\ \text{SUBCAT} < > \\ \text{LEX} \;-\end{bmatrix}\\ \text{CONTENT} \mid \text{PARAMETER } |i| & \end{bmatrix}\end{bmatrix}$$

(25)
$$\begin{matrix}\text{S}\\ |i|\end{matrix}\;\begin{bmatrix}\text{LOC} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{HEAD verb}\\ \text{SUBCAT} < > \\ \text{LEX} \;-\end{bmatrix}\\ \text{CONTENT } |i| & \end{bmatrix}\end{bmatrix}$$

(26)
$$\begin{matrix}\text{VP}\\ |i|\end{matrix}\;\begin{bmatrix}\text{LOC} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{HEAD verb}\\ \text{SUBCAT} \langle \text{synsem} \rangle \\ \text{LEX} \;-\end{bmatrix}\\ \text{CONTENT} \quad |i| & \end{bmatrix}\end{bmatrix}$$

9. Phrasal Signs and two Principles of Universal Grammar

The subsort phrasal sign allows to specify well-formed phrases of a language in the descriptive format of AVMs. In this section I will consider the structure of some simple phrases and discuss the HPSG's Head Feature Principle and the Subcategorization Principle of Universal Grammar.

Unlike lexical signs, phrasal signs have the additional attribute DTRS whose value is an AVM of type *constituent structure*. It encodes the immediate dominance information of a phrase. There are several subtypes of constituent structure, such as headed-structure, head-complement-structure, coordinate-structure, head-filler-structure and head-adjunct-structure. They take the following values: HEAD-DTR, COMP-DTR, COORD-DTR, FILLER-DTR, ADJUNCT-DTR. As every headed structure has a unique head or a unique filler daughter, but may have several complement daughters, COMP—DTRS takes a list of signs as its value. The order on the list is by increasing obliqueness. HEAD-DTR and FIlER-DTR simply take an AVM of type sign as their value.

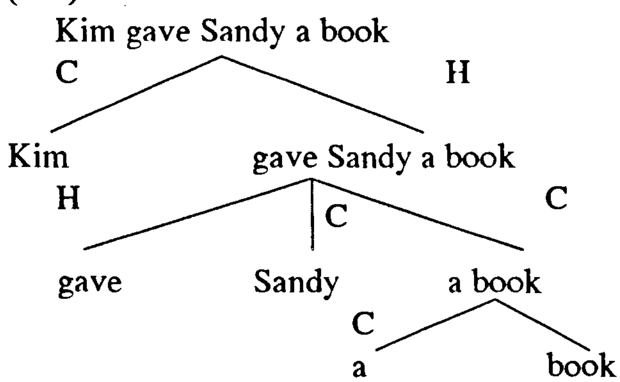
9.1. *Some Sample Phrasal Signs*

(27)

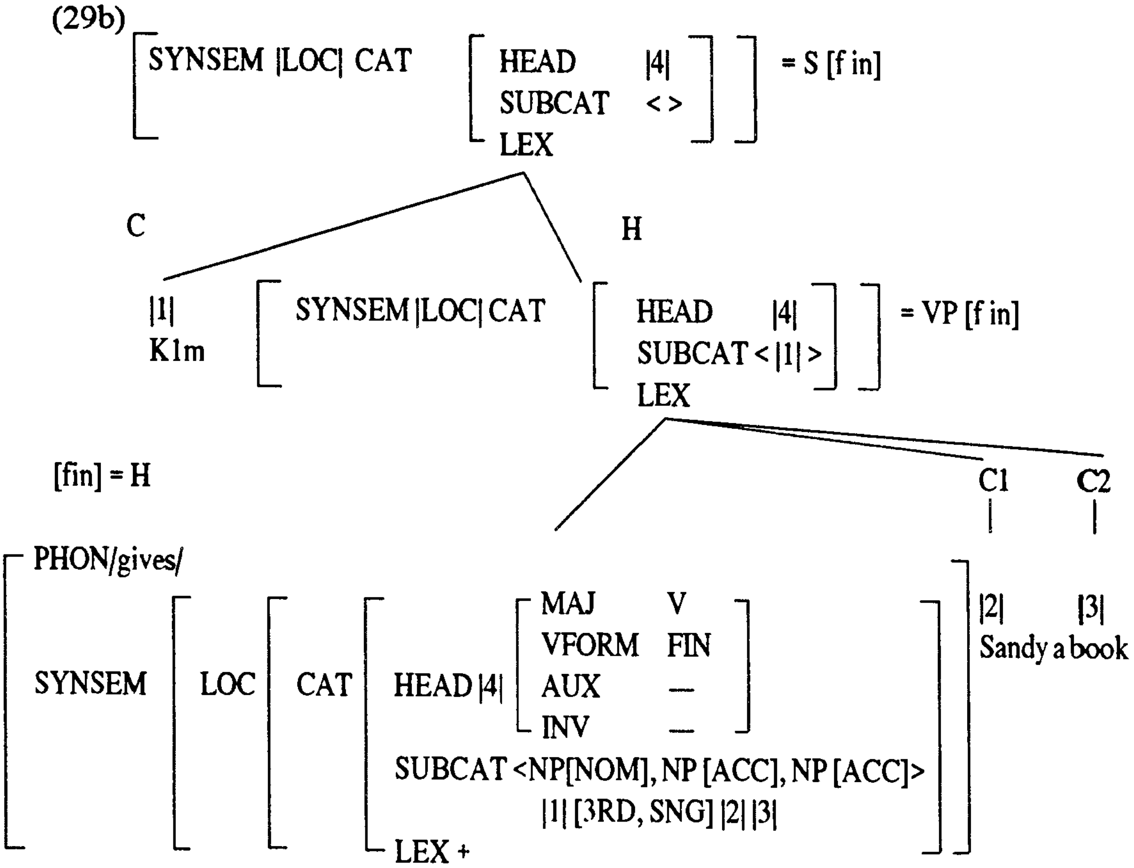
```
[ PHON /a buk/
  SYNSEM [ LOC [ CAT [ HEAD |2| [ MAJ    NOUN
                                  NFORM  NORM
                                  CASE   ...  ]
                       SUBCAT < >
                       LEX –                  ] ] ]
  DTRS [ HEAD-DTR [ PHON /buk/
                    SYNSEM [ LOC [ CAT [ HEAD |2| [ MAJ NOUN
                                                    NFORMNORM
                                                    CASE      ]
                                         SUBCAT <|1|>
                                         LEX +                ] ] ]
                    DTRS...                                   ]
         COMP-DTRS< [ PHON /a/
                      SYNSEM [ LOC [ CAT [ HEAD |1|   [MAJ DET]
                                           SUBCAT < >
                                           LEX +              ] ] ] ] ]
```

Now consider (29a) and (29b) in a simplified tree notation to enhance legibility:

(29a)



(29b)



Here , the lexical head *gives*, requires a sub ject |1|, a direct object, |2| and a second object, |3|. Moving up the tree, the first phrasal projection is the VP gives *Sandy a book*, whose SUBCAT requirements |2| and |3|, have already been satisfied by the NPs *Sandy* and a *book*. Only the sub ject value 111 needs to be satisfied. The S node takes an empty SUBCAT list, indicating that the remaining requirements on the VP's SUBCAT list have now also been met by the NP Kim. Thus, the value of the SUBCAT list reflects different degrees of saturation. If the SUBCAT value is the empty list [< >], then the constituent is saturated.

