Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

Document
D-94-15

German Nominal Syntax in HPSG

– On Syntactic Categories and Syntagmatic Relations –

Stephan Oepen

December 1994

Deutsches Forschungszentrum für Künstliche Intelligenz GmbH

Postfach 20 80
67608 Kaiserslautern, FRG
Tel.: + 49 (631) 205-3211
Fax: + 49 (631) 205-3210

Stuhlsatzenhausweg 3
66123 Saarbrücken, FRG
Tel.: + 49 (681) 302-5252
Fax: + 49 (681) 302-5341

Deutsches Forschungszentrum für Künstliche Intelligenz

The German Research Center for Artificial Intelligence (Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI) with sites in Kaiserslautern and Saarbrücken is a non-profit organization which was founded in 1988. The shareholder companies are Atlas Elektronik, Daimler-Benz, Fraunhofer Gesellschaft, GMD, IBM, Insiders, Mannesmann-Kienzle, Sema Group, Siemens and Siemens-Nixdorf. Research projects conducted at the DFKI are funded by the German Ministry for Research and Technology, by the shareholder companies, or by other industrial contracts.

The DFKI conducts application-oriented basic research in the field of artificial intelligence and other related subfields of computer science. The overall goal is to construct systems with technical knowledge and common sense which - by using AI methods - implement a problem solution for a selected application area. Currently, there are the following research areas at the DFKI:

- ☐ Intelligent Engineering Systems
- ☐ Intelligent User Interfaces
- ☐ Computer Linguistics
- ☐ Programming Systems
- ☐ Deduction and Multiagent Systems
- ☐ Document Analysis and Office Automation.

The DFKI strives at making its research results available to the scientific community. There exist many contacts to domestic and foreign research institutions, both in academy and industry. The DFKI hosts technology transfer workshops for shareholders and other interested groups in order to inform about the current state of research.

From its beginning, the DFKI has provided an attractive working environment for AI researchers from Germany and from all over the world. The goal is to have a staff of about 100 researchers at the end of the building-up phase.

Dr. Dr. D. Ruland

Director

German Nominal Syntax in HPSG

– On Syntactic Categories and Syntagmatic Relations –

Stephan Oepen

DFKI-D-94-15

This work has been supported by a grant from The Federal Ministry for Research and Technology (FKZ 01 IV 101 K/1).

ISSN 0946-0098

German Nominal Syntax in HPSG

– On Syntactic Categories and Syntagmatic Relations –

Stephan Oepen (oe@dfki.uni-sb.de)

December 1994

Abstract

The German nominal group — even when reduced to the core inventory of nouns, determiners and attributive adjectives — is a morphologically and syntactically complex structure.

In this paper it is suggested that a detailed understanding of the (morpho-) syntactic categories and the syntagmatic relations exhibited in the core nominal group is a prerequisite to an adequate analysis. It will be argued that the two fundamental syntagmatic relations holding within the nominal group, viz. GOVERNMENT and AGREEMENT, have to figure as theoretically primitive concepts in any reasonably detailed account of nominal structures. Explicating government and agreement relations and especially separating one from the other, will presuppose a sufficient inventory of formal descriptive devices in any particular theory of grammar.

The paper is settled in the framework of Head-Driven Phrase Structure Grammar (HPSG). Recent HPSG analyses for the German nominal group that have been put forth in [Pollard and Sag 1994] and [Netter 1994] are studied in detail contrasting them to (semi-) formal proposals from other linguistic frameworks; potential problems as well as some abstract joint properties of the two HPSG approaches are exemplified. Building on this comparison it is concluded that in exactly the linguistic stipulations shared by the two accounts, two important generalizations about the inherent structure of the German nominal group are to be found. At the same time the [Pollard and Sag 1994] analysis is tentatively reformulated.

To my parents and Constanze
— for the encouragement
and true love

Contents

1 Introduction

The German nominal group — even when reduced to the core inventory of nouns, determiners and attributive adjectives — is a morphologically and syntactically complex structure. Accordingly, it is not surprising that the academic dispute on how nominal structures are to be analysed adequately is an issue with a long-standing tradition in the linguistics literature.

Looking at familiar examples like those in (1) to (3), we find a substantial number of morphosyntactic categories that interweave the elements of the nominal group in various syntagmatic relations, each of them individually contributing to the inflectional shape of the whole.

(1) *ein kühles Bier*

(2) *das kühle Bier*

(3) *[der Genuß] kühlen Biers*

In this paper it will be suggested that a detailed understanding of the (morpho-) syntactic categories and the syntagmatic relations exhibited in the core nominal group is a prerequisite to addressing the question whether it is appropriate to think of the German nominal group as a noun or a determiner phrase (i.e. the bone of contention in the so-called NP vs. DP opposition). Although we will not attempt to bring the issue to a firm conclusion (which we doubt is to be found in following either of the two streams of argumentation exclusively), in looking at two fundamentally different analyses advocating the two trains of thought, we will study the systematic covariation of inflectional properties and how it can be accounted for in much detail.

It will be argued that without a fairly high degree of formalization several aspects of the linguistic structure of the German nominal group and how the individual elements relate to each other cannot be adequately captured. In fact, it is claimed, the two essential syntagmatic relations holding within the nominal group, viz. GOVERNMENT and AGREEMENT, have to figure as theoretically primitive concepts in any reasonably detailed account of nominal structures. Explicating government and agreement relations and especially separating one from the other, will presuppose a sufficient inventory of formal descriptive devices in any particular theory of grammar.

This paper has choosen to settle its study of the German nominal group in the framework of Head-Driven Phrase Structure Grammar (HPSG). As HPSG is a linguistic theory and formalism that grew out of computational linguistics research (which, we feel, is an academic subject that yet has to be better acknowledged in its relevance to 'traditional' linguistics), we provide a brief introduction into the basic HPSG ideas and its logical foundations in section 2. In the main part of the paper we review various approaches to syntactic categories and government and agreement relations (including the related concept of syntactic

headship) as they have been suggested in the school of Integrational Syntax (section 3.1) and two meta-theoretical essays by Arnold M. Zwicky (section 3.2), mutually relating them to each other and to the HPSG account of syntagmatic relations.

Section 4 then provides a fairly detailed review of two analyses for German nominals that recently have been suggested in the HPSG literature (viz. by [Pollard and Sag 1994] and [Netter 1994]) and — in a contrastive study — points to potential problems in the two approaches as well as to some abstract joint properties that they have in common. Finally, in the concluding section 5 it is argued that in exactly the linguistic stipulations shared by the two accounts, two important generalizations about the inherent structure of the German nominal group are to be found. At the same time the [Pollard and Sag 1994] analysis is tentatively reformulated to eliminate some of the problematic issues noted in section 4.1. The relevant facts from the German distributional data have been incorporated into the reviews of HPSG analyses for the German nominal group where it was deemed appropriate.

The work underlying this paper was carried out in the environment of the computational linguistics projects DISCO (Dialogue System for Autonomous Cooperating Agents) and VerbMobil hosted at the German Research Center for Artificial Intelligence (DFKI) in Saarbrücken.

The paper is based on a manuscript that was accepted for a *magister artium* (M.A.) degree at the department for German Studies of the Free University Berlin in autumn 1993. When originally writing the thesis, it was almost exactly two years since I had moved to Saarbrücken to get into doing it; however, being involved in ongoing project activities, it took me far longer than originally planned to actually start thinking about it. Nevertheless, I would not have wanted to miss the practical experience and exciting stimuli that arise from time to time in applying theoretical grammatical knowledge to a functional and implemented natural language processing system. During the process of writing the thesis I especially learned to value and highly regard the cooperativity and friendship among the members of the DISCO and VerbMobil projects.

This paper owes a lot to discussions with Klaus Netter and especially to the constructive criticism of Walter Kasper. I am most grateful to John Nerbonne for the constant encouragement and to Andrew P. White for the laborious proof reading (giving him the opportunity to read such brilliance). I appreciate the support and never ending patience of Peter Eisenberg and Hans Uszkoreit in teaching me the fundamentals of linguistics, backing my thesis and hosting me in Saarbrücken. Furthermore, my thanks go to the Saarbrücken colleagues and friends for freeing me from everyday obligations and finally making it happen. I acknowledge the genius of Don Knuth for designing TEX — the marvellous typesetting system employed in writing this paper.

2 HPSG — a State-of-the-Art Unification Grammar

HPSG has gained a predominant position in the area of so-called feature structure based or unification grammars, the family of linguistic formalisms and theories that has received its name due to its common data type and the one fundamental operation — the unification of complex feature structures.

With its clear distinction between the underlying descriptive formalism and the linguistic theory itself (as it is outlined in [Pollard and Sag 1987] and refined in [Pollard and Sag 1994]), HPSG goes well beyond some rather formalism-type frameworks like Functional Unification Grammar (FUG) [Kay 1986] or PATR-II [Shieber *et al.* 1983] in the degree of formality, but at the same time has surpassed most (if not all) of its more theory-driven unification-based predecessors — e.g. Generalized Phrase Structure Grammar (GPSG) [Gazdar *et al.* 1985], Lexical Functional Grammar (LFG) [Kaplan and Bresnan 1982] or Categorial Unification Grammar (CUG) [Uszkoreit 1986] — in the number and range of linguistic phenomena it has been applied to.

The following sections will give a short introduction into the overall epistemological setup of HPSG (section 2.1), its formal foundations in typed feature logic (2.2), the core inventory of the theory (2.3) and an overview of a practical and implemented HPSG system, including some diversions from the 'standard' theory making it both better applicable to German grammar and computationally more tractable (2.4). This introductory part of the paper is intended to provide the HPSG novice with the formal prerequisites to follow the discussion of the specific phenomena from German syntax (within the HPSG framework) in the sections to come. Still, some basic interest and knowledge of formal grammar theories will be presupposed presently.

2.1 HPSG: A System of Signs

A common understanding of linguistic theory is to aim for an abstract MODEL of natural language phenomena. Accordingly, the overall picture of HPSG as one particular theory of natural language grammar, basically, falls into three parts: (i) the EMPIRICAL DOMAIN, (ii) the MODELLING DOMAIN and (iii) the FORMAL SYSTEM deployed in the specification of the modelling domain (see figure 1). In the following paragraphs we will briefly characterize each of the three HPSG components.

Very generally, the empirical domain of HPSG is to be understood as "the universe of possible linguistic objects, [...] the system of linguistic types, that makes communication possible" [Pollard and Sag 1994, xiv and xxii][1]. However, as the HPSG theory of grammar

[1] Both [Pollard and Sag 1987] and [Pollard and Sag 1994] intentionally leave the question on the underlying ontology of linguistic objects unanswered.

Although the central notion of SIGN (see presently) employed in HPSG might suggest a conceptualist view roughly in the Saussurean sense (conceive of signs as mental objects — associative links between *significant* and *signifié*), the converse holds for the incorporation of ideas from Situation Semantics

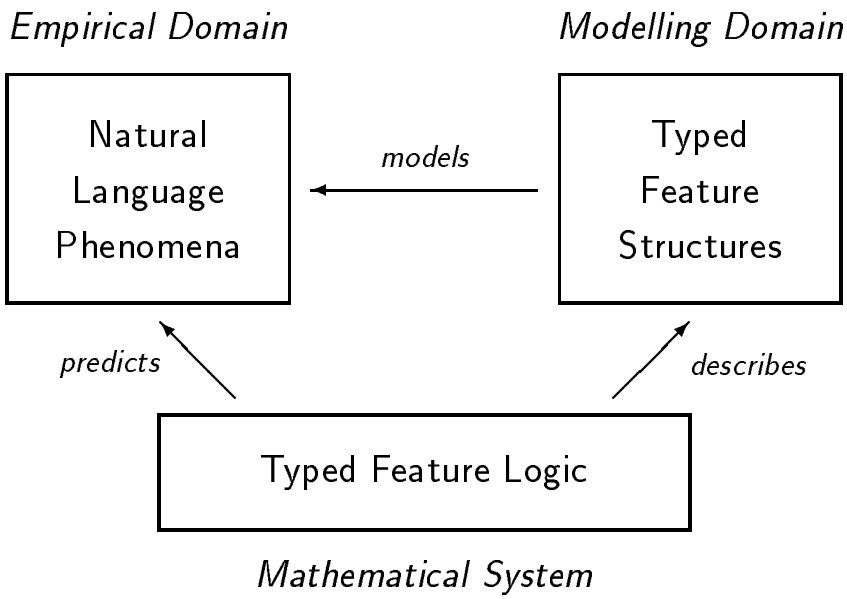


Figure 1: Three-component outline of HPSG as a formal theory; including the relations postulated between the empirical and modelling domains and the mathematical system deployed in linguistic specification (this diagram being shamelessly borrowed from [Pollard and Sag 1994, xvi]).

does NOT attempt to directly characterize the entities from the empirical domain of natural language phenomena (but instead gives a precise and formal characterization of the entities in its modelling domain), there will be little to say on this aspect of linguistic theory in the following.

Ideally, there should be a one-to-one mapping between objects from the empirical domain and the objects postulated in the modelling domain of the theory (and likewise for all parts of these entities). HPSG being a well formalized theory, has chosen to settle its model of types of linguistic objects on TYPED FEATURE STRUCTURES — mathematical objects that (can) have complex structure and contain several distinct levels of linguistic abstraction — as its basic units.

For the feature structures populating the HPSG modelling domain the notion of SIGNS has been adopted[2], which are to be taken as "structured complexes of phonological, syn-

(i.e. the explication of linguistic meaning as a correlation between non-mental (utterance) events and real-world objects, properties or situations) into the theory.

Diplomaticly [Pollard and Sag 1994, xxiii] conclude: "Our concern [...] will be with the internal architecture of the system that linguistic types form, not with that system's ultimate ontological status."

[2] According to [Pollard and Sag 1987, 2 – 6], there is a weak correspondence from the HPSG domain of signs to the Saussurean conception of *langue* as a system of *signes* linking together what nowadays might be called some sort of phonological structure (be it a mental sound image or a physical utter-

tactic, semantic, discourse and phrase-structural information" [Pollard and Sag 1994, 1]. HPSG signs are assumed to model the elementary objects forming a system that is shared knowledge in a "linguistic community" [Pollard and Sag 1994, xxii] (a language) allowing the members of a particular community to communicate through the exchange of INFORMATION.

HPSG signs are the basic units bearing and (used in) conveying information. Attributes in the modelling feature structures hence are to be interpreted as properties of the object that they are contained in, e.g. phonology, syntax, semantics et al. are properties of the top-level object[3], the sign itself, whereas case, number or gender, say, may be deeply embedded into other structures denoting properties of some (syntactic or semantic) subpart of a sign.

Finally, the purpose of the formal system underlying the HPSG theory of grammar (an instance of TYPED FEATURE LOGIC — see presently) is to give a sufficiently formal characterization of the system making up the modelling domain, viz. the various types of feature structures postulated, compositional principles used in building up complex structures from more basic ones, constraints on the well-formedness (mostly in terms of syntactic properties) of signs et al. HPSG presumably among contemporary theories of grammar exhibits the highest degree of descriptive uniformity and representational parsimony in requiring all entities of its modelling domain to to satisfy specifications within the logic of typed feature structures, thus excluding additional theoretical devices like meta-rules (in GPSG), functional uncertainty (especially in LFG) or movement operations (in transformational grammar). By their specifying the objects in the modelling domain, the stipulations making up the (formal) system of the linguistic grammar theory simultaneously are predictions on the empirical domain, viz. the given natural language phenomena at hand. This basically is the nature of the intended (one-to-one) modelling relation between the types of linguistic objects in the empirical and the hypothesized types of feature structures in the HPSG modelling domain.