In a simplified AVM notation, (the SYNSEM values have been omitted) (29a,b) amount to (30):

(30)

```
PHON /Kim gave Sandy a book/
DTR  HEAD-DTR  PHON /gave Sandy a book/
               HEAD-DTR  HEAD-DTR [Phon/gave/]
                         DTRS  COMP-DTRS < [PHON/Kim/],  PHON /a book/
                                                         DTRS  HEAD-DTR [Phon/book/
                                                               COMP-DTRS < [Phon/a/] >   >
               COMP-DTRS < [PHON/Kim/] >
```

Each of the complement daughters of the VP has its SYNSEM value (indicated by the tags |2|, |3| in (29b)) unified with one of the elements on the head daughter's SUBCAT list, and the4 SUBCAT value of the VP itself consists of the head daughter's SUBCAT list minus those requirements that were satisfied by one of the complement daughters.

9.2. *The subcategorization principle*

The first general principle to be illustrated here is the Subcategorization Principle which states that if a constituent is a headed phrase, then its subcategorization requirements which remain to be satisfied are those of the constituent's head daughter minus any requirements that were satisfied by the constituent's complement daughters. Another way of saying the same thing is (31):

(31) Subcategorization Principle
In a headed phrase (i.e. a phrasal sign whose DTRS value is of sort headed-structure), the SUBCAT value of the head daughter is the concatenation of the phrase's SUBCAT list with the list (in order of increasing obliqueness)[4] of SYNSEM values of the complement daughters. (Pollard & Sag, forthcoming)

This may be stated as an AVM of the following form:

```
[DTRS head-struc [ ] ] =>  SYNSEM  [LOC|CAT|SUBCAT |2|]
                           DTRS    HEAD-DTR|SYNSEM|LOC|CAT|SUBCAT append(|1|, |2|)
                                   COMP-DTRS |1|
```

The principle uses a functional dependency (append) to capture the relationship between the SUBCAT list and the COMP-DTRS list. The arrow of the logical implication is to be read "any description which unifies with the antecedent must also unify with the consequent." The intended effect of this principle is to guarantee a stepwise unification of the lexical head's SUBCAT requirements with the SYNSEM values of the corresponding complements.

9.3. The Head Feature Principle

Another important principle of Universal Grammar is the Head Feature Principle. It states that with any phrasal sign of type *headed-structure*, the HEAD value of the mother is token-identical with the HEAD value of the head daughter, cf. (32):

(32) Head Feature Principle.
The HEAD value of any headed phrase is unified with the HEAD value of the head daughter. (Pollard/Sag, forthcoming)

In terms of an AVM notation this amounts to (33):

(33)

$$[\text{DTRS head-struc } [\,]\,] \Rightarrow \begin{bmatrix} \text{SYNSEM|LOC|CAT|HEAD} & |1| \\ \text{DTRS|HEAD-DTR|SYNSEM|LOC|CAT|HEAD} & |1| \end{bmatrix}$$

The HFP guarantees that headed phrases are "projections" of their lexical heads.

Thus, in any saturated phrase with a [SUBCAT < >] value where every phrasal constituent is headed, it will always be the case that

1. every SUBCAT requirement of the lexical head is unified with the SYNSEM value of precisely one complement daughter of the phrasal projection of this lexical head (due to the Subcategorisation Principle);

2. the HEAD value of the entire phrase will be unified with that of the lexical head (due to the Head Feature Principle).

10. The Syntax / Semantics Interface

The SUBCAT attribute allows to link syntactic and semantic roles. Part of its effects rest on the Semantics Principle of HPSG which ensures that the SYNSEM|LOC|CONTENT value for a phrase is the SYNSEM|LOC|CONTENT value of its head (cf. Pollard and Sag, 1987). (For notational convenience, subscripted indices are used to indicate the value of the SYNSEM|LOC|CONTENT attribute of the elements on the SUBCAT list (cf. the lexical signs *sees, walks, gives*).

For raising verbs (e.g. *appear, seem, believe, expect, tend* etc.) and control verbs (e.g. subject equi: promise as in *John promised Mary to read the book* and object equi: *persuade*, as in *John persuaded Mary to read the book*) the effects of the SUBCAT attribute and the Subcategorization principle are slightly more difficult to understand.

In the case of *appears* (34),

(34)

PHON /appears/
SYNSEM [LOC [CAT [HEAD [MAJ V, VFORM FIN, AUX –, INV –]
SUBCAT<NP [NOM], VP [inf, SUBCAT <|1|>] |3| >
|1| [3RD, SNG]
LEX +]
CONTENT [RELN appear, THEME |3|]]]

the missing subject of the embedded infinitival VP is unified with the subject NP of the sentence which is the first element on the verb's SUBCAT list . Thus, if the embedded verb requires an NP with unusual CASE and NFORM features, these subcategorisation requirements will be passed up to the matrix verb. A lexical rule will deal with those cases where appears takes a specific impersonal NFORM *it*.

In the HPSG analysis, control phenomena, raising dependencies and equi dependencies subcategorize for different types of unsaturated VP complements, i.e. complements which are not specified as [SUBCAT <]

The group of raising verbs, *appear, seem, believe, expect, tend,* and also adjectives as *likely* behave in the same fashion and accept a VP[inf] complement, (seem to be happy, *expect Sandy* to leave;

begin, keep, be take VP[prd] complements *(kept sleeeping late). become, seem, be* accept AP[+PRP] complements *(become sick); become, be* take NP[+PRD] complements *(were good little housekeepers); regard, strike* take PP[+PRD] complements marked by as *(strikes them as a disaster); be, seem* take PP[+PRD] complements *(be in trouble).*

Equi verbs (and adjectives) similarly, take different kinds of complements; i.e. *try, hope, eager, persuade, promise* accept VP[inf] complements *(try to leave,.* persuade Sally to leave);

consider, try accept VP[ger] complements *(consider selling the car);*
feel, look take an AP[+PRD] complement *(feel ill);*
make accepts an NP[+PRD] complement *(made the boy a good housekeeper)* and
got, count admits PP[+PRD] complements *(got under the table; count Kim among our friends)* (material from Pollard & Sag forthcoming, chapt. 3).

In terms of AVMs, the lexical signs for the equi verbs *persuades* and *promises* look like (35) and (36):

(35)

```
[ PHON /promises/
[ SYNSEM [ LOC [ CAT [ HEAD [ MAJ    V
                             VFORM  FIN
                             AUX    –
                             INV    –   ]
                       SUBCAT <NP [NOM], NP[ACC], VP [INF, SUBCAT <NP |1|>] |3|>
                               |1| [3RD, SNG], |2|
                       LEX + ]
                 CONTENT [ RELN     promise
                           AGENT    |1|
                           PATIENT  |2|
                           THEME    |3| ] ] ] ]
```

(36)

```
[ PHON /persuades/
[ SYNSEM [ LOC [ CAT [ HEAD [ MAJ    V
                             VFORM  FIN
                             AUX    –
                             INV    –   ]
                       SUBCAT <NP [NOM], NP[ACC], VP [INF, SUBCAT <NP |2|>] |3|>
                               |1| [3RD, SNG], |2|
                       LEX + ]
                 CONTENT [ RELN        persuade
                           PERSUADER   |1|
                           PERSUADEE   |2|
                           STAKE       |3| ] ] ] ]
```

(In the case of *promises*, it is John, while in the case of *persuades* it is Mary who reads the book. With subject equi verbs like *promise,* the subject fills the role linked to the subject of the embedded VP whereas in the case of object equi verbs like *persuade* it is the object which fills the role linked to the subject of the embedded VP).

In the HPSG analysis of control phenomena, the difference between raising and equi verbs is seen in two facts:

1. The controllers of equi verbs (and adjectives) are assigned one more semantic role than raising verbs: all subcategorized arguments of equi verbs are assigned a semantic role; raising verbs always fail to assign a semantic role to one argument that they subcategorize for; and

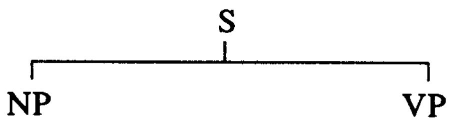
2. for equi verbs, the controller (e.g. the VP complement's unexpressed subject, its SUBCAT member) shares a parameter (is co-parameterized) with one of the other syntactic dependents (subject in subject equi and object in object equi verbs). In the case of raising verbs, the entire SYNSEM value of the subject of the VP complement is unified with that of either the subject (subject raising) or the object (object raising) (Pollard & Sag forthcoming; chap. 3).

11. The ID-Rules and LP-Statements

Phrase structure rules specifying constituent structure and order, and the corresponding PS-trees like (37),

(37a) S => NP VP

(37b)



encode information about immediate dominance (ID) and about linear precedence (LP) of constituents. Thus, the mother node S immediately dominates the daughters NP and VP. The same rule also says that NP linearly (from left to right) precedes its sister node VP. HPSG factors out the ID-relationship by considering it part of the Universal Grammar and treats LP-statements as language specific facts of constituent order.

With the separation of ID and LP relations, HPSG's grammar rules are really just highly schematic immediate dominance templates (descriptions of phrases), such that every non-lexical constituent must be an extension of one such schematic description. With the recent trend towards syntactic minimalism and the lexicalisation of grammar in which HPSG participates, the set of ID schemata is reduced to a very small number. Pollard and Sag (forthcoming) disjunctively join them in the Immediate Dominance Principle. It contains roughly 4 schemata.

(38) Schema 1: Non-lexical Heads and Subjects:

A saturated Phrase ([SUBCAT < >]) with DTRS value of sort head complement structure in which the HEAD-DTR value is a phrasal sign and the COMP-DTR value is a list of length one, is a well-formed phrase.

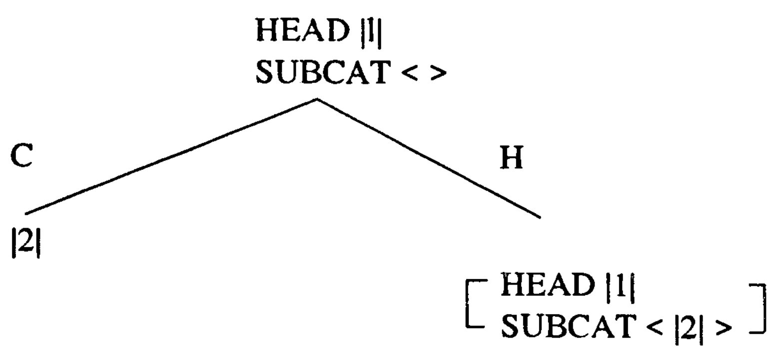
The AVM format (39)

(39)

[SYNSEM|LOC|CAT|SUBCAT < >
  DTRS [HEAD-DTR|SYNSEM|LOC|CAT|LEX -
         COMP-DTRS <[ ]>]]

simply says that one of the possibilities of an English phrasal sign is to be a saturated sign which has as constituents a single complement daughter and a HEAD-DTR which is a phrasal sign. The Subcategorization Principle demands that the complement daughter must satisfy (i.e. have its SYNSEM value unified with) the single element on the SUBCAT list of the head daughter. This element must be the least oblique SUBCAT element, i.e. the subject of the lexical head. The Head Feature Principle in turn demands that the lexical head's features are passed on to its mother node. A phrase licensed by Schema 1 and following the SP and HFP will have the form (40)

(40)



Schema 1 which, in an informal rewrite notation looks like (41),

(41) [SUBCAT < >] => H [LEX -], C

corresponds to the following PS-Rules (42):

(42)	S	=>	NP VP
	NP	=>	DET NP
	NP	=>	NP[GEN] NP

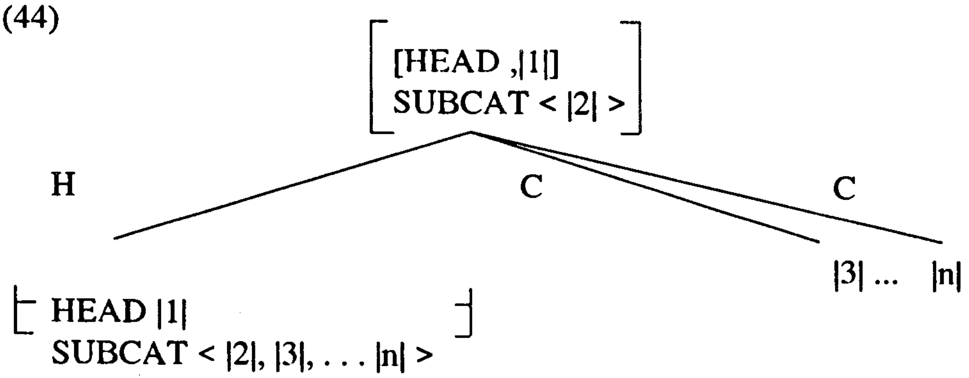
The second schema is given in (43):

(43) Schema 2: Lexical Heads + Non-Subject Complements:

An almost saturated phrase (SUBCAT list of length one) with DTRS value of sort head-complement-structure in which the HEAD-DTR value is a lexical sign, is a well-formed phrase.

This schema licenses phrases with a lexical daughter and zero or more complement daughters which have satisfied all their subcategorization requirements except the least oblique one, the subject; i.e. almost saturated phrases consisting of a lexical head and all but one of its complements. Any phrase licensed by schema 2 will have the general form (44)

(44)



This schema subsumes the phrase structure rules in (45)

(45)	VP	=>	V	
	VP	=>	V	NP
	VP	=>	V	NP PP
	NP	=>	N	
	NP	=>	N P$_{[of]}$	

and looks like (46) in AVM notation, and like (47) in informal rewrite notation:

(46)

SYNSEM|LOC|CAT|SUBCAT	<[ ]>	
DTRS|HEAD-DTR|SYNSEM|LOC|CAT	HEAD	
	LEX +	

(47) [SUBCAT < [ ] >] => H [LEX +], C*.

This schema interacts with the HFP, the SP, and the English Constituent Ordering Principle. The HFP identifies the phrase's head feature values with those on the head daughter; the SP guarantees that the lexical head will have a SUBCAT list whose value is one element greater than the number of complement daughters which actually satisfy the subcategorisation restrictions on the lexical head; and the Constituent Ordering Principle ensures that in the corresponding utterance tokens the lexical head will precede the complements. The Constituent Ordering Principle for English implements the LP-rule that nonlexical heads follow their complements.

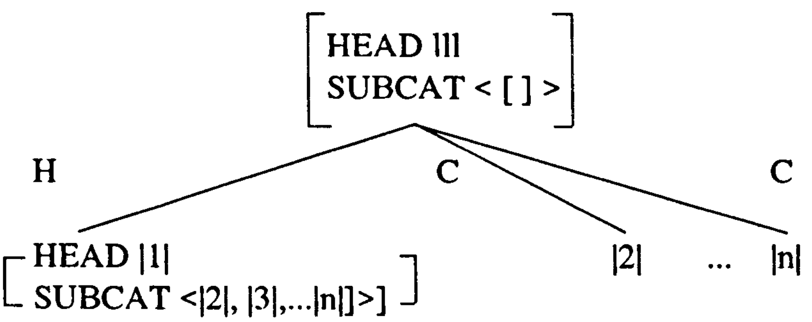
Schemata 1 and 2 subsume a large number of traditional PS-Rules; they account for all non-inverted head-complement structures and cover a substantial fragment of English syntax. Schema 3 is designed specifically for inverted structures such as English direct questions.