2.2 Typed Feature Logic and Unification

Informally speaking, feature structures are sets of attribute-value pairs with attributes denoting (names of) properties of linguistic objects (see above) and their values typically ranging over atomic symbols (atomic feature structures) and, recursively, complex feature structures themselves. Additionally, it has become common practice to allow for disjunctions and lists or sets of feature structures as supplementary data types. In the various unification based frameworks the terms F-STRUCTURE (LFG), FEATURE BUNDLE,

ance event, i.e. an acoustic wave) with some kind of meaning representation (potentially including pragmatic aspects of the use of a particular type of sign).

[3] This, at least, holds for the feature structure geometry given in [Pollard and Sag 1987] which has undergone some rearrangement in [Pollard and Sag 1994]. In section 2.4 we will see that the Saarbrücken HPSG system as the framework for this paper chooses yet another layout for the top-level attributes of signs.

ATTRIBUTE VALUE MATRIX, FUNCTIONAL STRUCTURE (FUG), TERM (DCG) or DAG have been in use more or less synonymously.

Conceiving of feature structures as bearing information (and, hence, denoting sets of linguistic entities) yields a partial ordering on their informational content; e.g. (1) can be said to contain more information than (2) (thus (2) potentially denotes a larger set of entities than (1)) because it is more specific on the properties of the value of CASE. This (partial) ordering on feature structures has become known as the SUBSUMPTION relation (formally written as '$\sqsupseteq$'), i.e. it is said that (2) subsumes (1) ((2) $\sqsupseteq$ (1)) as it has less informational content (is less specific).

(1) $\begin{bmatrix} \texttt{CASE} \begin{bmatrix} \texttt{OBL} & - \\ \texttt{GOV} & - \end{bmatrix} \end{bmatrix}$

(2) $\begin{bmatrix} \texttt{CASE} \begin{bmatrix} \texttt{GOV} & - \end{bmatrix} \end{bmatrix}$

Naturally, subsumption does NOT hold between any arbitrary pair of feature structures. Feature structures containing conflicting information are said to be incompatible (denoting the empty set) and, in turn, there may be pairs of feature structures containing information on mutually unrelated properties. — E.g. the feature structure in (3) is incompatible with (1) because it contains conflicting information for the OBL feature but at the same time is simply unrelated to (2) because it spells out a different property for the CASE value.

(3) $\begin{bmatrix} \texttt{CASE} \begin{bmatrix} \texttt{OBL} & + \end{bmatrix} \end{bmatrix}$

Even in the vanilla-flavoured feature structure examples given so far two fundamental notions from the logic of feature structures have been suggested already, viz. (i) the idea of partial information structures and (ii) the concept of informational compatibility.

Examples (1) – (3) above can be interpreted as information on the CASE value of linguistic objects (i.e. they are entities in the HPSG modelling domain[4]. that will typically be embedded as substructures into the morphological or syntactic properties of a sign)[5]. The major advantage of using a pair of binary features instead of the four atomic case values

[4] In the sections to come the fine-grained distinction between the HPSG modelling domain and the postulated correspondence to the empirical domain of natural language phenomena will mostly be lost.

Except where explicitly stated (or clear from the context) all reference will be to HPSG signs (or parts of them), i.e. typed feature structures, the elementary class of objects in the HPSG grammar model.

[5] The partioning of CASE into two binary attributes OBL and GOV originally is due to Manfred Bierwisch and has been taken up (among others) in [Zwicky 1986], [Wunderlich 1988] and [Netter 1994].

(for German) is that it allows to refer to (natural) classes of CASE values (for the linguistic motivation see presently) by simple UNDERSPECIFICATION of CASE properties.

For instance, assuming a distribution of features as in

(4)

GOV \ OBL	−	+
−	*nominative*	*genitive*
+	*accusative*	*dative*

the feature structures in (2) and (3) denote the classes of *nominative* or *genitive* case (the top row in the table) and *genitive* or *dative* case (the right column) respectively.

Although a similar class building mechanism could be encoded in disjunctive feature specifications like the one in (5) as well,

(5) $\left[\texttt{CASE}\left\{\textit{nominative}\ \ \textit{genitive}\right\}\right]$

there are both linguistic and technical reasons to prefer a direct encoding using atomic properties over a disjunctive one, viz. that (i) disjunctive specifications are to be interpreted as missing generalizations on the properties involved and (ii) the processing of disjunctive feature structures typically causes inefficiency in the unification algorithm because it involves extensive backtracking.

Now UNIFICATION, the primary operation on feature structures, is the process of merging compatible information from two (or more) structures into a single object. Hence, unifying the feature structures (2) and (3) (written as (2) $\sqcap$ (3)) will collect the information for the GOV and OBL attributes into

(6) $\left[\texttt{CASE}\left[\begin{array}{l}\texttt{OBL}\ +\\ \texttt{GOV}\ -\end{array}\right]\right]$

Somewhat more formally, the unification of two feature structures D_1 and D_2 is the least specific feature structure D_0 subsumed by both, D_1 and D_2 (thus, unification corresponds to the intersection operation on the sets of objects denoted by D_1 and D_2)[6].

The universe of feature structures under the subsumption relation can be thought of as a (semi-) lattice stretching between two special concepts known as TOP and BOTTOM (written as $\top$ and $\bot$ respectively), with $\top$ being the most general feature structure (containing no

[6] For a formal introduction to the logic of (typed) feature structures see [Shieber 1986] and especially [Carpenter 1992].

information, i.e. denoting the set of all objects) and $\perp$ being the most specific structure containing inconsistent information (thus denoting the empty set).

Using $\top$ and $\perp$, the following equations will hold for arbitrary feature structures D

(i) $\top \sqsupseteq D \sqsupseteq \perp$ ($\top$ subsumes everything; everything subsumes $\perp$)

(ii) $D \sqcap D = D$ (unification is idempotent)

(iii) $D \sqcap \top = D$ ($\top$ is a neutral element)

(iv) $D \sqcap \perp = \perp$ ($\perp$ is a constant element)

and, likewise, the unification of incompatible structures *per definitionem* is $\perp$, because this is the only structure allowed to contain conflicting information.

Finally, a concept generally playing an important role in unification grammars (but especially in the HPSG account of syntagmatic relations like government and agreement) is that of REENTRANCY or COREFERENCE of feature structures. Coreference of two feature structures is to be understood as (a model of) actual token identity of objects as opposed to mere type identity.

For instance, looking at (7) and (8), we find that (7) is more specific than (8) because it constrains the values of CASE in the morphology and CASE as a syntactic HEAD feature (see presently) to be reentrant (token identical) whereas in (8) the two attributes simply happen to have compatible (type identical) values.

(7)
$$\begin{bmatrix} \text{MORPH} & \begin{bmatrix} \text{INFL} & \begin{bmatrix} \text{CASE} & \boxed{1} \begin{bmatrix} \text{OBL} & - \\ \text{GOV} & - \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{SYN} & \begin{bmatrix} \text{LOC} \mid \text{HEAD} & \begin{bmatrix} \text{CASE} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

(8)
$$\begin{bmatrix} \text{MORPH} & \begin{bmatrix} \text{INFL} & \begin{bmatrix} \text{CASE} & \begin{bmatrix} \text{OBL} & - \\ \text{GOV} & - \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{SYN} & \begin{bmatrix} \text{LOC} \mid \text{HEAD} & \begin{bmatrix} \text{CASE} & \begin{bmatrix} \text{OBL} & - \\ \text{GOV} & - \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

A tentative interpretation of the reentrancy relation found in (7) actually is as a generalization on lexical items, maintaining that morphological case marking determines syntactic properties of word forms and their phrasal projections, e.g. in the case of nominal structures.

So far, we have only been looking at untyped feature structures as complex, structured bundles of information (with, maybe, the basic distinction between atomic and complex feature structures). However, from the idea of organizing the universe of feature structures into a lattice stretched between $\top$ and $\bot$ (partially ordered by the subsumption relation) it is not really a large step to think of the same set of objects as a type lattice organized according to (possibly multiple) inheritance relations[7]. Types in this sense are to be intuitively understood as names for classes of linguistic objects satisfying the (feature structure) constraints that are associated with a particular type. Accordingly, in a typed universe of feature structures the subsumption relation corresponds to the notion of supertypes (resulting from type inheritance specifications) and the unification of two typed feature structures ϕ and ψ yields the most general type ω in the type lattice that is a subtype to both ϕ and ψ (i.e. ω is the greatest lower bound of ϕ and ψ).

Besides the greater expressive power of a TYPED feature logic the major advantage of associating a type lattice with the domain of feature structures is that type inheritance can give an appropriate account of the relations holding between the various kinds of linguistic objects. Additionally, by having the properties (i.e. feature names) of some object being defined for its respective type there is an appropriateness condition on attributes of linguistic objects and on the type of value a given attribute may have, i.e. for a given linguistic object from its associated type we know (i) which properties are defined to be appropriate for it (i.e. will eventually be specified); and (ii) what the domain (value restriction) for any of these properties is.

Slightly anticipating the sections to come, let us consider the specification of the most basic HPSG type, the *sign*, and some of its subtypes. As has been outlined in section 2.1 HPSG signs (can) contain phonological, syntactic, semantic, pragmatic and phrase-structural information. Because the notion of phrase structure, however, can only be meaningfully applied to non-lexical objects, the type *sign* is partioned into the two subtypes *word* and *phrase*[8]. Assuming the top-level feature geometry of the Saarbrücken HPSG implementation for the moment[9] the type specifications in (9) – (11) straightforwardly reflect these facts (see figure 2 for the corresponding type lattice):

(9) $$sign \equiv (word \sqcup phrase) \sqcap \left[\texttt{CAT}\ category\left[\ \right]\right]$$

(10) $$word \equiv sign$$

[7] In [Pollard and Sag 1987] and [Pollard and Sag 1994] the term SORT is used for what is called a type in this paper. Often both names have been used in the literature non-distinctively. However, the Saarbrücken HPSG system (see presently) reserves the term sort for atomic objects that undergo closed-world reasoning whereas the general type lattice follows open-world assumptions.

[8] [Pollard and Sag 1987] originally have suggested the names *lexical-sign* and *phrasal-sign* but — probably for typesetting reasons — *word* and *phrase* have become equally common in the HPSG literature.

[9] See section 2.4 for the motivation to diverge from the feature structure layout used in [Pollard and Sag 1994].

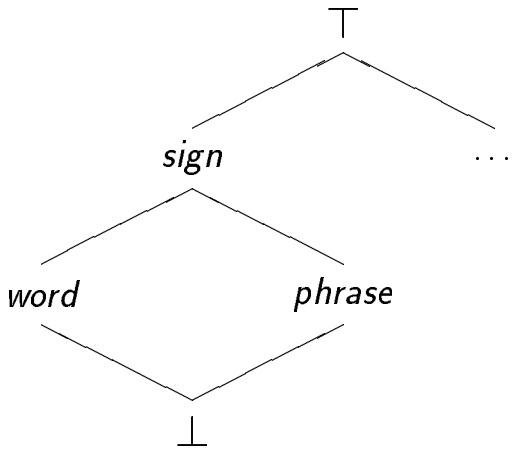


Figure 2: Type lattice resulting from the HPSG specifications for *sign* and its subtypes ((9) to (11)); note that the incompatibility between *word* and *phrase* results from the partioning declaration (using '⊔') in (9).

$$(11) \quad \mathit{phrase} \equiv \mathit{sign} \sqcap \left[\texttt{DTRS}\ \mathit{daughters}\right]$$

Basically, (9) defines *sign* to have the single attribute `CAT` with its value being restricted to feature structures of the type *category* (the linguistic category of some object). Following, (10) makes *word* a subtype of *sign* (thus having it inherit the `CAT` attribute) introducing no additional properties (except, maybe, its more specific type), while (11) defines *phrase* as another subtype of *sign* containing the additional `DTRS` feature which itself is restricted in its value to objects of type *daughters* (i.e., as we will see, tree representations of phrase structure.

2.3 Some HPSG Essentials

HPSG is a truly lexicalized theory of grammar, in that the lexicon is designed as the primary linguistic knowledge base interacting with a set of general wellformedness principles and a very small number of highly abstract phrase structure schemata.

At the same time HPSG qualifies as a non-derivational grammar conception, in that there are no transformations or a concept of movement employed. Instead, the crucial mechanism in accounting for the relations holding between signs or substructures of signs, e.g. in modelling of government and agreement relations, assignment of thematic roles or coindexation, is that of structure sharing

Section 2.2 gave a preview of the gross feature structure geometry of HPSG signs with `CAT` and `DTRS` (for phrasal signs only) as the top-level attributes. The feature `CAT` is restricted in its value to objects of type *category* thus embedding all the information that

contributes to the category of a linguistic token (be it a single lexical entry without any phrase structure or a whole sentence containing a large number of signs recursively nested in its DTRS feature).

As has already been pointed out, HPSG basically accounts for phonological, syntactic, semantic and pragmatic properties of categories by having each of them contribute a particular feature with appropriate value restrictions to the type *category*. Yet, there has been little work on the incorporation of a formal theory of morphophonology into HPSG up until today (but see [Krieger *et al.* 1993] for an attempt to encode finite state morphology in feature structures); therefore, HPSG PHON values are usually orthographically glossed as unstructured lists intended to simply represent the surface string of some word or phrase[10].

Except where it will be argued that a syntactic theory of agreement is to be preferred over the inherently semantic account given in [Pollard and Sag 1994] for the phenomena exhibited in German nominals, we will not be concerned with semantic or pragmatic properties of categories presently. Hence, the focus of interest in this paper will be the remaining *category* attribute SYN containing the syntactic information of signs as structured objects of type *syntax*.

Based on a simplified example, the remaining part of this section will introduce the relevant properties of *syntax* objects and the related HPSG grammar principles; (12) gives (parts of) the syntactic category of the German transitive verb (form) *sieht* (sees)[11]:

---

[10] As we will see in section 2.4, in the Saarbrücken HPSG system the attribute MORPH has been substituted for the usual HPSG PHON, mainly to interface the grammar and lexicon to a separate two-level morphological component.

In this particular case the concept of linguistic types (as in the value restriction on the MORPH feature) serves an additional purpose, viz. as a formal interface specification between the type lattice forming the syntactic and semantic parts of the grammar and a dedicated morphological module internally using a completely different data type (finite state automata).

[11] According to [Pollard and Sag 1994] (12) is merely a FEATURE STRUCTURE DESCRIPTION than a feature structure in its own right because it is only a PARTIAL model of the category of a German word. Accordingly, feature structures are required to be TOTAL models of linguistic entities in that they be both totally well-typed and sort-resolved. With reference to [Carpenter 1992], [Pollard and Sag 1994] require feature structures to be of most-specific sorts (types in our terminology), i.e. leafs in the type lattice, and for any given sort to contain all the attributes that have been defined appropriate for that sort.

Clearly, the distinction between feature structures themselves and feature structure descriptions has consequences for the resulting linguistic ontology (the world of feature structures) but, nevertheless, is not that significant in the description of concrete natural language phenomena. As it is common practise to use partial and underspecified structures in the HPSG literature (e.g. by ignoring attributes that are clear from the context or simply unrelated to some topic at hand), this paper will not attempt to build on the feature structure vs. feature structure description distinction.

Still, however, it is worth bearing in mind that it is part of a formalized version of HPSG to explicitly construct its type lattice with appropriate features and value types.

(12)
$$\textit{category}\left[\begin{array}{l}\texttt{PHON}\ \langle \textit{sieht}\rangle \\ \texttt{SYN}\ \left[\textit{syntax}\left[\begin{array}{l}\texttt{LOC}\ \left[\textit{local}\left[\begin{array}{l}\texttt{HEAD}\ \textit{verb}\left[\begin{array}{l}\texttt{VFORM}\ \textit{fin}\\ \texttt{TENSE}\ \textit{pres}\end{array}\right] \\ \texttt{SUBCAT}\ \langle \text{NP[nom] NP[acc]}\rangle\end{array}\right]\right]\end{array}\right]\end{array}\right]$$

Objects of type *syntax* basically fall into two parts — local and non-local information — where LOC is to be understood as the set of properties local to some category and NON-LOC is the home of information that is shared between signs in constructions that are traditionally characterized as unbounded dependencies (because they typically extend a single head domain). Again, as all syntactic structure in German nominals predominantly is of a strictly local nature, there will be nothing to say on the non-local properties of syntactic categories in the following sections.

As with CAT and SYN the attribute LOC is restricted in its value to objects of a certain type (viz. *local*) which, likewise, comprises the set of properties that are appropriate for the local information in syntactic categories. Looking at (12), we find the features HEAD and SUBCAT.

Based on the assumption that all fundamental phrase structure has one constituent serving as its head and comprising its HEAD information all along the projection line, HPSG HEAD features (sort of resembling GPSG in this respect) play a very crucial role in the account of constituent structure. Therefore, there is a dedicated wellformedness principle on all HPSG (phrasal) signs enforcing the percolation of HEAD properties, viz. the

> Head Feature Principle (HFP)
>
> The HEAD value of any headed phrase is structure-shared with the HEAD value of the head daughter. [Pollard and Sag 1994, 24]

In the epistemological conception of grammar HPSG wellformedness constraints are interpreted as linguistic universals (possibly employing parameters for cross language variation) such that any token linguistic object is required to be compatible with them[12]. Thus, the

(13)
$$\textit{phrase}\left[\texttt{DTRS}\ \textit{headed-structure}\left[\ \right]\right] \Longrightarrow \left[\begin{array}{l}\texttt{CAT}\ \left[\texttt{SYN}\,|\,\texttt{LOC}\left[\texttt{HEAD}\ \boxed{1}\right]\right] \\ \texttt{DTRS}\left[\texttt{H-DTR}\,|\,\ldots\,|\,\texttt{LOC}\left[\texttt{HEAD}\ \boxed{1}\right]\right]\end{array}\right]$$

[12] In [Pollard and Sag 1987] HPSG wellformedness constraints are formalized as implications on the type *phrase*, e.g.

resulting grammar can be formalized as the conjunction of all the principles with the disjunction of the lexicon and the phrase structure schemata (see presently), so that for some language L (German, say) it demands that

$$(14) \quad L \equiv \mathrm{P}_1 \sqcap ... \sqcap \mathrm{P}_l \sqcap (\mathrm{L}_1 \sqcup ... \sqcup \mathrm{L}_m \sqcup \mathrm{S}_1 \sqcup ... \sqcup \mathrm{S}_n)$$

Intuitively speaking, (14) requires that in L any entity in the HPSG modelling domain will fulfill ALL the wellformedness principles and, at the same time, be an instance of (at least) one lexical entry or a phrase structure schemata.

The second *local* feature (besides `HEAD`) found in (12) is `SUBCAT`, the locus of SUBCATEGORIZATION information. As we will see, the HPSG `SUBCAT` mechanism serves a number of different purposes in the theory, the basic idea being that the list-valued `SUBCAT` feature of lexical entries and (yet) unsaturated phrases specifies the VALENCE of these signs.

Without anticipating the discussion of the HPSG model of licensing and government in the context of German nominals (see section 3) it should be clear from example (12) that the `SUBCAT` list encodes (at least) two distinct bits of information on the combinatoric potential of the German verb form *sieht*, viz. (i) the number of arguments (complements in usual HPSG terminology, including the subject) it takes; and (ii) the selectional restrictions imposed on them.

Note that the category symbols with abbreviated properties in square brackets found in (12) (e.g. 'NP[nom]') are to be taken as a shorthand notation for complex feature structures, e.g. NP[nom] for a *category* type object with a nominal head (a `HEAD` value of appropriate type) that qualifies as a saturated or maximal projection (which, in standard HPSG, is taken to comprise an empty `SUBCAT` list). In turn, the 'nom' in square brackets abbreviates an attribute value somewhere embedded in the category (`SYN`$|\ldots|$`CASE` for this example) where the appropriate feature has to be uniquely identifiable from the value itself and the context.

Assuming the *case* partioning from section 2.2 and the standard HPSG `HEAD` feature geometry for the moment[13] together with the aforementioned (minimal) condition on saturation[14], NP[nom] is then to be understood as follows[15]:

However, as [Netter 1993] has pointed out a maybe more straightforward approach in encoding universal phrase structure principals is to have them be part of the grammar type lattice and simply be inherited by *phrase* and all of its subtypes.

[13] In section 3.4 it will be argued, though, that in a syntactic account of German agreement phenomena it is reasonable to have an additional level of abstraction on `HEAD` features, viz. the grouping of properties involved in agreement relations into an `AGR` or `INFL` attribute.

[14] [Netter 1994] shows that the saturation of the `SUBCAT` list (or similar selectional features — see presently) is insufficient in defining the maximal projections of nominal heads. Therefore, this part of (15) will have to undergo revision in the forthcoming sections, too.

[15] Here, as in all following feature structure descriptions, the type symbols (lower left subscripts to feature structures) will often be left out when clear from the context.

(15)
$$\begin{bmatrix} category \\ \text{SYN} \begin{bmatrix} \text{LOC} \begin{bmatrix} \text{HEAD} & noun \begin{bmatrix} \text{CASE} \begin{bmatrix} \text{OBL} & - \\ \text{GOV} & - \end{bmatrix} \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

— Yet, there is more information (implicitly) encoded in the *list* data type in the structure $\left[\text{SUBCAT}\ \langle \text{NP[nom]}\ \text{NP[acc]} \rangle\right]$ — viz. a (partial) ORDERING RELATION ON COMPLEMENTS.

Originally, [Pollard and Sag 1987] organized SUBCAT lists according to an obliqueness relation (roughly, on grammatical functions) which, in [Pollard and Sag 1994, Chapter 6], is taken to substitute for the notion of c-command in the HPSG binding theory. However, there is a long-standing discussion on how to obtain the hierarchy of complements (with its relation to the surface word order) and whether to include subjects into the list of arguments at all.

As [Pollard and Sag 1994] themselves (in the final chapter slightly contradicting the rest of the book) speak in favour of a separate selectional mechanism for subjects (by means of a supplementary *local* feature SUBJ) and, at the same time, there have been various suggestions to have SUBCAT be set-valued (at least) for non-configurational languages like German, we will leave the question on how exactly to interpret the order of SUBCAT lists open for now. Fortunately, there seems to be nothing in German nominal phrase structure strictly depending on the one or other analysis, so we will feel free to adopt whatever selectional mechanism appears to be most appropriate.

As an introduction to HPSG essentials, however, it has to be noted that traditionally the subject is taken to occupy the first position in a SUBCAT list, followed by the direct object and possibly, indirect or prepositional objects.

Again, there is a dedicated wellformedness principle maintaining the cancellation of SUBCAT elements in constituent structures containing a head daughter and one or more COMPLEMENT DAUGHTERS, the

> SUBCATEGORIZATION PRINCIPLE (SP)
>
> In a headed phrase (i.e. a phrasal sign whose DTRS value is of type *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) of CAT values of the complement daughters[16].

Having the elements of SUBCAT be *category* objects instead of signs enforces what in [Pollard and Sag 1987] had to be made a principle: the locality of selection from heads.

[16] Diverging from the original wording in [Pollard and Sag 1994, 24], CAT has been substituted for SYNSEM in order to meet the top-level feature geometry assumed in this paper.

If heads were subcategorizing for entire signs, via the DTRS feature there would be unrestricted access to the constituent structure of complements so that in principle there was no boundary for government; heads in fact would be free to impose selectional restrictions on complements — as well as on other heads or modifiers — deeply embedded into the phrase structure of their own complements. As unbounded government clearly is linguistically undesirable, the top-level feature distribution STRUCTURALLY reflects the linguistic stipulation that selectional restrictions be limited to the domain of a single head projection.

Though [Pollard and Sag 1994] avoid to give a feature structure representation of the subcategorization principle, it should be clear that its formalization requires additional descriptive devices in the logic of typed feature structures not mentioned in section 2.2. Both the list concatenation and the indirection from DTRS values to their respective categories have to be encoded as RELATIONAL DEPENDENCIES on feature structures, e.g. the concatenation operation as a relation holding between three lists l_1, l_2 and l with $l = \texttt{append}(l_1, l_2)$ (but see section 2.4 for a variant that gets by without relational constraints).

Before looking at the HPSG phrase structure account there are still two essential concepts remaining to be introduced, viz. (i) the organisation of the top-level DTRS attribute and (ii) the role of phrase structure schemata as very abstract grammar rules.

Depending on the kind of construction, HPSG distinguishes a set of subtypes of *daughters* that introduce appropriate features to give a local tree-like representation of the immediate daughters (the entire *sign* objects) of a particular phrasal sign. The subtypes to *daughters* are, roughly speaking, structured according to an abstract notion of grammatical functions (in the sense of distinguishing arguments from modifiers, functional elements et al. but not, say, different types of objects). In looking at the structure of German nominals we will exclusively be concerned with two of the subtypes assumed in [Pollard and Sag 1994]: *head-complement-structure* and *head-adjunct-structure*.

As the type names suggest, both constructions agree in that they contain a head (the value of the H-DTR attribute), thus qualifying as headed structures (see above), but differ in the type of additional constituents the head combines with. In a *head-complement-structure* the head binds one or more of its arguments, i.e. signs subcategorized for by their categories via the SUBCAT mechanism. As we have seen in the discussion of the subcategorization principle already, complement daughters are represented in a list valued *daughters* feature C-DTRS.

Departing from [Pollard and Sag 1987] (where a head is marked for a set of modifiers it can possibly combine with), [Pollard and Sag 1994] assume that adjuncts are selecting the signs they modify through the *category* valued HEAD feature MOD. Accordingly, adjuncts form their own class of daughters in that they neither are heads nor being licensed (subcategorized for) by heads, which is accounted for in a second *sign* valued A-DTR attribute in structures of the type *head-adjunct-structure*.

Summing up the phrase structure type definitions introduced so far we get[17]

[17] Note that the meta-syntactic $_{sign}[]^+$ in (18) is to be read as the sequence of any number greater or

(16) $\mathit{daughters} \equiv \mathit{headed\text{-}structure} \sqcup \ldots$

(17) $\mathit{headed\text{-}structure} \equiv \left[\texttt{H-DTR}\ \mathit{sign}\left[\ \right]\right]$

(18) $\mathit{head\text{-}complement\text{-}structure} \equiv \mathit{headed\text{-}structure} \sqcap \left[\texttt{C-DTRS}\ \left\langle \mathit{sign}\left[\ \right]^{+}\right\rangle\right]$

(19) $\mathit{head\text{-}adjunct\text{-}structure} \equiv \mathit{headed\text{-}structure} \sqcap \left[\texttt{A-DTR}\ \mathit{sign}\left[\ \right]\right]$

Now, the final part missing in our subset of HPSG that will be necessary in looking at the syntactic structure of German nominals are the combinatory rules or schemata used in building up phrase structure from lexical signs. As HPSG attempts to encode the major parts of linguistic knowledge in the (structured) lexicon, [Pollard and Sag 1994] assume no more than a small number of heavily underspecified phrase structure or immediate dominance schemata, some of them having more or less direct counterparts in the framework of Government and Binding theory and its $\overline{\text{X}}$-schemata.

As with the selection of subjects there is some discussion in the HPSG literature on the number and structure of immediate dominance schemata; and, similar to some other HPSG universals, these may be subject to cross language variation and parameterization. Nevertheless, schemata (20) and (21) are fairly uncontroversial and can be understood as the HPSG equivalents of the common $\overline{\text{X}}$-schemata (22) and (23) respectively:

SCHEMA 1 (HEAD + SUBJECT)

(20) $$\left[\begin{array}{ll}\texttt{CAT} & \left[\texttt{SYN}\,|\,\texttt{LOC}\,|\,\texttt{SUBCAT}\ \langle\rangle\right] \\ \texttt{DTRS} & \mathit{head\text{-}complement\text{-}structure}\left[\begin{array}{ll}\texttt{H-DTR} & \mathit{phrase}\left[\ \right] \\ \texttt{C-DTRS} & \left\langle\left[\ \right]\right\rangle\end{array}\right]\end{array}\right]$$

SCHEMA 2 (HEAD + COMPLEMENTS)

(21) $$\left[\begin{array}{ll}\texttt{CAT} & \left[\texttt{SYN}\,|\,\texttt{LOC}\,|\,\texttt{SUBCAT}\ \left\langle\left[\ \right]\right\rangle\right] \\ \texttt{DTRS} & \mathit{head\text{-}complement\text{-}structure}\left[\texttt{H-DTR}\ \mathit{word}\left[\ \right]\right]\end{array}\right]$$

equal to one (but NOT zero) objects of type *sign*. As a direct consequence of this definition, head complement or head adjunct structures comprising no more than a single head daughter can not be wellformed. Hence, unary branching structures (in case they were desirable) would have to be accounted for as an additional subtype of *daughters* or *headed-structure*.

(22) $\overline{\overline{X}} \longrightarrow \overline{\overline{Y}} \quad \overline{X}$

(23) $\overline{X} \longrightarrow X \quad \overline{\overline{Y}}$

Roughly speaking, (20) (known as Schema 1 in [Pollard and Sag 1994] and Rule 1 in [Pollard and Sag 1987]) combines a phrasal head that has satisfied all its subcategorization requirements except for its least oblique argument with its subject (or specifier in GB terminology) thus subsuming a class of traditional phrase structure rules like those in (24) or (25)

(24) $S \longrightarrow NP \quad VP$

(25) $NP \longrightarrow Det \quad \overline{N}$

Similarly, (21) allows a lexical head to bind all of its complements except for the subject, again standing for a set of rules of the type found in (26) to (28)

(26) $VP \longrightarrow V \quad NP$

(27) $VP \longrightarrow V \quad NP \quad NP$

(28) $VP \longrightarrow V \quad NP \quad PP$

It has already been mentioned that in the HPSG conception of grammar universal principles, immediate dominance schemata and lexical information interact in that all wellformed signs have to satisfy each of the principles and be licensed by at least one lexical entry or one phrase structure schema, where both properties, i.e. satisfaction of the principles and licensing, are defined through the subsumption relation on typed feature structures. In fact, because the entire grammatical knowledge is encoded in feature structures, all information in a token linguistic object — a sentence, say — is the result of the strict monotonic cumulation of constraints from the various parts of the grammar, i.e. the unification of the HPSG principles with compatible lexical data and appropriate immediate dominance schemata licensing the phrase structure[18].