(48) Schema 3: Inverted Structures:

A saturated ([SUBCAT ~ >]) phrase with DTRS value of sort head-complement-structure in which the HEAD-DTR value is a lexical sign is a well-formed phrase.

Schema 3 allows phrases in which all complements, including the subject, are realized as sisters of the lexical head. By the HFP and the SP, phrases licensed by Schema 3 will have the general form shown in (49):

(49)



This rule licenses phrases such as *Did Mary give the book to John?*. As the phrase is lexically headed, the English Constituent Ordering Principle ensures that this head fills the phrase initial position.

All non-auxiliary verbs must be marked as [HEAD|INV -]; auxiliaries are simply treated as verbs marked AUX + and the majority of them are left unspecified for the head feature INV. (Exceptions to this are auxiliaries like *aren 't* = [INV-] and better = [INV+][5]); thus, only auxiliaries can head phrases licensed by schema 3. Please note, that by allowing schemata 1 and 2 to take care of uninverted phrases and to employ schema 3 for inverted phrases, there must be the additional (language specific) condition on the HEAD values of schemata 1 and 2 that they bear the specification INV+.

In AVM notation, this schema looks like (50) and in the informal rewrite notation like (51):

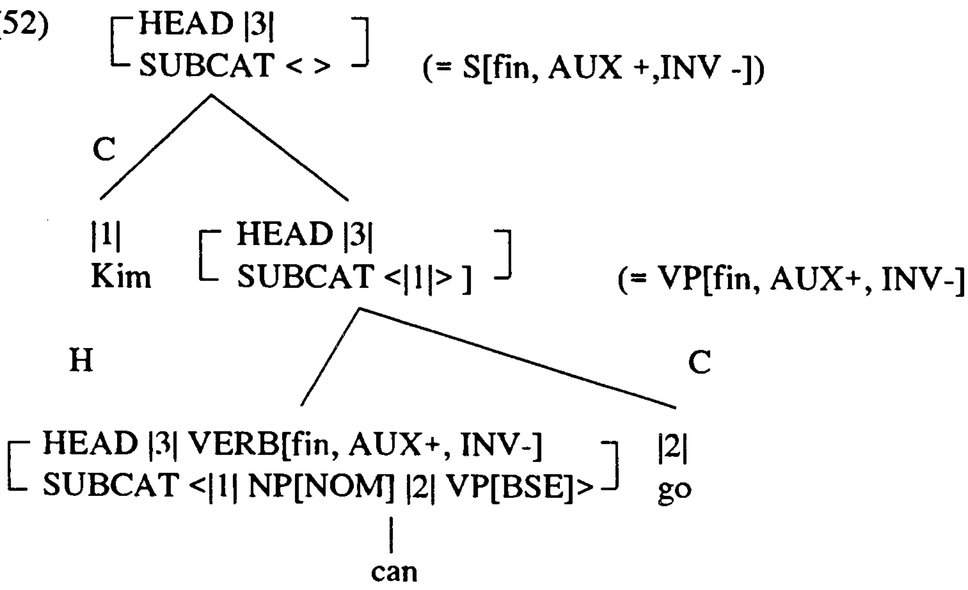
(50)

[SYNSEM|LOC|CAT|SUBCAT < >
  DTRS|HEAD-DTR|SYNSEM|LOC|CAT [HEAD | INV +
                                  LEX +]]

(51) [SUBCAT < >] => H[INV +, LEX +] C*

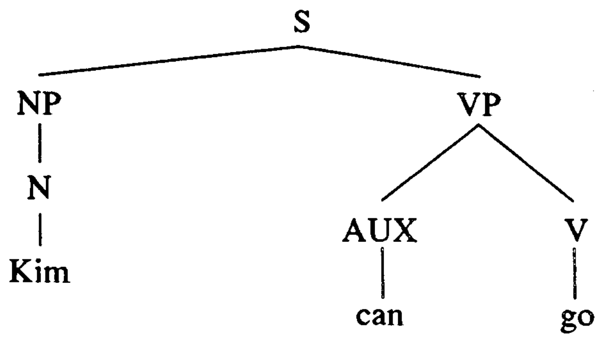
The functions of the above schemata can nicely be demonstrated with the two sentences *Can Kim go?* and *Kim can go.* (52), (53).

(52)



Here, the S and VP nodes are licensed by Schema 1 and 2 respectively. Except for the fact that *can* is treated as a verb, the analysis is a rather traditional one.

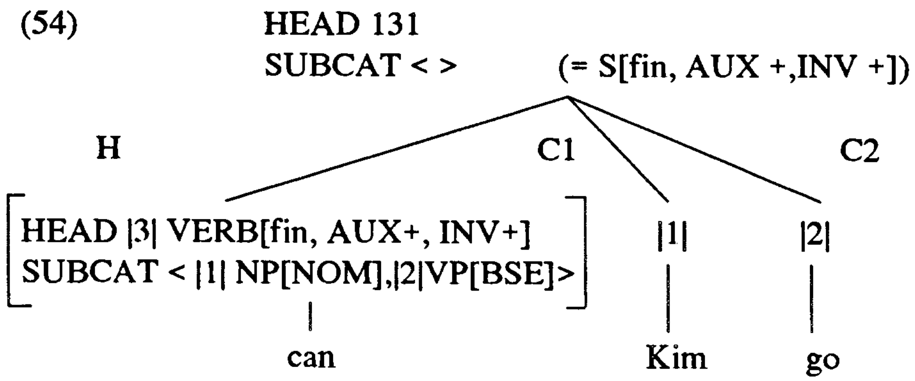
53)



With the inverted sentence, the situation is different:

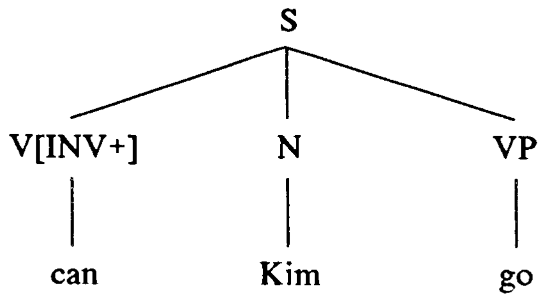
HPSG does not have a constituent *can go* in its analysis of the inverted phrase: cf. (54)

(54)



Represented in the form of a tree diagram, (52) looks like (53):

(53)



12. The Lexicon Component in HPSG

The lexicon plays a central role in HPSG, given the massive relocation into the lexicon of linguistic information that was formerly carried by a complex apparatus of syntactic rules, e.g. phrase structure rules and transformations in gTG or metarules in GPSG. HPSG's grammar contains less than 20 rules versus some 350 in a GPSG implementation and some 1500 in a generative framework. The syntactic simplification is made possible by an enrichment of the structure and content of the lexicon. However, the linguistic information contained in the lexicon should be represented in an efficient, conceptually clear, readily modifiable format and must allow to express generalisations so that every piece of information appears only in one single lexical entry. An outline of the structure of an HPSG lexicon and the mechanisms used in its representation will be given below.

The complexity of information stored in the lexicon is evident from the partial description of the lexical signs. HPSG deals with this complexity by utilizing two mechanisms: multiple inheritance within a hierarchy of lexical signs and lexical (redundancy) rules. Language users have knowledge of a system of lexical types and subtypes. By factoring out information about words which can be predicted from their membership in a class of lexical elements (its supertype) whose properties will be stated only once but inherited by all its members (its subtypes), the amount of idiosyncratic information in individual lexical entries is dramatically reduced. Lexical rules provide an elegant and completely lexicon-internal account of many phenomena that apply to classes of lexical signs; such rules take care of phenomena like inflectional and derivational patterns but also of syntactic phenoma known by their transformational names like *passive, tough - movement; i t - extraposition.*

12.1. *Inheritance in the Lexicon and the Hierarchy of Lexical Types*

HPSG strives to eliminate redundant information in lexical entries by not fully specifying the individual entries and by distributing the information pertaining to each entry over a large number of word classes organised in a non-trivial set of hierarchies. Two assumptions underlie this position:

1. Each distinct lexical property exhibited by one or more elements of the lexicon is introduced exactely once in the lexicon.