Consider (29) as a straightforward example. (29) is a phrasal sign containing five constituents (including itself) resulting from combining Schema 1 with Schema 2 (as the embedded `H-DTR`), three lexical items and the head feature and subcategorization principles

[18] Although from time to time it is hard to avoid the allusion to specific processing regimes (i.e. parsing or generation), it should be clear that there is nothing inherently directional in the feature structure formalism or the HPSG theory of grammar. Because of the monotonicity property of the unification operation it is truly irrelevant in what order the combination of constraints, i.e. the unification of feature structures, is carried out.

$$
\left[\begin{array}{l}
\textsc{cat}\ \left[\textsc{syn}\,|\,\textsc{loc}\ \left[\begin{array}{l}\textsc{head}\ \ \boxed{1}\\ \textsc{subcat}\ \langle\rangle\end{array}\right]\right]\\[2ex]
\textsc{dtrs}\ \left[\begin{array}{l}\textsc{h-dtr}\ \left[\begin{array}{l}\textsc{cat}\ \left[\ldots\ \left[\begin{array}{l}\textsc{head}\ \ \boxed{1}\\ \textsc{subcat}\ \langle\boxed{2}\rangle\end{array}\right]\right]\\ \textsc{dtrs}\ \left[\begin{array}{l}\textsc{h-dtr}\ \left[\ldots\ \left[\begin{array}{l}\textsc{head}\ \ \boxed{1}\\ \textsc{subcat}\ \langle\boxed{2}\ \boxed{3}\rangle\end{array}\right]\right]\\ \textsc{c-dtrs}\ \langle\left[\textsc{cat}\ \boxed{3}\right]\rangle\end{array}\right]\end{array}\right]\\ \textsc{c-dtrs}\ \langle\left[\textsc{cat}\ \boxed{2}\right]\rangle\end{array}\right]
\end{array}\right]
\tag{29}
$$

A token instance of (29) could be the German sentence *Peter sieht Maria* (Peter sees Mary)[19] with the transitive verb form *sieht* as its lexical head (see (12) above) and nominative and accusative NP complements as in (30) and (31) respectively.

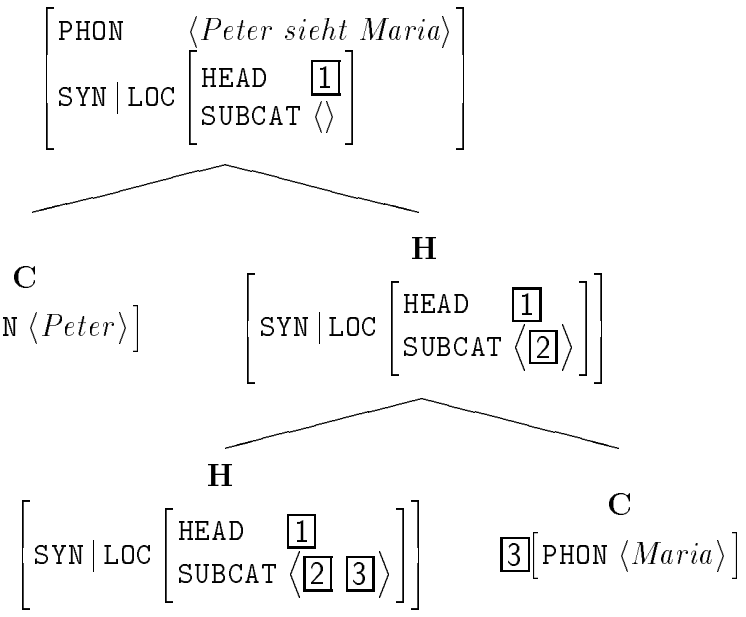
$$
\textit{word}\left[\begin{array}{l}\textsc{cat}\ \left[\begin{array}{l}\textsc{phon}\ \ \langle \textit{Peter}\rangle\\ \textsc{syn}\,|\,\textsc{loc}\ \left[\begin{array}{l}\textsc{head}\ \ \textit{noun}\left[\textsc{case}\ \left[\begin{array}{l}\textsc{obl}\ -\\ \textsc{gov}\ -\end{array}\right]\right]\\ \textsc{subcat}\ \langle\rangle\end{array}\right]\end{array}\right]\end{array}\right]
\tag{30}
$$

$$
\textit{word}\left[\begin{array}{l}\textsc{cat}\ \left[\begin{array}{l}\textsc{phon}\ \ \langle \textit{Maria}\rangle\\ \textsc{syn}\,|\,\textsc{loc}\ \left[\begin{array}{l}\textsc{head}\ \ \textit{noun}\left[\textsc{case}\ \left[\begin{array}{l}\textsc{obl}\ -\\ \textsc{gov}\ +\end{array}\right]\right]\\ \textsc{subcat}\ \langle\rangle\end{array}\right]\end{array}\right]\end{array}\right]
\tag{31}
$$

Besides the feature structure representation, phrasal signs are often given in a more traditional tree-like notation with nodes labelled according to the local CAT values and arc labels encoding the daughters structure:

[19] Actually, if desirable, (29) could be taken to account for word order variation of the sentence (like *[weil] Peter Maria sieht* ([because] Peter sees Mary)) as well. As the constituent structure (the DTRS value) itself relates to the surface string only through the PHON feature (which, in Schemata 1 and 2 we intentionally left out), in HPSG there are a number of ways open to tackle the so-called free word order of languages like German; see [Nerbonne *et al.* 1994] for a summary on the variety of suggestions.

(32)



2.4 The Saarbrücken HPSG System

Although there has been constantly growing interest in the HPSG theory in computational and formal linguistics since the publication of [Pollard and Sag 1987], the number of implemented systems and computationally exploited HPSG grammars is considerably small. What has been referred to as the Saarbrücken HPSG system in the introductory sections already, presumably is one of the most advanced natural language processing implementations including a substantial HPSG grammar of German[20].

As major parts of the work underlying this paper have been carried out and partly been implemented in the framework of the DISCO project, at least some of its distinctive features as they are related to the topic at hand will be briefly reported on here.

Top-level Feature Structure Geometry As has been outlined already, the Saarbrücken HPSG implementation has chosen to slightly diverge from the top-level (*sign*) feature structure layout given in [Pollard and Sag 1994]. Originally, [Pollard and Sag 1987] had the features PHON, SYN, SEM and DTRS all as properties of (phrasal) *sign* type objects, so that, accordingly, lexical heads were subcategorizing for the entire *sign* structure. However, to structurally prevent unbounded government (see section 2.3 above) [Pollard and Sag 1994] combine syntactic and semantic properties in a single attribute (SYNSEM) and assume the elements on SUBCAT lists to be of the type *synsem* instead of *sign*. The resulting top-level feature distribution in [Pollard and Sag 1994] hence is the following:

[20] The system has been designed and implemented in the DFKI project DISCO (Dialogue System for Autonomous Cooperating Agents) (in cooperation with the ASL — Architectures for Spoken Language and VerbMobil projects) which ran from 1990 to 1993.

(33)
$$\begin{bmatrix} \texttt{PHON} & \langle \ldots \rangle \\ \texttt{SYNSEM} & \begin{bmatrix} \texttt{LOC} & \begin{bmatrix} \texttt{CAT} & \begin{bmatrix} \texttt{HEAD} & head \boxed{} \\ \texttt{SUBCAT} & \left\langle synsem \boxed{} * \right\rangle \end{bmatrix} \\ \texttt{CONTENT} & content \boxed{} \end{bmatrix} \\ \texttt{NON-LOC} & non\text{-}local \boxed{} \end{bmatrix} \\ \texttt{DTRS} & daughters \boxed{} \end{bmatrix}$$

Obviously, looking at the Saarbücken top-level geometry found in (34), it draws a clearer distinction between phrase-structural information and non-phrase-structural (in some sense: local or categorial) information. Whatever comprises to the category of a constituent is bundled into the CAT value, whereas its constituent structure (as in [Pollard and Sag 1994]) is represented in the *daughters* structure embedded under the DTRS attribute.

(34)
$$\begin{bmatrix} \texttt{CAT} & \begin{bmatrix} \texttt{MORPH} & morphology \boxed{} \\ \texttt{SYN} & \begin{bmatrix} \texttt{LOC} & \begin{bmatrix} \texttt{HEAD} & head \boxed{} \\ \texttt{SUBCAT} & \left\langle category \boxed{} * \right\rangle \end{bmatrix} \\ \texttt{NON-LOC} & non\text{-}local \boxed{} \end{bmatrix} \\ \texttt{SEM} & content \boxed{} \end{bmatrix} \\ \texttt{DTRS} & daughters \boxed{} \end{bmatrix}$$

There seem to be both linguistic and computational reasons for a feature distribution as in (34). First, the structural encoding of the local governing domain of heads is achieved without recourse to the artificial SYNSEM feature by having elements on the SUBCAT list be of type *category*. And second, the *category* attribute MORPH as the home of morphological information in lexical signs and the interface to a two-level morphological component allows for the selection of morphological properties, e.g. a certain lexeme, in collocational and idiomatic constructions. As for the computational benefits of the Saarbrücken top-level feature geometry, it is worth noting that the clear separation of categorial from phrase-structural information allows for the design of processing devices (a chart parser, say) that need not duplicate phrase structure representations (i.e. the parse tree) over and over, as

all information collapsed into the HPSG DTRS feature is more efficiently represented in the internal data structures of the parser.

Binary Constituent Structure Again for linguistic and computational reasons, the Saarbrücken HPSG grammar is restricted to strictly binary branching constituent structures. So, diverging from standard HPSG assumptions for English phrase structure, there are exactly two possible configurations for immediate dominance schemata, viz. head-initial and head-final structures (sort of corresponding to right and left functional application in categorial grammar).

Fortunately, the basic assumption of binary phrase structure nicely coincides with what is found in (basic) German nominal structures: no matter which constituent will be considered the head in determiner plus noun combinations, there will be strict binary branching, as both categories typically contribute to the nominal group no more than once. In the case of adjunction (be it pre- or post-nominal) for semantic reasons [Pollard and Sag 1994] themselves are committed to binary structures already; and as adjunction is usually understood to be of an iterative or recursive nature (i.e. not changing the category of the modified constituent), assuming modification to take place one at a time seems to be a reasonable way of encoding this state of affairs.

However, restricting the HPSG phrase structure schemata to binary branching allows for a significant simplification of the subcategorization principle (see above), viz. a feature structure representation like that in (35) (while [Pollard and Sag 1994] give their subcategorization principle in the informal, textual representation only) that accounts for complement cancellation without recourse to functional constraints:

(35)
$$\left[\begin{array}{ll} \textsc{cat} & \left[\textsc{syn}\,|\,\textsc{loc}\left[\textsc{subcat}\ \boxed{2}\right]\right] \\ \textsc{dtrs} & \left[\begin{array}{l} \textsc{h-dtr}\left[\textsc{cat}\,|\,\textsc{syn}\,|\,\textsc{loc}\left[\textsc{subcat}\ \left\langle \boxed{1} \bullet \boxed{2} \right\rangle\right]\right] \\ \textsc{c-dtr}\left[\textsc{cat}\ \boxed{1}\right] \end{array}\right] \end{array}\right]$$

The simplified version of the subcategorization principle introduces the two minor changes to the types *category* and *head-complement-structure* found in section 2.3: (i) the feature C-DTR (formerly C-DTRS) is no longer list-valued but takes as its value a single *sign* type object and (ii) the order of arguments on the SUBCAT list has been reverted so that complements are cancelled from the front of the list (through the bullet or concatenation operator '•' on the head daughter in (35)) instead of from the tail of the list.

Named Disjunctions On the formalism side, a concept that is becoming more and more common in the logic of feature structures is that of so-called NAMED or DISTRIBUTED DISJUNCTIONS. The basic idea in assigning names to disjunctive contexts is to link disjuncts

according to the order in which they appear in disjunctions of the same name (i.e. the same disjunctive context); in fact, as we will see, adding named disjunctions to the formalism is no logical extension to the feature structure calculus but merely an elegant way of keeping disjunctive specifications local to a structure.

Looking at examples (36) and (37) we find that (36) links the values for OBL and GOV into a named disjunctive context (the subscripted tag $\boxed{1}$ in the case of named disjunctions has nothing in common with ordinary reentrancy tags), so that whenever one of the disjuncts fails (becomes $\perp$) the (by order) corresponding disjunct in all other disjunctions bearing the same name will fail too.

(36)
$$\left[\begin{array}{l}\textsc{case}\left[\begin{array}{ll}\textsc{obl} & \boxed{1}\{-\ +\}\\ \textsc{gov} & \boxed{1}\{+\ -\}\end{array}\right]\end{array}\right]$$

(37)
$$\left[\begin{array}{l}\textsc{case}\left\{\left[\begin{array}{l}\textsc{obl}\ -\\ \textsc{gov}\ +\end{array}\right]\left[\begin{array}{l}\textsc{obl}\ +\\ \textsc{gov}\ -\end{array}\right]\right\}\end{array}\right]$$

It is easily seen that the feature structures in (37) and (36) are really equivalent in that they constrain OBL and GOV to be one of + or − and at the same time mutually incompatible (i.e. CASE to be either *accusative* or *genitive*). However, (37) can be understood as a disjunctive normal form representation of (36), so that — especially if there was additional non-disjunctive information at the level of OBL and GOV or the disjunctive contexts were embedded under different feature paths — the disjunctive specification in (36) can be said to be more local than in (37).

Named Disjunctions have been proven to be especially comfortable in spelling out the cross dependencies between different clusters of properties, e.g. between the syntax, morphological properties and the surface form (PHON) in paradigmatic inflectional variation of a morphological stem.

3 Syntagmatic Relations: Government and Agreement (No Binding)

3.1 Integrational Syntax: On the Formal Status of Categories

Integrational Syntax as a contemporary stream of linguistic research is to be understood as the syntactic part of the theory of Integrational Linguistics, a school of linguistic theory that has been founded by Hans-Heinrich Lieb and fellow researchers mostly at the department for German Studies at the Free University Berlin.

There are (at least) two reasons to contrast our HPSG analysis of German nominal structures with insights from Integrational Syntax as a vantage point, viz. (i) its sufficient degree of formalization in a basic set-theoretic calculus and (ii) the profound applications of Integrational Linguistics to German and especially German syntax. Nevertheless, this thesis is settled in the HPSG framework and there will be no attempt to contrast the formalization of Integrational Linguistics with the logic of typed feature structures. Therefore, focussing on the application of Integrational Syntax to German and the consequences for our analysis of nominal structures, the following sections will be primarily based on [Eisenberg 1989], a reference grammar of contemporary German within the integrational framework (but see [Lieb 1983] and [Lieb 1993] for the general conception of Integrational Syntax and the details of its formalization).

In an introductory section on syntactic categories [Eisenberg 1989, 33 – 35] draws a basic distinction between what he calls CATEGORIZATIONS and the traditional concept of CATEGORIES. Categories, according to Eisenberg, are to be understood as sets of linguistic entities (roughly of a morphological or syntactic nature, i.e. what he calls word forms, in the case of syntactic categories) whereas categorizations (acting, in a sense, as meta-categories) form sets of categories themselves. As an example of the categorization vs. category distinction Eisenberg quotes case as the class (i.e. categorization) of the categories *nominative*, *genitive*, *dative* and *accusative* (for German).

Most, if not all of the categories that will be important in our analysis of the syntactic structure of German nominals are in some sense related to inflectional marking and the concept of (syntactic) PARADIGMS. Paradigms, in the terminology of Eisenberg, are to be taken as sets of word forms that have an internal organization by syntactic categories. Likewise, word forms are (informally) defined as the linguistic entities that actually make up linguistic tokens or utterances, i.e. the inflected variations of (lexical) words[1].

On the basis of the concept of syntactic paradigms and their positions (slots or — in

[1] For a somewhat stronger formalization of the concepts of (syntactic) words, word forms and paradigms see [Lieb 1983, 104 – 110] and [Lieb 1980].

As the intuitive explications taken from Eisenberg will be exemplified in their application to German nominal structures, we feel justified in not reconstructing the integrational terminology from its basic elements here.

the integrational jargon — units) [Eisenberg 1989, 36 – 44] then suggests to distinguish categories (or their categorizations) that apply to the single units of a paradigm from those that hold for entire paradigms. As a straightforward example consider (1), the paradigm of the German noun *Mann* (man) in the traditional table notation:

(1)

Case \ Number	*singular*	*plural*
nominative	Mann	Männer
genitive	Mannes	Männer
dative	Manne	Männern
accusative	Mann	Männer

Obviously, number and case are categorizations accounting for the internal organization of the paradigm, whereas gender in the case of German nouns applies to all units from the paradigm, because there is no inflectional variation according to grammatical gender[2]. Hence, *singular* and *plural* and *nominative* to *accusative* for German nouns have a different categorial status than the German gender categories *masculine*, *feminine* and *neuter*. Therefore, in Integrational Syntax the former are classified as UNIT CATEGORIES and the latter as PARADIGM CATEGORIES.

Now, if we apply these definitions to the paradigm of German adjectives (or inflected determiners) we find that gender seems to have a different categorial status for adjectives than it has for nouns because adjectives clearly are inflected for gender. Eisenberg concludes that gender in the case of adjectival paradigms therefore is to be understood as a unit categorization, whereas for german nouns, as we saw, it is a paradigm categorization. In fact, he argues, it is a common misunderstanding to assume the two categorizations to be the same; in the integrational analysis gender on nouns and adjectival gender are taken to be two DIFFERENT categorizations that — for whatever reasons — happen to share the same name.

Next, [Eisenberg 1989, 52 – 57] introduces SYNTAGMATIC RELATIONS[3] as the link between what he calls syntactic means (linear order, inflectional marking and intonation)

[2] We will have to reconsider this point when looking at the restricted set of German nouns that actually exhibit morphological variation along the gender dimension, i.e. the well-known *Beamter* (civil servant) and *Verwandter* (relative) class. It will be argued then that these are to be taken as substantively used forms from adjectival paradigms rather than actual nominalizations.

However, as the syntactic distribution of these forms is obviously that of common nouns, their major classification in the HPSG lexicon will be very similar to that of nouns except for the additional inflectional variation according to grammatical gender.

[3] As a matter of terminology, Eisenberg reserves the term SYNTACTIC RELATION for what often is called a grammatical function, i.e. notions like subject, predicate, object et al.

Syntactic relations in this sense will be closely related to semantic interpretation and presuppose a syntactical structure analysis (in that, say, the opposition between subject and object can be defined on the basis of a syntactic structure and syntagmatic relations holding therein) whereas — as we will see — syntagmatic relations are taken to be the key to syntactic structures in general.

and the syntactic (i.e. constituent) structure. From the four syntagmatic relations that Eisenberg assumes (viz. government, agreement, identity and positional binding) we will only be concerned with government and agreement in the context of this paper.

The characterization of syntagmatic relations as the links between syntactic means and syntactic structure results from the fact that, say, the order of words and their inflectional marking are visible surface properties of linguistic utterances whereas their constituent structure does NOT in any sense reflect itself in the (surface) form of spoken or written linguistic tokens. However, there is broad consensus on the existence of a (presumably tree-like) hierarchical structuring of linguistic utterances and their parts, although the concrete nature of these structures and their relation to the surface form has always been the topic of heavy debating.

Eisenberg suggests his notion of syntagmatic "relations between parts of linguistic utterances among each other on all levels of the system" [Eisenberg 1989, 52] as direct evidence for specific constituent structures, because he assumes constituents to be both in the domain and range of syntagmatic relations. As among others government and agreement in the integrational understanding will be defined in terms of paradigm and unit categories (and maybe other surface properties), i.e. elements from the domain of what has been called the syntactic means, they can indeed be seen as linking the surface form to its syntactic structure, in potentially motivating the assumption of some specific constituent structure and rejecting another. We will come back to this key role of syntagmatic relations when looking at the encoding of government and agreement relations in HPSG and attempting to spell out which of them hold in what parts of German nominal structures.

Now for the definitions; [Eisenberg 1989] gives the following:

GOVERNMENT

(2) A constituent f_1 governs a constituent f_2 when the form of f_2 is determined by a paradigm category of f_1.

AGREEMENT

(3) A constituent f_1 agrees with a constituent f_2 when f_1 depends in at least one of its unit categories on a unit category of f_2.

Commenting on (2) and (3), first, it should be noted that the definition of government is more general than that of agreement in that it rather abstractly requires the form of f_2 to be determined by a paradigm category of f_1, whereas the definiton of agreement makes much more specific recourse to unit categories (hence again, inflectional properties) on both f_1 and f_2. From the examples given by Eisenberg it is clear that the determination of the syntactic form can range over unit categories (e.g. case values on verbal complements) and paradigm categories (the major category or part of speech of verbal complements) as well as other lexical properties like in the selection of a specific preposition and a certain case governed by that preposition as a prepositional object to a given verb.

Second, the formulation of the agreement relation suggests that agreement be a directional, i.e. not (necessarily) symmetric, relation in that it requires unit categories of f_1 to DEPEND on unit categories of f_2. Additionally, the concept of directionality of agreement is verbatimely reflected in the annotations to constituent diagrams given by Eisenberg that represent agreement relations as arrows that are (in opposition to the arrows marking the additional syntagmatic identity relation) headed at only one side.

In applying the above definitions to the combination of determiner (optional attributive adjective) and noun in German, Eisenberg observes that — in his terminology — there is both government and agreement within these syntagmas: as we saw, gender is a paradigm categorization for nouns but a unit categorization in the determiner (and adjectival) paradigms, so that the distribution of gender values within nominal groups is taken to be governed from the noun and not (as Eisenberg claims is to be found in most traditional grammars of German) an instance of agreement.

On the other hand, the two remaining categorizations that we have been looking at so far, viz. number and case, are unit categories for nouns as well as for determiners and adjectives; hence, their covariation within the German nominal group according to Eisenberg actually is the result of agreement between the respective constituents.

In particular, in the arrow diagrams given, Eisenberg makes the noun the 'source' and the determiner or adjectives the 'targets' of the agreement relations accounting for number and case distribution (i.e. he has arrows pointing from the unit categories on the noun to the respective categories on the adjective and determiner). Although, neither in the introductory sections on syntagmatic relations, nor in the more detailed analysis of German nominal groups Eisenberg gives explicit arguments for this choice of direction in number and case agreement, the intuition seems to be that the noun being the governor with respect to gender has a special status within the nominal group and therefore in some sense controls agreement relations holding in the same context as well[4].

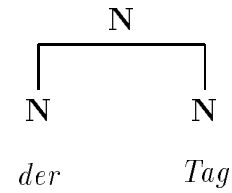
Regarding the syntactic (constituent) structure of linguistic tokens, there are only very moderate restrictions on the structure format to be found in [Eisenberg 1989]. In a very general sense there seems to be the tendency to favour 'flat' phrase structure analyses (e.g. (5) over (6)) and move the majority of information into what is called the MARKING STRUCTURE: the annotation of constituent structure diagrams with unit and paradigm categories and the syntagmatic relations holding between them. However, based on the definitions of nominal constituent categories (see [Eisenberg 1989, 41 and 42]) the mother

[4] Although throughout [Eisenberg 1989] the concept of (syntactic) heads does not seem to play any role in the theory of Integrational Syntax, there is actually the notion of NUCLEUS (*Kern*) found in the context of groups of constituents as well as in the application to morphological composition.

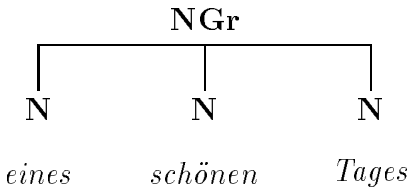
Still, the nucleus idea — as it is a rarely and only informally used concept — can not be taken to account for the directionality in agreement relations observed by Eisenberg. From his distinction between the constituent categories N (determiners, adjectives, nouns AND simple determiner plus noun combinations) and NGr (more complex nominal groups, e.g. incorporating additional nominal attributes) it follows that Eisenberg assumes both attributive adjectives and nouns (but not determiners) to be nuclei in the nominal group.

node in (6) would have to be labelled as a simple N because it does not satisfy the requirement on nominal groups (NGr) to contain (at least) two nominal nuclei (where determiners do obviously NOT qualify as nuclei of nominal groups as we see from (4)) — a restriction on potential nominal phrase structure that Eisenberg fails to mention in his claim that both (5) and (6) (with a label NGr on the mother) could be assigned to a sequence of three nominals (see [Eisenberg 1989, 49 – 51]).

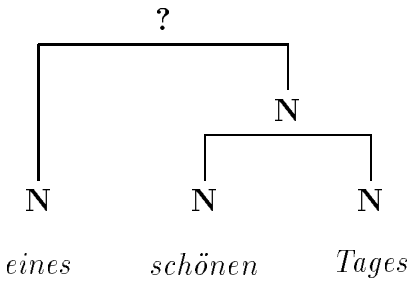
(4)



(5)



(6)



Summing up our review of syntagmatic relations and the role of constituent structure in the integrational framework, there are two generalizations to be mentioned that we may want to reconsider in the context of a forthcoming HPSG analysis of German nominals.

First, in [Eisenberg 1989, 55] government is assumed to be a group building mechanism in that "the government relation formally adjoins constituents so that they jointly form a higher constituent." In other words, instances of government license (or maybe enforce) phrasal nodes in the constituent structure like the mother nodes N and NGr in (4) and (5).

Actually, a straightforward consequence of this explication of government and its impact on phrase structure would be a notion of the GOVERNING DOMAIN as the range of constituents that under government are grouped into a superior phrasal node. Hence, as the noun in (5) is taken to govern both the determiner and the adjective according to the gender categorization (see above), there is now a second reason to rule out structures like (6), viz. that the lexical noun *Tages* (days) would have to extend the governing domain

of the embedding nominal group to require *masculine* gender as a unit category on the determiner *eines* (one)[5] Nevertheless, we will have to come back to this point in the light of the distribution of declension classes governed by the determiner — a state of affairs that [Eisenberg 1989, 238] has to move beyond the scope of his formal apparatus.

Second, Eisenberg clearly rules out a UNARY (non-branching) constituent structure by virtue of — as he argues — the underlying concept of syntactic categories rather than by an arbitrary stipulation in the descriptive formalism. As constituent categories shall be taken to characterize a token constituent with respect to its syntactic properties, according to [Eisenberg 1989, 51] there is no sense in which the same constituent can simultaneously be both a simple noun (N) and a nominal group (NGr). Hence, what this claim results in, is the assumption that constituent categories are DISJOINT sets of constituents, such that there will be no linguistic token allowed to be a member of more than one of them.

Finally, Eisenberg suggests to distinguish two types of government, viz. what he calls lexical government from so-called categorial government. Lexical government, he argues, is a property of individual lexicon entries (like the specific case governed by a preposition) whereas categorial government is a property of a certain category as a whole, e.g. the ability of all nouns (a paradigm category) to take nominal attributes assigning them *genitive* case. Although this distinction will not play a role in our HPSG analysis of nominal structures (actually, Eisenberg himself makes no further reference to it when it comes to syntagmatic relations in the German nominal group), we will briefly want to compare it to the different types of government assumed in [Zwicky 1986] (see section 3.2.4).

3.2 [Zwicky 1985]: Properties of Syntactic Heads

> The intuition to be captured with the notion HEAD is that in certain syntactic constructs one constituent in some sense 'characterizes' or 'dominates' the whole. From these basic ideas [*sic*], however, it is possible to move in many directions. [Zwicky 1985, 2]

Among the very frequent matters of dispute in the analysis of nominal structures is the opposition between the traditionally so-called NP analysis and what recently has become known under the label DP analysis. Very generally, what the issue amounts to is — as the names shall suggest — an ongoing argument on whether it is preferable to think of a group of nominals (e.g. a simple determiner noun combination) as a projection of the noun

[5] Obviously, this argumentation only holds as long as the government of gender on the determiner is not taken to be a property of the entire nominal group (or $\overline{\text{N}}$ in more traditional terminology), although this move would imply that lexical government properties can be inherited by superordinate constituents.

However, in the HPSG framework this exactly falls out as a very natural assumption in that we will see that government relations originate from subcategorization requirements in the lexicon and are inherited to intermediate projections by virtue of the `SUBCAT` principle.

(hence, a noun phrase) or rather as being projected from the determiner (a determiner phrase).

Naturally, there is a lot of variation in terminology and underlying stipulations involved in the NP vs. DP opposition, especially depending on the specific formal framework or linguistic theory at hand. Nevertheless, there seems to be one common concept that besides the considerable degree of variation in its definition and use is taken to account for the fundamental difference between the noun or determiner phrase analyses: the notion of SYNTACTIC HEADSHIP.

Using the term 'head', the rough characterization of nominal groups as being a projection of either a noun or a determiner from the previous paragraph can be taken to mean that either the noun or the determiner serves as the syntactic head of the whole construction. Thus, as — quite uncontroversially for frameworks that assume the existence of a VP phrasal node — the verb is assumed to be the head in the verb phrase, the category NP implies that there is a noun heading the phrase, while speaking of a determiner phrase suggests to think of the determiner as the actual head of the very same construction.

As we saw in section 2.3, the idea of headship plays a very central role in the HPSG phrase structure account, viz. in the type *headed-structure* (and its subtypes) and the head feature principle. Differing from the analysis we sketched from Integrational Syntax (which has mostly discarded the concept of syntactic headship) in the last section, we expect any HPSG account of German nominal structures to be confronted with the question of which constituent it is going to make the head daughter in the nominal group. Therefore in this section we shall try to single out some of the properties that are intuitively attributed to syntactic heads and discuss their relation to other (syntactic) devices that in one way or the other can be found in HPSG (although not all of them manifest as primitive HPSG concepts with an overt correlate in the HPSG ontology, the type lattice or the set of grammar principles).