2. A property shared by more than one lexical item is introduced in a word class common to those items (or these items should be related by a lexical rule) (Flickinger et al. 1985; Flickinger 1987).

HPSG makes use of word classes to capture generalizations about the lexical elements in order to make predictions about the behaviour and distribution of a lexical item on the basis of the classes it belongs to. A word class will be stipulated if it can be demonstrated that some particular morphological or syntactic property is shared by a number of lexical items. All lexical entries that have this property will be elements of this class. As each lexical item has several distinct properties, it will have to be cross-classified and no simple branching hierarchy to represent the lexicon will be in order: a hierarchy must be partitioned into different dimensions.

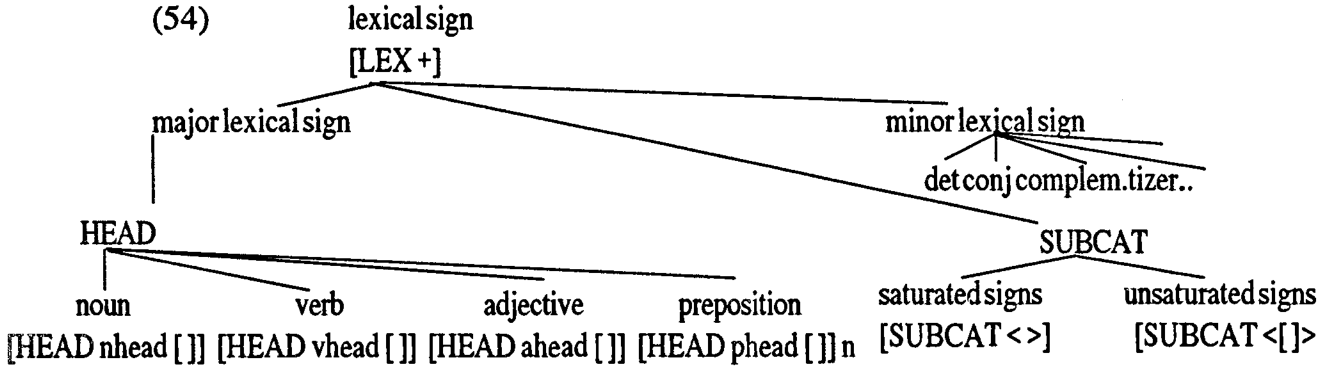
A simple illustration of the need for assigning words to more than one class involves 2 independent criteria for partitioning nouns in English:

One partition of nouns rests on the distinction of noun types: common nouns, (cn = *dog, book*), proper nouns (pn = *John, Kim*), personal pronouns (ppro = she, *they*, . . .), anaphors (ana = *himself*); expletive pronouns (expl = *it, there*) . Each of these subtypes of noun exhibits syntactic and semantic properties not shared by the other types; e.g. common nouns name/refer to classes of entities in the world and syntactically permit or require one obligatory argument, and allow a range of adjuncts (adjectives and restrictive relative clauses) not permitted with proper nouns. On the other hand, proper nouns name/refer to individuals, permit no obligatory argument but may take titles as adjuncts which are not permitted for common nouns.

The second partition distinguishes nouns according to their number property by separating singular (sn), plural(pn) and mass nouns (mn). (In German and also Spanish, the class of nouns has a third dimension: gender).

The set of feature structures types used in the description of the English lexicon is assumed to be partially ordered by subsumption, i.e. some feature structures are supertypes of subtypes so that the former subsume the latter. Subtypes inherit all the attributes and corresponding type restrictions on their values from all of their supertypes. The hierarchy will allow multiple inheritance if lexical signs are allowed to be members of several cross-cutting partitions. It is assumed that the fundamental dimensions for classifying lexical entries are part of speech and valence. The root of the hierarchy, beginning with lexical sign is as follows

(54)



Sofar, the lexical hierarchy indicates only syntactic information associated with the lexical types in question; the specific kind of semantic attributes connected with the CONTENT attribute also constrain the lexical types; lexical signs of type noun and verb have the following CONTENT values (55):

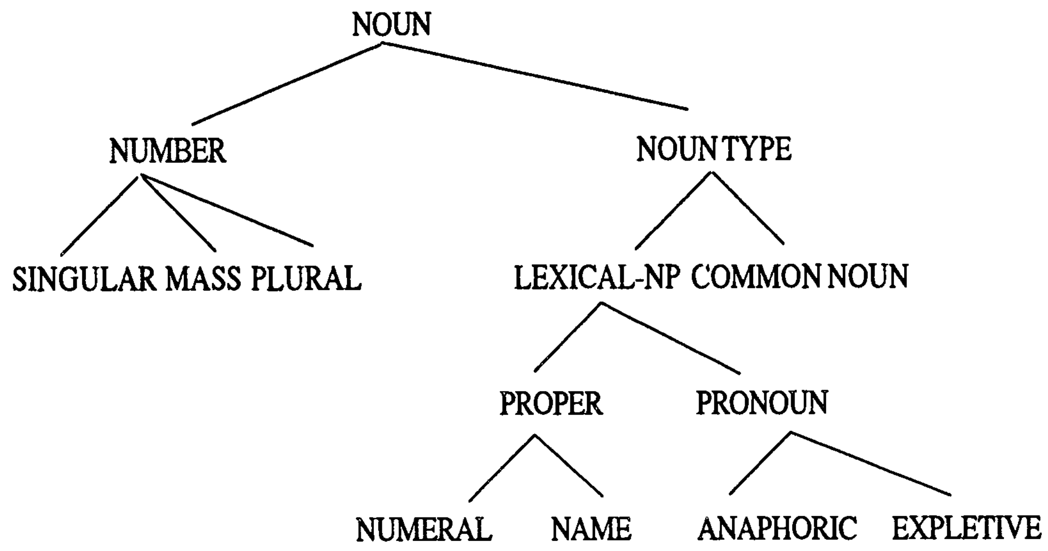
(55)

$$\text{noun}[\ ]_{\text{major lex sign}} = \begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD} & \text{nhead}[\ ] \\ \text{CONTENT} & \text{indexed-obj}[\ ] \end{bmatrix}$$

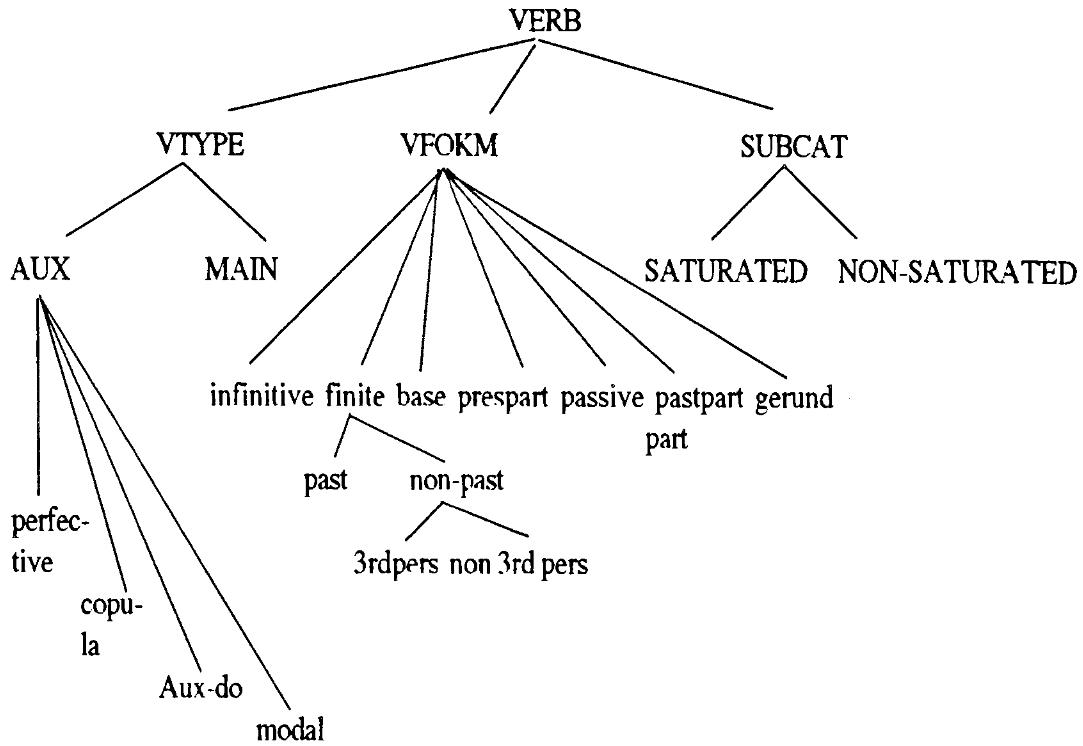
$$\text{noun}[\ ]_{\text{major lex sign}} = \begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD} & \text{vhead}[\ ] \\ \text{CONTENT} & \text{basic circum stance}[\ ] \end{bmatrix}$$

In (56) and (57) the subclasses of nominal and verbal lexical signs are given in order to make the notion of lexical hierachy more concrete. The system of verbal signs constitutes the regular aspects of the English verbal inflectional paradigm, cross-classifies according to the main-verb - auxiliary distinction and has valence as a third dimension.

(56) Subclasses of NOUN



(57)



The basic idea then is simple: the lexicon contains (abstract, non-lexical) entries specifying properties of the above classes and subclasses; a lexical entry contains the information that specifies its membership in these classes. Thus, the information contained in an abstract entry will be "inherited" by every lexical entry specified for membership in the abstract class. As an example, consider the almost fully specified entry for the common noun dog which refers to a female dog in this case (20)

(20)

```
PHON /dog/
SYNSEM   [ LOC  [ CAT      [ HEAD          [ MAJ NOUN
                                             NFORM NORM
                                             CASE ... ]
                             SUBCAT <DET>
                             LEX + ]
                  CONTENT  [ PARA |1|        [ INDEX  [ PER 3RD
                                                        NUM SNG
                                                        GEND NEUT ] ]
                             RESTRICT   {    [ RELATION DOG
                                               INSTANCE |1| ]     } ]
                  CONTEXT  [ BACKGR     {    [ RELN female
                                               INSTANCE |1| ]     } ] ]
           NON
           LOC        { }  ]
QSTORE
```

The entry dog shares with all words the information that the sign is a lexical sign

(58) [SYNSEM | LOC | CAT | LEX +]

It shares with all nouns the property of being a noun

(59)

```
[ SYNSEM  [ LOC [ CAT      [HEAD      [MAJ NOUN]]
                  CONTENT
                                 indexed-object  [     ]

                  CONTEXT           context  [     ] ] ] ]
```

It shares with all common nouns among other things the specific SUBCAT potential of English common nouns

(60)

```
[ SYNSEM [ LOC [ CAT     [ HEAD   MAJ NOUN                 ] ] ] ]
                                  NFORM NORM
                                  CASE...
                           SUBCAT <DET v POSP>
                           LEX +
                 CONTENT [ PARA|1|
                                           [indexed-obj[ ]]
                           RESTRICT { [ RELATION          ] } ]
                                      [ INSTANCE|1|       ]
                 CONTEXT   [BAKGROUND [INSTANCE|1|]]
```

It shares with all nouns that are 3rd singular common nouns as opposed to other combinations of number and person values

(61)

```
[SYNSEM [LOC [CONTENT [PARAMETER [INDEX [ PER 3RD   ]]]]]]
                                        [ NUM SNG   ]
                                        [ GEND NEUT ]
```

So, the information specific to the lexical sign *dog* that is contained in this entry is very little:

(62)

```
[ PHON /dog/                                          ]
[ SYNSEM|LOC|CONTENT|PARAM|RESTRCT|RELN DOG           ]
```

The information, that the lexical entry for *dog* is a lexical sign of type noun, and a 3rd person singular common noun is inherited from its membership in the abstract classes of lexical sign, noun etc.

This method of inheritance from a multiple hierarchy of classes also holds for verbs except that verbs are also grouped on the basis of their SUBCAT potential. The verb *walks* for example has a lexical entry (63)

(63) [PHON /walks/
SYNSEM | LOC | CAT | SUBCAT [NP [nom] |1| [3rd, sing]]
|CONTENT [RELN walk
AGENT |1|]]

The following information is shared

(64) a. Shared by all words
[SYNSEM | LOC | CAT | LEX +]

b. Shared by all verbs (in addition to a)
[SYNSEM | LOC | CAT | HEAD | MAJ v
CONTENT basic circumstance [ ]]

c. Shared by all main verbs (in addition to a, b)
[SYNSEM | LOC | CAT | HEAD [AUX –
INV –]]

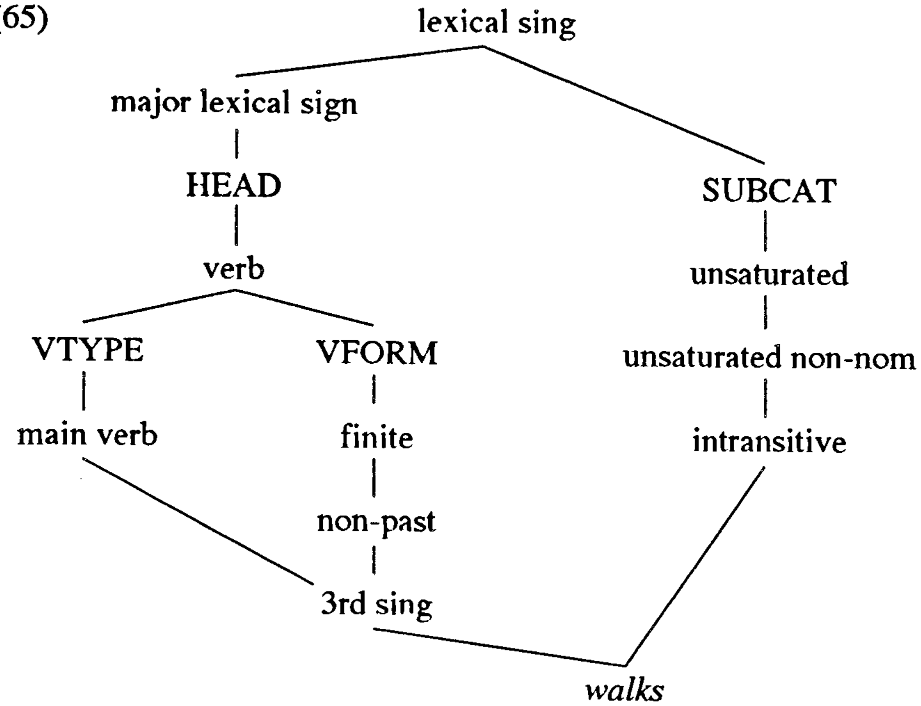
d. Shard by all finite verbs (in addition to a, b, c)
[SYNSEM [LOC [CAT [HEAD [VFORM FIN
PRD –]]
CONTENT [RELN walk
AGENT]]]]

e. Shared by all intransitive verbs
[SYNSEM [LOC [CAT [SUBCAT <NP [nom] |1|>]]]]

f. Shared by all 3rd-person-verbs
[SYNSEM [LOC [CAT [CONTENT [AGENT [PERSON 3RD
NUMBER SING]]]]]]