Whereas in the history of syntactic theory and especially the school of generative (or transformational) grammar based on the work of Noam Chomsky it has been traditional to favour variants of the NP analysis, in recent development in what is now commonly referred to as Government Binding Theory (GB) the treatment of nominal groups as being headed by determiners has gained a very strong position. Since the suggestion to treat determiners as functional categories in analogy to complementizers and the GB inflection element in [Abney 1987] (originally applied to the English noun phrase), the resulting DP analysis has indeed become widely accepted in the application of Government Binding ideas to German nominal structures.[6] However, for the moment we shall resist the temptation to look into the details of the GB analyses suggested for German and the serious differences in how they account for government and agreement phenomena or determinerless constructions, but instead will try to sharpen our notion of syntactic heads without committing ourselves

[6] E.g. [Haider 1988] and [Olsen 1990] give evidence for this move in the German GB literature towards the treatment of determiners as functional heads selecting an NP complement, rather than as specifiers to $\overline{\text{N}}$.

to the specific phrase structural constraints commonly held in $\overline{\text{X}}$-theory.

In searching for criteria on syntactic headship we will mostly review work by Arnold M. Zwicky published in [Zwicky 1985] and [Zwicky 1986] that is not really tied to a specific linguistic school or syntactic formalism (although clearly closer related to the GPSG (Generalized Phrase Structure Grammar) framework than to the tradition of transformational grammars). In the sections to come it will then be possible to evaluate the existing HPSG analyses for German nominal groups by use of the nomenclature borrowed from Zwicky.

In [Zwicky 1985] it is claimed that despite the growing use of the term 'head' in both morphology and syntactic theory there is a wide variety of often very intuitive and informal notions associated with syntactic headship. As a vantage point, Zwicky therefore suggests to establish a set of primitive syntactic concepts that (to his opinion) in some form must be incorporated into any particular theory of grammar. With this core inventory, Zwicky argues, it will then be unlikely to find another independent primitive concept for headship, but instead either one of the fundamental devices in syntax will turn out to subsume all the phenomena that have been traditionally taken to be instances of (syntactic) headship or — more likely — we will see that for different purposes there will be different criteria on heads, such that one would either have to discard the term at all or identify it for a given purpose with one of the primitive concepts from the core inventory.

The actual grammatical concepts that Zwicky considers with respect to their potential head properties are the following:[7]

- The semantic argument

  "[..] the constituent acting as the semantic argument , as opposed to the semantic functor in a syntactic combination."

- The Subcategorizand

  "[...] the constituent that is lexically subcategorized with respect to the sister constituents it can occur with."

- The morphosyntactic locus

  "[...] the constituent on which inflectional marks will be located."

- The governor

  "[...] the constituent determining government."

- The determinant of concord

  "[...] the constituent determining government."

[7] All citations in this list are taken from [Zwicky 1985, 3].

As [Zwicky 1985] really avoids the recourse to a specific syntactical framework, he can hardly provide a formal definition for any of these concepts. Instead, he uses a set of common constituent configurations together with example phrases from English (see table (7)) to illustrate each of the five notions, trying to explicate what the idea is behind each of them. Still, however, there remains a substantial degree of vagueness in most cases, especially as the number of examples given in [Zwicky 1985] is considerably small and (mostly) limited to English.[8]

(7)

Concept \ Construction	Det + N	V + NP	NP + VP
Semantic Argument	N	NP	NP
Subcategorizand	Det	V	–
Morphosyntactic Locus	N	V	VP
Governor	–	V	VP
Determinant of Concord	N	NP	NP

3.2.1 The Semantic Argument

Although the focus of [Zwicky 1985] clearly is on SYNTACTIC headship, in an introductory section it is argued that three of the five head-like notions in (7) are to be understood as interfacing the syntax to the semantics, the lexicon and the morphology respectively, so that there should be no principal obstacle against semantic considerations in the determination of heads in the syntax. Despite the counter arguments that will be given against this position (regarding the status of subcategorization and inflectional marking in the syntax), we will follow Zwicky for a moment and briefly review his concept of semantic arguments[9].

According to Zwicky, for many constituent combinations X + Y there is a strong intuition that "X is the 'semantic head' if, speaking very crudely, X + Y describes a kind of the thing described by X" [Zwicky 1985, 4]. So, still very crudely, Zwicky suggests that in Det + N the semantic head is N (because *that penguin* describes a kind of penguin) and that on semantic grounds it is the functor argument distinction that can be taken to account for the 'kind of' relation quoted before. As in the interpretation of Det + N the determiner acts as a semantic functor on the interpretation of N, Zwicky cautiously concludes: "We

[8] Besides the constructions found in table (7), Zwicky from time to time gives examples of Aux + VP, Comp + S and P + NP combinations as well. Because Zwicky himself admits that the division of categories and potential syntactic analyses for these examples might be much less uncontroversial than for the constructions from (7), we will restrict ourself to Det + N, V + NP and NP + VP combinations presently. We feel justified in doing so even more as not all of the headship candidates in table (7) can be applied to Aux + VP and Comp + S, so that these anyhow play a somewhat inferior role in the overall argumentation.

[9] As we will see in section 4.1 the HPSG analysis of nominal structures suggested in [Pollard and Sag 1994] too is strongly influenced by semantic considerations (like roughly the functor argument distinction made by Zwicky) and an inherently semantic theory of agreement.

might then propose that in X + Y, X is the 'semantic head' if in the semantic interpretation of X + Y, Y represents a functor on an ARGUMENT represented by X" [Zwicky 1985, 4].

However, Zwicky himself seems to be rather unhappy with this move of identifying the intuitive 'kind of' relation with the concept of headship in semantics and the semantic functor argument distinction. Although he schematically deduces that in the interpretation of both V + NP and NP + VP constructions the NP is usually taken to be the argument to functors represented by the verb and verb phrase, respectively, so that it has to be identified as the semantic head of these constructions, he does NOT apply this idea to the corresponding set of examples. In analogy to the Det + N case with his explication of semantic heads Zwicky would be forced to assume that *control that penguin* describes a kind of penguin whereas *Peter controls that penguin* describes a kind of Peter. Obviously, for the V + NP and NP + VP examples the identification of the semantic head as the semantic argument runs against the common intuition that *control that penguin* and *Peter controls that penguin* roughly describe some kind of controlling.

3.2.2 The Subcategorizand

Subcategorization, according to Zwicky, is the property of lexical items to specify a set of sister constituents that they can combine with, e.g.

> The verb *give* is subcategorized to occur with either NP NP or NP *to* + NP as its sisters (*give Kim money*, *give money to Kim*); *donate* is subcategorized to occur only in the second of these two constructions (*donate money to Kim*). [Zwicky 1985, 5]

Together with the assumption that "subcategorization is associated with rules of constituent combinations, namely those introducing lexical categories" [Zwicky 1985, 5], at first glance the subcategorization concept of Zwicky looks like a somewhat informal variant of the HPSG `SUBCAT` mechanism as it was introduced in section 2.3. However, in the examples cited above we notice that Zwicky restricts subcategorization frames to the bare constituent categories (and the surface form of the preposition in the *to* + NP case) and excludes all information on, say, the different case values that the verb *give* might assign to its direct and indirect objects.

Both the specification of the set of (atomic) categories that some lexical item will combine with and the determination of additional (syntactic) properties on those co-constituents are instances of SELECTIONAL RESTRICTIONS imposed by one constituent onto its sister constituents. However, Zwicky wants to maintain a difference between the determination of the general category of a co-constituent on the one hand and the selection of the concrete morphosyntactic shape this constituent is going to exhibit on the other hand. The former, he claims, constitutes subcategorization and is restricted to the lexicon, whereas the latter is called government and can hold between lexical as well as phrasal constituents.

Using the integrational terminology introduced in the last section, one could possibly rephrase the suggested distinction between subcategorization and government as the se-

lection of properties from either the constituent structure (i.e. constituent categories like NP) or from the marking structure (the domain of paradigm and unit categories). Still, in Integrational Syntax both processes would be subsumed under the concept of government as this has generally been defined as the relation holding between any two constituents with a paradigm category on one constituent determining the form of the other.

Intuitively speaking, what Zwicky is trying to achieve in the subcategorization vs. government distinction is the separation of the bare licensing of co-constituents from the actual selection of the form of those items. However, this presupposes that there really is a fundamental oppositon between a category label like NP (an entity that lexical items can be subcategorized for) and, say, the different case values that can be assigned to an NP. Given the extended HPSG notion of categories (see section 2.3) as a complex type, encoding whatever morphological, syntactic and semantic properties a constituent may have (including the part of speech and in a sense its bar or projection level, i.e. the information represented in the atomic constituent category labels of Zwicky) the distinction between subcategorization and government will hardly persist; indeed, in section 3.3 we will see that in the HPSG theory of grammar government phenomena and the `SUBCAT` mechanism are inseparably bound to each other.

Yet, the key argument that Zwicky gives to motivate the separation of licensing and selection (viz. into the concepts of subcategorization and government) is that in the association of subcategorization with the rules of constituent combination and the assumption that constituent combination has to be paired with the principles of semantic interpretation, "there will necessarily be a close relationship between subcategorization and semantic interpretation" whereas "[...] government lacks this semantic correlate" [Zwicky 1985, 5 and 7]. But again, the argument only holds as long as one accepts subcategorization as a syntactical device distinct from government, so that it can be linked to constituent combining rules in its own right. Furthermore, to its full extent, the conclusion amounts to the claim that in a ditransitive verb combining with two noun phrase objects (e.g. *give Kim money*) the actual case values (and maybe additional syntactic properties) assigned to the objects do actually not take part in the semantic interpretation of the phrase. For languages with a richer inflectional inventory than English this claim is — at best — less than uncontroversial.

There is yet another questionable point in the concept of subcategorization put forth in [Zwicky 1985, 5 – 7]. As has been mentioned already, Zwicky thinks of subcategorization as the interface between the syntax and the lexicon (however the two will actually be separated), so that subcategorization frames are taken to be properties of lexical items only. Accordingly, looking at table 7, we find that in V + NP the lexical verb is assumed to be the subcategorizand (the constituent that is subcategorized with respect to its sisters) whereas in NP + VP neither the noun phrase nor the verb phrase can be assigned the function of the subcategorizand, simply because none of them is lexical.

There are a number of rather strong presuppositions about the nature of grammar theory implicitly hidden in restricting subcategorization to the lexicon. First, there is no way

in which subjects can be taken to be subcategorized for by verbs similar to the specification of objects in the subcategorization frame (which is at odds with the widely held HPSG approach to let the subject indeed be a member of the verbal SUBCAT list). Second, subcategorized for arguments will be required to be bound all at once in a flat structure — if not so, there would be no straightforward way to account for the relation between the subcategorization information on a lexical constituent and the valence (i.e. selectional restrictions) of a phrasal projection of this constituent. Presumably, all subcategorization requirements would have to be fulfilled in trees of depth one (possibly including modifiers that are not lexically licensed), because otherwise the theory of grammar would have to allow subcategorization frames to be passed up to phrasal nodes (which is roughly what HPSG uses the subcategorization principle for).

In looking for the subcategorizand in Det + N, Zwicky claims that it is a fairly non-controversial position "that a construction like *those penguins* results not from a single rule combining Det with N, but rather from (at least) two rules, one combining Det with a phrasal category Nom, and the other permitting an N unmodified by adjectives to be one of the realizations for Nom" [Zwicky 1985, 5]. Therefore, if in a noun phrase only the determiner really is lexical, it follows that Det must be the subcategorizand, which, it is argued, "is a welcome consequence, given the familiar fact that determiners are lexically subcategorized according to whether they can combine with singular count nouns (*each penguin* [...]), plural count nouns (*many penguins* [...]), or mass nouns (*much sand*)" [Zwicky 1985, 5 – 6][10].

We will consider the question of whether it is appropriate to think of the noun licensing its determiner or vice versa (which, once more, can be taken as an instance of the noun phrase vs. determiner phrase argument) in detail in the comparison of two HPSG analyses of (German) nominal structures that diverge in exactly this matter (see section 4). Nevertheless, it should have become clear that for Zwicky the only reason to make the determiner the subcategorizand in determiner noun combinations is the bare fact that he has restricted his notion of subcategorization to the lexicon. If one was willing to associate subcategorization frames with phrasal categories, the claim that determiners select the class (including number properties) of nouns they combine with could easily be reverted.

3.2.3 The Morphosyntactic Locus

From the five head-like notions listed in table 7, the morphosyntactic locus is the one that Zwicky suggests to be actually identified with the term head in syntax, at least for the purpose of syntactic percolation, as we will see. The morphosyntactic locus of a phrase,

[10] We note in passing that *singular* and *plural* are morphosyntactic properties that go well beyond the level of atomic category labels as they were used in the *give* and *donate* examples above.

Going back to our earlier suggestion to interprete the suggested opposition between subcategorization and government as the distinction between constituent and marking structure in the integrational framework, we find that in this case number (a unit categorization) clearly is a concept that has its home in the marking structure.

according to Zwicky, is to be understood as the constituent bearing "the morphosyntactic marks of syntactic relations between the construct and other syntactic units" [Zwicky 1985, 6], where morphosyntactic marks can either be taken to be (i) the actual inflectional realization of morphosyntactic properties like, say, number, gender, case and person in languages that have a sufficiently rich inflectional morphology or (ii) as rather abstract properties of constituents, percolating up from a lexical node to its projections. Zwicky suggests the terms INFLECTIONAL LOCUS for the former, MORPHOSYNTACTIC LOCUS for the latter explication of morphosyntactic marks and claims that the actual inflectional locus (if realized in a language) of a construction usually will serve as an operational criterion in locating the abstract morphosyntactic locus.

The motivation for Zwicky to propose that the morphosyntactic locus be a good equivalent for the term head in syntax, is the common observation that in a token group of constituents (e.g. a nominal group) inflectional marks as abstract morphosyntactic properties may be shared between the mother node and one of its daughters. Although he gives no formal detail on how one would account for properties being shared between constituents, by using the term 'percolation' Zwicky implicitly suggests that there is one cluster of properties being realized on one of the (lexcial) daughter constituents flowing up to the whole phrase, thus marking the mother node for exactly the same set of properties.

It has been noted already that in a unification-based framework the idea of shared properties (i.e. the percolation of features) can be straightforwardly encoded in terms of reentrant (token identical) feature structures. So obviously what Zwicky calls the morphosyntactic locus in a phrase nicely maps onto the HPSG (or GPSG) concept of `HEAD` features being percolated by some sort of head feature principle (see section 2.3 for the details on the HPSG variant). Accordingly, the conclusion that the morphosyntactic locus (i.e. the bearer of `HEAD` features) is to be identified as the "explication of headship in syntactic theory" [Zwicky 1985, 6] directly squares with the HPSG assumption that the `HEAD` features of a (headed) phrase always be comprised by the head daughter (the value of the `H-DTR` attribute in the *headed-structure* component of HPSG phrasal signs).

How then does Zwicky locate the morphosyntactic loci in the three example constructions we have been using in illustrating the other head-like concepts already? Probably the most uncontroversial case is V + NP, like in *control those penguins* as opposed to *controls those penguins* — clearly, the difference along the person dimension (anything but third person for the first, third person for the second example) on the whole phrase can not be linked to a person distinction on the object NP so that it must be the finite verb comprising the person (as well as number, tense and mood) properties in the VP.

Somewhat similarly for the NP + VP example, the tense and mood properties of the whole sentence can not be explained on the basis of inflectional marks on the subject NP but only as being percolated up from the verb again. Hence, the morphosyntactic locus in NP + VP is the verb phrase by virtue of being headed by a finite verb and sharing its morphosyntactic features with it. Although (for the sake of the argument assuming the phrase structure resulting from the V + NP and NP + VP constructions given by Zwicky)

neither the sentence nor the VP phrasal nodes can be said to bear inflectional marks in the sense of a morphologically realized inflectional locus, it is commonly assumed that they are marked for properties like number, person and tense[11]. So, taking the finite verb to be the morphosyntactic locus in the verb phrase which itself comprises the morphosyntactic locus for the whole sentence, seems to be a reasonable motivation to assume that inflectional marks realized on the head verb (as suffixes, say) percolate up as abstract properties to the VP and S nodes, thus straightforwardly accounting for the distributional opposition in sentences like *the children control those penguins* and *Peter controls those penguins.*

For the Det + NP sample construction, Zwicky very briefly claims that N must be the inflectional locus in the nominal group because "the distinction between singular *the child* and plural *the children* is linked to number distinctions in the VP" [Zwicky 1985, 6]. Number, the argument seems to go, is a relevant property of the whole NP as it can be involved in syntagmatic relations holding between the noun phrase and other constituents (e.g. in what is commonly referred to as subject verb agreement in NP + VP), and — as morphological marks for the number distinction are realized on the respective nouns in *the child* vs. *the children* but not on the determiner — it must be the number information from the lexical noun percolating up to the NP node.

From the data given, it is not really clear whether Zwicky assumes English determiners to be marked for number at all. However, following the arguments for treating the determiner as the subcategorizand in the nominal group quoted earlier, it would be possible to consider all instances of English determiners that actually exhibit number variation (e.g. *this penguin* vs. *these penguins*) as being lexically subcategorized for a specific class of nouns, instead of bearing morphological marks for singular or plural themselves. Although this move would be especially consistent with what Zwicky set out for the opposition between *each penguin* and *many penguins*, viz. that the determiners subcategorize for either singular or plural count nouns respectively, the resulting analysis would be unable to treat, say, *this* and *these* as being systematically linked to each other (e.g. — most naturally — as paradigmatically varying with respect to number).

No matter whether Zwicky would in fact take number to be an inflectional property of English determiners or not, looking at examples from German we find inflectional variation along the number dimension (and several others) in the determiner paradigm. The resulting problem, of course, is that, as soon as both constituents in the Det + NP combination are overtly marked for number and (as we will assume it for the moment) all other relevant HEAD properties of the embedding noun phrase, either the inflectional locus will no longer serve as an operational criterion in locating the actual morphosyntactic locus, or one would have to assign both the determiner and the noun in Det + NP a head-like status for the

[11] For the moment we do not really care about the actual inventory of HEAD features stipulated for either S, VP or NP (in English) by Zwicky. Only in the context of how HPSG analyses for nominal structures assign morphosyntactic features to the elements of the noun phrase, we will actually compare the suggested accounts for German nominals according to which properties they stipulate for which levels of the structure.

purpose of syntactic percolation. As a matter of fact, in German roughly the same set of morphosyntactic categories relevant to the nominal group as a whole (i.e. — at least — number, gender and case) is equally well realized not only on determiners and nouns but as well in the adjectival paradigm. Therefore, the latter variant (viz. stick to identifying the inflectional and morphosyntactic loci in the nominal group) would obviously lead to abandoning the concept of syntactic headship as it was suggested by Zwicky. — If in the combination of two or three lexical items to a nominal group in fact all the constituents (including, as would trivially follow, intermediate levels of projection) were considered the heads of the construction, the intuion behind this concept, viz. the superior status of a particular element, would be completely lost.

So, given examples like *die Frau* (the woman) vs. *die Frauen* (the women) but *der Lehrer* (the teacher) vs. *die Lehrer* (the teachers), the distribution of concrete morphological marks can hardly be taken as evidence for assuming either the determiner or the noun to be the morphosyntactic locus in German nominal structures (maintaining, of course, that the morphosyntactic locus as a head-like concept in syntax be assigned systematically in the construction Det + NP and not arbitrarily to the noun in the first, but the determiner in the second pair of examples). Still, though giving up the criterion of identifying the morphosyntactic with the inflectional loci where the latter are inflectionally realized, we might want to preserve the idea that syntactic headship indeed is linked to the percolation of morphosyntactic properties.

As with the subcategorizand, in applying the concept to the German nominal group it will be necessary to individually consider the properties of lexical items in their relations to their sister as well as to embedding constituents. As we find both the morphosyntactic locus concept and its function in syntactic percolation encoded in the core HPSG inventory (viz. in the head feature principle), we will again postpone the detailed examination of the consequences of making either the determiner or the noun the head to section 4, where — in reviewing two analyses for German nominals suggested within the HPSG framework — it will provide us with a useful measure for the linguistic adequacy of the different proposals.

3.2.4 The Governor

Both government and concord, according to Zwicky, differ from the three head-like notions we have seen so far, in that they do not interface the syntax to some other part of the grammar (like the semantics, lexicon or morpholgy). Instead, Zwicky claims, government and concord are syntactic phenomena "in which one constituent in a construct can intuitively be said to 'dominate' another" [Zwicky 1985, 7] so that any reasonably formal theory of grammar will have to account for them as theoretically primitive syntactic notions in their own right.

However, major parts of the discussion on government in [Zwicky 1985] are devoted to arguments for its separation from subcategorization, most of which have been quoted in the section on subcategorization already (see above). So, nearly all we find in terms of a

defining explication of the concept is the following:

> Syntactic government, speaking rather loosely, is the selection of the morphosyntactic shape of one constituent (the GOVERNED, or SUBORDINATE, constituent) by virtue of its combining with another (the GOVERNOR).
>
> [...]
>
> There are, in fact, two rather different sorts of government. In government of the first type, the governed constituent bears features simply by virtue of its occuring in a construction. In government of the second type, a lexical cleavage within the governor is projected as inflectional marks on the governed constituent. [Zwicky 1985, 7]

Since in [Zwicky 1986] (besides the reference to the 1985 paper) we find a far more elaborate conception of government relations and the two texts happen to be sufficiently close to each other in both the topic and degree of formalization, we will feel free to integrate it into the review of head-like devices from [Zwicky 1985] presently.

Configurationally, Zwicky argues, there are two basic types of government: VERTICAL — with the governor being the mother of the governed constituent — and HORIZONTAL government where governor and governed are sisters to each other[12]. Likewise, it is then suggested, instances of horizontal government do fall into two classes, depending on whether the governor is PHRASAL or LEXICAL respectively. Lexical government in turn can additionally be subclassified into what Zwicky calls SYSTEMATIC vs. IDIOSYNCRATIC government.

If [Zwicky 1986, 969] in doubting the existence of actual natural language phenomena exhibiting examples of phrasal horizontal government was right (which in the intuitive sense coincides with the HPSG perspective that selectional restrictions originate from the lexicon, but — at the same time — conflicts with a formal explication of government relations within the HPSG framework put forth in section 3.3), then the configurational classification into vertical vs. horizontal and phrasal vs. lexical government would collapse to the more traditional distinction between structural and lexical government. In fact, the use of the terms "occuring in a construction" and "lexical cleavage" in the explication of the two government classes given in the above quote from [Zwicky 1985], suggests that the former indeed refers to structural properties of a group of constituents and the latter to lexical properties of the governor.

Among others, we find the three sample constructions we have been looking at so far in the examples given in [Zwicky 1986]. For English, it is argued, V + NP and NP + VP are

[12] Although in [Zwicky 1986, 968] it originally reads "[...] VERTICAL and HORIZONTAL, according as the governing element is the mother or the daughter [*sic*], respectively, of the governed element" — we take it from the immediately preceding paragraph and the examples given, that what is intended, is in fact the sister relation (e.g. in verbs and prepositions governing a specific case on their objects or auxiliaries selecting the form of the verb they combine with).

instances of vertical (or structural) government[13], with the NP being assigned accusative and nominative case simply by virtue of being a daughter to VP and S respectively, whereas in languages like German the case assignment in V + NP varies as a lexical property of the verb, so that it is to be analyzed as horizontal government. Moreover, because (at least for German) the individual case values governed by the verb can not be linked systematically to other syntactic or semantic properties of the governor, Zwicky suggests to consider V + NP in German as an example for idiosyncratic government.

In his 1986 paper Zwicky does not cite a clear example of systematic lexical government and applying his criterion on systematic government, viz. that "the set of governors [...] might constitute a syntactic and/or semantic class" [Zwicky 1986, 969], will hardly obtain for the common cases of government, e.g. case assignment by prepositions, verb form selection by auxiliaries or complementizers governing a specific sentential mood. In fact, if we take into account that the distinction between subcategorization and government was based on the stipulation that government properties are never systematically paired with principles of semantic interpretation, it is tempting to conclude that government phenomena systematically building a semantic class would actually have to be taken as instances of subcategorization.

Yet another potential instance of systematic lexical government could be what in [Eisenberg 1989] is called categorial government (see section 3.1 above), e.g. the property of arbitrary nouns (members of the paradigm category SUBST) to potentially combine with a nominal attribute assigning it *genitive* case[14]. However, in an alternate analysis of genitive attribute constructions, the case assignment could be taken to result from the bare fact that the NP modifying a nominal is itself the daughter to a nominal constituent (be it NP, NGr or $\overline{\text{N}}$) and thus, as a structural property of the specific group of constituents instead of as a lexical property of the (head) noun.

If indeed there were no cases of systematic lexical government, the three-level classification of government phenomena outlined in [Zwicky 1986] would entirely coincide with the commonly acknowledged opposition between structural and lexical government. This move however would imply a crucial modification to the conception of government cited earlier from Integrational Syntax, viz. the enlargement of the domain of government sources to include not only paradigm categories but arbitrary properties of constituents.

Far more interesting than the abstract classification of government relations in the con-

[13] It is somewhat unclear though, how Zwicky motivates the choice of V and VP as the actual governors in table 7, as — speaking of vertical or structural government — one would expect the respective mother nodes (i.e. VP and S) to be the governors.

[14] Nevertheless, we note that proper nouns seem to be quite restricted in their combination with genitive attributes; the only grammatical configurations we are aware of run along the lines of

(8) Bremens Otto schlägt Bayerns Lattek.

(9) Deutschlands Steffi vertritt Berlins Diepgen.

— i.e. the combination of two proper nouns with the attribute being used rather metaphorically.

text of this paperis how Zwicky applies his concept of government to the German noun phrase. While in [Zwicky 1985] (for English) it is assumed that in Det + N neither the determiner nor the noun act as governors (thus, it follows, there is no government in the English NP), in the detailed analysis of the distribution of inflectional marks within German nominals in [Zwicky 1986, 962 and 981] it is argued, that there are in fact two distinct government relations exhibited in combinations of a determiner with (one or more) attribute adjectives and a noun. First, (similar to [Eisenberg 1989]) Zwicky realizes that gender on the noun has a very different status than on determiners and adjectives and concludes that the distribution of gender within the German nominal group is in fact determined by the invariable, lexical gender property of the head noun.

Second, in looking at the patterns of declension class variation between determiners and adjectives (and maybe, the restricted class of nouns inflecting for gender) that is traditionally accounted for in a three-valued distinction between strong, weak and mixed declension classes, [Zwicky 1986, 979 – 980] comes to the genuinely new insight, that this long-standing problem in formalizing the rules of inflectional variation within the German nominal group can be straightforwardly accounted for as lexical government induced by the determiner. As there obviously is no syntactic or semantic correlate to the assignment of a particular declension class by a given determiner, this — in the terminolgy of [Zwicky 1986] — is an instance of idiosyncratic government.

Now, what is really peculiar about this move by Zwicky is that it abandons the two most predominant assumptions held in almost all analyses of inflectional phenomena within the German nominal group: either (i) it is postulated that determiners are marked for declension class in some sense (including phonologically empty elements) and that the relation between the declension class of the determiner and attributive adjectives follows a (potentially very) complicated agreement pattern[15], or (ii) the distribution of inflectional marks in German nominals follows a principle that Zwicky calls a 'characteristic exponent approach' ("[...] the relation between determiner subtypes and adjective declensions follows from a general principle requiring characteristic — unambiguous and nonredundant — exponents of the morphosyntactic categories CASE GEND NUM" [Zwicky 1986, 972]) and shows to be too weak, though not in general unattractive, to capture the actual phenomena[16].

[15] E.g. in [Olsen 1990], citing diachronous arguments, it is called for an 'agreement chain' (holding between the determiner, attributive adjectives, the noun and all intermediate projections) interacting with dedicated principles accounting for 'morphological realization' and 'invisible categories' to maintain the declension distribution.

Somewhat more formalized but still unconvincing, [Maier and Steffens 1989] claim there be a 'disagreement' relation between the determiner and prenominal adjectives that reveals the 'janus face' of the declension class marking.

[16] In many respects the explication given in [Eisenberg 1989, 237 – 238] (with reference to work by Jósef Darski), that adjectival declension falls into a 'determinating' and a 'non-determinating' class, with the determination of case and gender being due to the adjective form only where it is not already realized on the determiner or noun the adjective combines with, is a tentative formalization of the characteristic exponent approach. Accordingly, it is not too difficult to think of examples that do not

We will come back in some more detail to the analysis put forth in [Zwicky 1986] when reviewing [Netter 1994] in section 4.2; roughly speaking, the treatment of the declension class of attributive adjectives as being lexically governed by the determiner they combine with will serve as one of the key motivations for the (DP-style) HPSG analysis put forth there.

3.2.5 The Determinant of Concord

Throughout [Zwicky 1985] we find the term CONCORD used with respect to phenomena that in [Zwicky 1986] are subsumed under the concept of AGREEMENT. As in the 1985 paper as well the term agreement is occasionally used as a synonym to concord, we take it that the conception of concord put forth by Zwicky is intended to emphasize the morphosyntactic nature of what is often called syntactic agreement: the covariation of inflectional properties on two or more constituents in combination with each other. In (nearly) avoiding the allusion to the more general concept of agreement, Zwicky obviously attempts to separate his notion of syntactic concord from genuinely semantic approaches to agreement phenomena like, for instance, the one suggested in [Pollard and Sag 1994].

Traditionally, Zwicky argues, government and concord have often and easily been confused and suggests to distinguish one from the other according to the following criterion:

> [...] in both phenomena morphosyntactic features of one constituent can determine the morphosyntactic features of a sister constituent[17], but in concord the same features are involved in the determining and the determined constituents, while in government different features are involved. [Zwicky 1985, 7]

Although not in any way formalized, this explication of the difference between government and concord nicely squares with the definitions we cited in section 3.1 from [Eisenberg 1989], in that it (i) assumes concord to be a directional determination as is government;

---

follow the postulated rule, e.g.

(10) *die starken Männer* (the strong men)

(11) *der klugen Frau* (the wise woman)

(12) *der lieben Verwandten* (the beloved relative(s))

where (10) is ambiguous between *nominative* and *accusative plural*, (11) between *genitive* and *dative singular* and (12) between *genitive* and *dative singular* and *genitive plural*. — Still, none of the adjectival forms is inflected following the so-called determinating declension pattern suggested by Eisenberg.