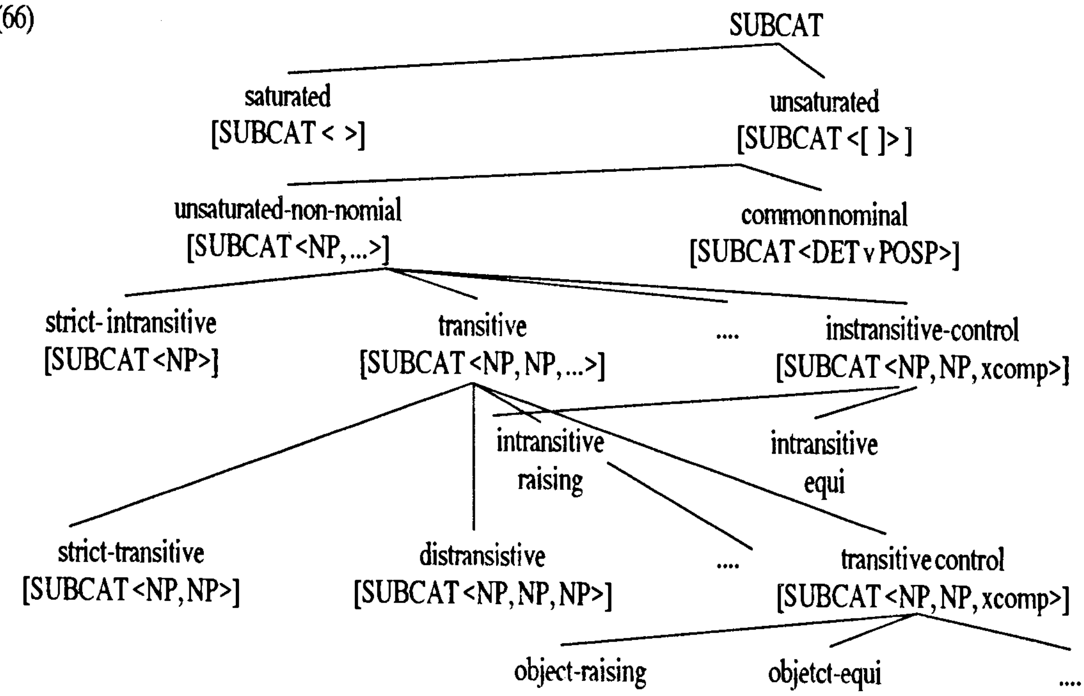
The position of the verb *walks* in the lexical hierarchy is roughly as shown in (65)

(65)



Before discussing the mechanism of lexical rules, a short note on the hierarchy of SUBCAT types may show the partial ordering of valence types which adds another dimension to the lexical type hierarchy.

(66)



12.2. *Lexical Rules*

In order to capture phonological, morphological and syntactic regularities and patterns such as nominal declension patterns in German or Latin, English verb conjugation, nominalization and un-prefixation or regular syntactic behaviour of verbs such as valence shifts due to passivization or dative shift, the mechanism of lexical (redundancy) rule has been used. A lexical rule is a tripel <name, input-specification, output-specification> where properties of an input are mapped into a set of outputproperties and everything not mentioned in the rule remains unchanged. Given a lexical rule and an input specified for that rule, the output can be predicted.

Take for example the lexical rule which will accept a verbal base of verbtype main as input and will produce the 3rd person singular verbform as output. The application of this rule to auxiliary verbs is blocked by specifiying that the input has to be of type main verb.

(67) 3RDSNG inflectional rule

$$\left[\begin{array}{ll} \text{PHON} & |1| \\ \text{SYNSEM|LOC|CAT|SUBCAT} & |2| \\ \text{|CONTENT} & |3| \end{array}\right]_{\text{base\&main}} \Rightarrow \left[\begin{array}{ll} \text{PHON} & f_{\text{3RDSNG}}(|1|) \\ \text{SYNSEM|LOC|CAT|SUBCAT} & |2| \\ \text{CONTENT} & |3| \end{array}\right]_{\text{3rdsng}}$$

Pollard & Sag (1987) give the following interpretation to this rule: the function 3RDSNG takes a lexical sign of types base and main with phonology |1|, subcategorization |2|, and semantic content |3| as input and unifies as output the results of that function (appending -s to the phonological base) with the unchanged input properties (SUBCAT and CONTENT).

A more subtle, informal version of the function 3RDSNG:

/s/ (spelled -s) will be added after bases ending in non-sibilant voiceless consonants;

/z/ (spelled -s) will be added after bases ending in voiced nonsibilant consonants or vowels;

/ɪz/ (spelled -es) will be added to a stem ending in sibilant.

It ought to be noted though, that besides the morphophonological effects, this function also adds new HEAD feature values and agreement information on the subcategorized subject.

Seen from the point of view of individual lexical entries, each entry contains idiosyncratic information about the word in question, including information about irregular forms which precludes the entry participating in a lexical rule. For example, the base entries of irregular German or Spanish verbs may list all the principal verbal forms[6]. Furthermore there are pointers to the minimal lexical types from which it inherits; both will be enough of a specification of required input properties for membership in a lexical rule.

Finally, to give a taste of the syntactic effects which lexical rules may have, (68) presents the passive lexical rule

(68) Passive Lexical Rule

```
            [ PHON |1|                                          ]
            [ PASTPART |2|                                      ]
            [ SYNSEM | LOC | CAT | SUBCAT <[ ] |3|, [ ] |4| >   ]   =>
            [                   |CONTENT |5|                    ]
base & trans

            [ PHON f_PSP (|1|, |2|                                     ]
            [ SYNSEM | LOC | CAT | SUBCAT <(PP [BY] |4|, ...., [ ] |3|> ]
            [                   |CONTENT |5|                           ]
passive
```

The core of this rule concerns the subcategorization: the input form's subject is simply omitted from the SUBCAT list and its index reassigned to an optional PP[BY]. Transitive verbs which will not passivize will need information in their lexical entry to this effect. (Pollard & Sag 1987).

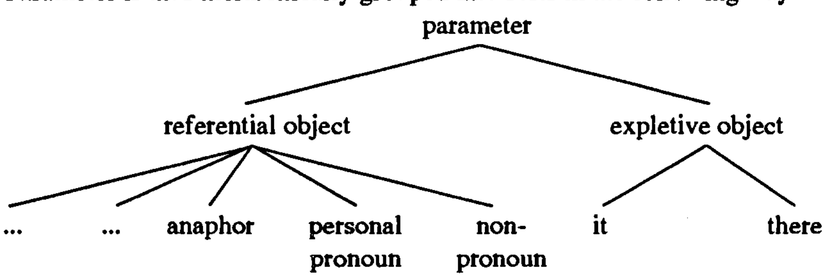
NOTAS

1 In this paper, I draw on prepublication versions of chapters 1, 1, 3, 4, S of Pollard & Sag (forthcoming), *Information-Based Syntax and Semantics, Vol. 2: Topics in Binding and Control.* Stanford: CSLI. I do not always mark my textual dependence on these materials. Errors of interpretation and presentation, however, are solely mine.

With kind permission from Richard Cooper, Centre for Cognitive Studies, University of Edinburgh, I have drawn on his paper"An Introduction to HSPG" presented at the European Summer School on Language, Logic an Information, Leuven, July 1990. Again I do not specifically mark my dependence on his material.

2 The head of a phrase is that daughter (inmediate constituent) or a phrase wich either is or contains the phrase's lexical head. The lexical head is that word which determines most of the syntactic properties of the whole phrase as a whole. Thus the lexical head of an NP is a noun, of a VP is a verb, of a PrepP is a preposition etc.

3 Parameter is thus hierarchically grouped into sorts in the following way



Only referential objects can bear semantic roles so that the parameters of equi-controllers (subject or object of the embedded VP) are always of sort *referential object.* Raising controllers are not assigned a semantic role in the CONTENT value of raising verbs. As a consequence, raising controllers are expletives in those cases in which the unsaturated complement of the raising verb subcategorizes for an expletive subject. If the unsaturated complement of a raising verb subcategorizes for a complement with a referential parameter, the raising-controller will be assigned a semantic role.

4 The theory of argument obliqueness HPSG subscribes to rest on a hierarchy of grammatical relations as suggested by Keenan and Comrie (1977). Pollard & Sag (forthcoming) suggest a reverse of the order suggested in Pollard & Sag (1977).

5 The exceptions are due to facts such as:

You better wash the dishes.
**Better you washthe dishes.*
Aren't I in charge of refreshments?
**I aren't in charge of refreshments.*

6 Consider for example German Ablautreihen as in

PHON /schreib-/	PHON /trink-/
PAST /-ie-/	PAST /-a-/
PARTICIPLE /-ie-/	PARTICIPLE /-u-/
SYNSEM/LOC/CAT	SYNSEM/LOC/CAT....

BIBLIOGRAPHY

H. Ait-Kaci (1984), "A New Model of Computation Based on a Calculus of Type Subsumption." Ph.D. Thesis, University of Pennsylvania, Philadelphia, PA.

J. Barwise; J. Perry (1983) Situations and Attitudes. Cambridge, Mass.: MIT Press.

R. Borsley (1987), Subjects and Complements in HPSG. CSLI Report CSLI 87-107

J. Bresnan (Ed.) (1982), The Mental Representation of Grammatical Relations. MIT Press, Cambridge, Mass.

R. Cooper (1990) Specifiers, Complements and Adjuncts in HPSG. Centre for Cognitive Science Research Paper, EUCCS/RP-42. Centre for Cognitive Science, University of Edinburgh.

D. Flickinger; C. Pollard; T. Wasow (1985), "Structure Sharing in Lexical Representation." In Proceedings of the 23rd Annual Meeting of the ACL, Chicago 1985.

D. Flickinger (1987), Lexical Rules in the Hierarchical Lexicon. PhD Dissertation. Stanford University 1987.

G. Gazdar; E. Klein, G. K. Pullum; I. A. Sag (1985), Generalized Phrase Structure Grammar. Harvard University Press, Cambridge, Mass.

R. A. Hudson (1984), Word Grammar. Oxford: Blackwell.

M. Johnson (1988) "Attribute-Value Logic and the Theory of Grammar." CSLI Lecture Notes 16. Stanford, CA.

L. Karttunen (1986), "Radical Lexicalism," CSLI Report No. 8668, CSLI, Stanford, CA.

R. T. Kasper; W. C. Rounds (1986) "A Logical Semantics for Feature Structures." In Proceedings of the 24th ACL, Columbia University, New York. pp 257-265.

M. Kay (1984), "Functional Unification Grammar: A Formalism for Machine Translation." In: Proceedings of COLING 1984.

M. Kay (1985), "Parsing in functional Unification Grammar." In D. Dowty, L. Karttunen, and A. Zwicky (eds.), Natural Language Parsing. Cambridge University Press, Cambridge.

C. J. Pollard; I. A. Sag (1988), An Information-Based Syntax and Semantics. Volume 1: Fundamentals, CSLI Lecture Notes 13, CSLI, Stanford, CA.

C. J. Pollard; I. A. Sag (forthcoming), An Information-Based Syntax and Semantics. Volume 2: Topics in Control and Bindillg. CSLI Lecture Note Series, Stanford: CSLI.

C. J. Pollard; I. A. Sag (1987), An Information-Based Theory of Agreement. CSLI Report # CSLI-88-132.

C. J. Pollard (1984), Generalized Conte.~rt-Free Grammars, Head Grammars and Natural ~anguage. PhD Thesis, Stanford University.

C. J. Pollard (1989), "Sorts in Unification-Based Grammar and What They Mean." In M. Pinkal & B. Gregor (eds.) Unification in Linguistic Analysís.

C. J. Pollard (1988), The Nature and Structure of a Computational Linguistic Theory. Proceedings of the First Republic of China Workshop on Computational Linguistics (198~J

C. J. Pollard (1989), "The Syntax-Semantics Interface in a Unification-Based Phrase Structure Grammar. " In Views of the Syntax/Semantics Interface. Proceedings of the Workshop "GPSG and Semantics", Berlin. pp 167-184

C. J. Pollard; D. Moshier (1989) "Unifying Partial Descriptions of Sets. " In P. Hanson (ed.) (1989): Information, Language, and Cognition.

W . Rounds ; R . Kasper (1986), "A Complete Logical Calculus for Record Structures Representing Linguistic Information. " In Proceedings of the 1st IEEE Symposium on Logic in Computer Science, pp 38-43.

W . Rounds (19 8 8), Se t Val ues f or Uni f i ca ti on -Based Grammar Formalisms and Logic Programming. CSLI Report # CSLI88-129 .

I. A. Sag (1986), Grammatical Hierarchy and Linear Precedence. CSLI Report # CSLI-86-60

S. M. Shieber (1986), An Introduction to Unification-Based Approaches to Grammar. CSLI Lecture Notes 4. CSLI, Stanford, CA.

S. Shieber; H. Uszkoreit; J. Robinson; M. Tyson (1983), "The Formalism and Implementation of PATR-II." In Research on Interactive Acquisi tion and Use of Knowl edge. SRI International, Menlo Park, CA.

S. Shieber (1984), "The Design of a Computer Language for Linguistic Information. " In S. Shieber, L. Karttunen, and F. Pereira (Eds.) Notes from the Unification Underground. SRI Technical Note 327, SRI International, Menlo Park, CA.

G. Smolka (1988), "A Feature Logic with Subsorts. " LILOG Report 33, IBM Germany, Stuttgart.

G. Smolka; H. Ait-Kaci (1987), "Inheritance Hierarchies: Semantics and Unification." MCC Report AI-057-87, MCC Austin, TX. To appear in Symbolic Computation.

S. Starosta (1988), The Case for Lexicase. An Outline of Lexícase Grammatical Theory. London, New York: Pinter

H. Uszkoreit (1986), "Categorial Unification Grammars. " In Proceedings of COLING ~6, Bonn. Also appeared as CSLI Report No. 86-66, CSLI, Stanford, CA.

H . Uszkoreit (1988) "From Feature Bundles to Abstract Data Types: New Directions in the Representation and Processing of Linguistic Knowledge. " In A. Blaser (Ed.) Natural Language on the Computer. Springer, Berlin, Heidelberg, New York.

H. Zeevat; E. Klein; J. Calder (1987), "Unification Categorial Grammar. " In Haddock, Klein, and Morrill (Eds.) Categorial Grammar, Unification Grammar, and Parsing. Edinburgh Working Papers in Cognitive Science, Vol. 1. Centre for Cognitive Science, University of Edinburgh, Edinburgh .