[17] Restricting governors to sisters of the governed elements in principle seems to disallow what in [Zwicky 1986] has been defined as vertical government. However, we already noted that even in instances of structural (hence vertical government) Zwicky somehow wants to maintain that the governor in, say, NP + VP is the verb phrase, although the assignment of *nominative* case to the NP (in English) is explained as a structural property of the whole sentence.

and (ii) explains the difference between the two relations on the basis of the properties (features for Zwicky, categories for Eisenberg) involved. However, the conception of Zwicky is more restrictive in both of the two aspects as it (i) explicitly requires that in concord phenomena one of the constituents involved serves as the DETERMINANT OF CONCORD while [Eisenberg 1989] in general does not comment on the nature of directionality in agreement relations (see above); and (ii) constrains instances of concord to the covariation of exactly the SAME properties on both the determining and the determined element, whereas in Integrational Syntax we saw that the only verbatim requirement on the categories involved in agreement is that they all be unit categories.

Another (even stricter) potential explication of the borderline between government and agreement (or concord in the sense of Zwicky) would be to require agreement phenomena to involve NON-DIRECTIONAL, i.e. SYMMETRIC relations and to think of any syntagmatic relations exhibiting the asymmetrical determination (or assignment) of properties of one element by properties of another constituent as instances of either subcategorization or government (which, as we saw, are hardly separated from each other and invariably linked within the HPSG framework). A non-directional conception of agreement would be in line with what [Barlow and Ferguson 1988, 12 – 13] suggest (but see the other papers of the volume for opposing positions) and in general better conforms to the tradition (if not spirit) in unification based grammars to account for shared information through unification (or reentrancy) of feature structures which is inherently non-directional. Moreover, conceiving of agreement phenomena as symmetric instead of directed relations seems to square nicely with commonly held intuitions on the majority of the prominent examples, viz. that, say, the systematic covariation of case and number in the nominal group acts as a wellformedness constraint on the whole construction rather than as a selectional restriction imposed by one of the elements onto its sisters. In section 3.4 we will pusue this idea somewhat further and confront it with the HPSG approach to agreement put forth in [Pollard and Sag 1994].

For the moment let us sum up how [Zwicky 1985] locates agreement relations and their alleged direction in the three sample constructions we have been studying under the microscope in the previous sections already.

First, Zwicky claims that so-called subject verb agreement in English NP + VP is a clear example of concord in that the "relevant feature [*sic*]" [Zwicky 1985, 8] (presumably number and person) is morphologically realized on both constituents[18]. However, it is in

[18] Which, obviously, is quite questionable insofar as only personal pronouns are inflectionally marked for person. For ordinary noun phrases, in fact, it is an open question whether to assume all nouns to be marked for *third* person in the lexicon or (as it is set out in [Eisenberg 1989, 286]) to think of the selection of *third* person *singular* on the finite verb as a government restriction imposed by the paradigm category of nouns.

The second alternative, however, implies the *prima facie* puzzling consequence that the distribution of properties in the combinations of finite verbs with either pronominal or truly nominal subjects in the first case is an instance of agreement but in the second follows an agreement relation with respect to the number property but a government relation regarding the person value on the verb.

the following argued, from sentences like *the penguin swims* as opposed to *the penguins swim* there is hardly any evidence on which of the constituents in NP + VP is the source and which the target of the agreement relation. So, quoting data from Swahili and the alleged fact that "the existence of inherently plural, but morphologically unmarked, nouns like *people*, together with the nonexistance of inherently singular, but morphologically unmarked, verbs, suggests that the NP is the concord determinant in English" [Zwicky 1985, 9], in NP + VP the noun phrase is assumed to determine the concord relations.

On a comparably small set of data and arguments for the V + NP case, Zwicky acknowledges that in English there is no evidence for either the verb or the noun phrase to act as the determinant of concord (as if there were agreement between English verbs and their objects at all) but — pointing to the mere fact that in Hungarian there is some covariation between V and NP — again concludes that the NP is to be considered the determinant of concord[19].

Finally for the Det + NP example, Zwicky goes back to the *this penguin* vs. *these penguins* case we have been considering in the discussion of morphosyntactic loci already (see above). Given this opposition in number and "the clear directionality of determination in languages with arbitrary gender, like French and German [...]" [Zwicky 1985, 9], it is postulated that the noun must be the determinant of concord in the nominal group. This, we note, is a by far weaker analysis of the syntagmatic relations holding within the combination of determiners and nouns than the one found in [Zwicky 1986] (reviewed briefly in section 3.2.4). In general, it seems that the location of the source of determination in syntactic agreement is more or less following what has been called the semantic argument before, since for both concepts we find the same patterns in table 7. However, we feel that the given data and the spare reference to languages other than English (without actually quoting a single example) is insufficient to motivate the stipulations put forth by Zwicky in the discussion of concord. Accordingly, in section 3.4 we will attempt to sharpen our doubts regarding a typical and explicit directionality in agreement relations and suggest an alternate, inherently symmetrical account of, at least, the agreement phenomena exhibited in German nominal structures.

3.2.6 A Provisional Conclusion

Despite the remaining vagueness, the (deliberate) lack of formalization and the questions left unanswered here and there, the review of [Zwicky 1985] and [Zwicky 1986][20] should have provided us with a better understanding of both the characteristics of the four no-

[19] Not that it was really relevant for the remaining parts of this paper, but admittedly we find it hard to follow this argumentation.

[20] Actually, the parts of the two publications that have been under discussion here make up not even half of the volumes of [Zwicky 1985] and [Zwicky 1986].

In the 1985 paper Zwicky attempts to transfer all of the head-like concepts into the morphological analysis in order to come to a comparison (and potential parallel) of headship in syntax and morphology.

tions subcategorization, morphosyntactic locus, government and concord as well as of the difficulties in applying them to (even very simple) natural language data. In the detailed discussion of the individual conceptions put forth by Zwicky and the comparison to the fundamental concepts introduced earlier from the HPSG and Integrational Syntax worlds (which despite all formal differences could be meaningfully related to each other), it should have become especially clear that (i) all of the four concepts in some form or another have to be an integral part of any theory of grammar; and that (ii) defining any one of them presupposes a sufficient degree of formalization and a reasonable inventory of theoretically primitive devices (like, say, category or constituent) in a given theory of grammar.

We found that the suggestion by Zwicky to identify headship in syntax with what has been called the morphosyntactic locus in syntactic percolation, directly coincides (in some sense as a streamlined version) with the HPSG concept of `HEAD` features and their percolation according to the HPSG head feature principle. Next, in sections 3.3 and 3.4 we will consider the questions of how subcategorization, government and agreement (in the application to German nominals) can be encoded in a well formalized theory of grammar like HPSG.

3.3 Government Relations in HPSG

If in section 3.2 it was claimed that in HPSG subcategorization and government are inseparably linked to each other, then this was only true in the sense that the interaction of `SUBCAT` lists with the HPSG subcategorization principle is the predominant selectional device in the theory. However, for some rather specific purposes additional selectional mechanisms (i.e. *category* valued selector features maintained by specialized principles and supplementary immediate dominance schemata) have been incorporated into the theory, so that one would either have to (i) broaden the coverage of the term subcategorization to include all instances of HPSG signs that through a dedicated feature on the *category* level select properties of other signs; or (ii) stay with the traditional usage to refer to the primary `SUBCAT` mechanism as SUBCATEGORIZATION while reserving a more general notion to cover all HPSG cases of licensing and selection. We suggest to take the latter move and use the term VALENCE PROPERTIES to refer to the class of HPSG devices employed in encoding both the licensing and selectional information (i.e. the combinatoric potential) of linguistic signs.

In looking at how the licensing of constituents and the selection of specific properties interact within the HPSG theory, this section will first spell out some more details on the percolation of information in phrasal signs incorporating the type *head-complement-structure* (based on an example of subcategorization) and then introduce one of the additional selectional devices, viz. the `SPEC` feature and principle, in its application to complementizers.

Likewise, besides the general discussion of government relations, [Zwicky 1986] contains a detailed overwiev of the inflectional inventory found on German determiners, adjectives and nouns and an encoding of his key insight that declension class distribution be lexically governed from the determiner in the GPSG framework.

Though maybe not at first glance obvious, the case of a complementizer governing properties of the subordinate sentence will be relevant to the analysis of German nominals, because it is an instance of a non-head that is allowed to select the shape of the head daughter it combines with, which — in one possible HPSG analysis — is very similar to the relations holding between a determiner and the head noun in the German nominal group.

Let us recall the lexicon entry for the transitive German verb *sieht* (sees) from the introductory section, repeated here for the convenience of the reader[21]:

(14)
$$\left[\begin{array}{l}\textsc{cat}\left[\begin{array}{l}\textsc{phon}\ \langle sieht\rangle\\ \textsc{syn}\ \left[\begin{array}{l}\textsc{loc}\ \left[\begin{array}{l}\textsc{head}\ \left[\begin{array}{l}\textsc{vform}\ fin\\ \textsc{tense}\ pres\end{array}\right]\\ \textsc{subcat}\ \langle \mathrm{NP[nom]}\ \mathrm{NP[acc]}\rangle\end{array}\right]\end{array}\right]\end{array}\right]\end{array}\right]$$

In section 2.3 it has already been outlined that the cancellation of complements from the verbal head and the assignment of case values is achieved in the interaction of the subcategorization principle with the appropriate immediate dominance schemata[22].

Now, how exactly is the determination of *nominative* and *accusative* case on the subject and direct object, respectively, carried out? First, the verb presumably binds its object as an instance of (a binary version of) Schema 1, yielding a phrasal sign embedding the finite verb as its head daughter (H-DTR) and a structure similar to (13) (of course, bearing lots of additional information but compatible with (13) in particular in its category, degree of saturation and case marking) as the complement daughter (C-DTR). Because the resulting phrase has a DTRS value of the type *headed-structure* (a supertype of

[21] Remember that in (14) symbols like 'NP[nom]' are merely used as a shorthand notion for *category* type objects, e.g for

(13)
$$\left[\begin{array}{l}\textsc{syn}\ \left[\begin{array}{l}\textsc{loc}\ \left[\begin{array}{l}\textsc{head}\ \left[\begin{array}{l}\textsc{case}\ \left[\begin{array}{l}\textsc{obl}\ -\\ \textsc{gov}\ -\end{array}\right]\end{array}\right]\\ \textsc{subcat}\ \langle\rangle\end{array}\right]\end{array}\right]\end{array}\right]$$

— a saturated nominal projection bearing *nominative* case.

[22] Note that through the fundamental assumption in the Saarbrücken HPSG world that all phrase structure be strictly binary branching, not only the SUBCAT principle could be substantially simplified, but similarly the set of phrase structure schemata becomes clearer.

Thus, at least for the basic type *head-complement-structure*, in a binary branching framework there can be no more than two configurations, as any head will either precede or follow its complement. Yet of course, we do not want to exclude the possibility to encode further generalizations, e.g. on the type and phrasal level, of both heads and complements in the phrase structure schemata — see [Netter 1992] for an approach to German word order in a binary branching framework.

head-complement-structure), the HPSG subcategorization principle (see section 2.4 for a feature structure representation of the binary version) will obtain, requiring that (i) the first element on the SUBCAT list of the head daughter be discarded while the tail of the list is passed up as the SUBCAT value to the mother node; and (ii) the feature structure representation of the bound (or discarded) complement be identified (i.e. unified) with the *category* of the complement daughter. Thus, by unification of the CAT value on the complement with the complement description in the subcategorization frame of the verbal head, it is achieved that, first, all properties common to both the complement description and the actual complement are guaranteed to be compatible; and, second, additional properties specified in the SUBCAT list will be percolated (or assigned) to the *category* of the object NP[23].

As the phrasal sign resulting from a transitive verb combining with its direct object still has one element on its SUBCAT list (viz. the subject), a very similar process (presumably instantiating a different immediate dominace schema) is going to allow for the cancellation of the subject NP, this time assigning it *nominative* case. So, the unification operation in the HPSG encoding of government relations really serves two distinct purposes which, intuitively, might be characterized as the selection and assignment of properties. Obviously, however, there is no meaningful way to theoretically separate one from the other in the HPSG framework. This we take to be justified from the discussion in sections 3.2.2 and 3.2.4, given the HPSG conception of the *category* type as a structured complex of morphosyntactic (and semantic) information that (lexical) heads are subcategorized for.

A slightly more complicated example for the encoding of government relations in HPSG involves complementization as it is outlined in [Pollard and Sag 1994, 37 – 41].

Diverging from standard GB assumptions, it is claimed, complementizers are best an-

[23] This also functions in the opposite direction, i.e. if the actual complement is more specific in its category than the description in the subcategorization frame. Still all properties will become shared between the two feature structures as unification is genuinely non-directional. However, in a well organized lexicon this should rarely be the case.

alyzed as non-heads instead of as the head of a complementizer phrase (CP)[24]. However, complementizers clearly impose selectional restrictions on the constituent they combine with, so that (as the SUBCAT mechanism is limited in its application to 'true' HPSG heads) [Pollard and Sag 1994] introduce a new part of speech MARKER that they characterize as follows:

> On our account, a marker is a word that is "functional" or "grammatical" as opposed to substantive, in the sense that its semantic content is purely logical in nature (perhaps even vacuous). A marker, so-called, because it formally MARKS the constituent in which it occurs, combines with another element that heads that constituent. [Pollard and Sag 1994, 39]

Markers, it is furthermore suggested, use a *category* valued HEAD feature SPEC to select the CAT value of another sign, thus governing the appropriate properties on it. Now, a marker (a complementizer, say) combining with a headed phrase yields a configuration that can not be licensed by one of the immediate dominance schemata introduced so far, as these are limited to the *daughters* subtype *head-complement-structure*. The solution adopted in [Pollard and Sag 1994] is to define a new subtype to *daughters* that will combine a head daughter with a marker daughter (M-DTR). Taking up the type definitions from section 2.3 (see page 16), this is achieved through the type specification in (19) allowing us to spell out a new phrase structure for head plus marker configurations:

---

[24] We are not really concerned about the motivation for this move in [Pollard and Sag 1994]. The main argument, however, builds on the opposition (in English) between (15) and (16):

(15) I know that he leaves as soon as possible.

(16) I demand that he leave as soon as possible.

If the complementizer were to be the head in (15) and (16), it is argued, it would be difficult to maintain that *know* obviously selects a finitely tensed sentence while *demand* governs the base form of the verbal head in its complement.

The example actually seems to be somewhat similar to German verbs that subcategorize for a PP complement headed by a particular preposition and additionally select a specific case governed by that preposition, e.g. (from [Eisenberg 1989, 78])

(17) Sie hängt an ihrer elektrischen Eisenbahn.

(18) Sie denkt an ihre Vergangenheit.

Actually, both phenomena seem to involve a sort of functional elements (viz. English complementizers and German prepositions, respectively) that adopt properties of the constituents they combine with. Thus, if one wanted to maintain the assumption that the head domain limits the scope of government relations, we would either have to stipulate two distinct prepositions an_1 vs. an_2 in the lexicon, or assume that the prepositional phrase as a whole receives case marking from its complement daughter.

An elaborate theory of functional heads in HPSG (along the lines of [Netter 1992], [Netter 1994] and [Netter 1995]) might well provide additional insight into the nature of these examples.

(19) *head-marker-structure*
$\equiv$ *headed-structure* $\sqcap$ [M-DTR *sign*[ ]]

SCHEMA 4 (HEAD + MARKER)

(20)
$$\left[\begin{array}{ll} \texttt{CAT} & \left[\texttt{SYN}\,|\,\texttt{LOC}\left[\texttt{MARKING}\ \boxed{1}\right]\right] \\ \texttt{DTRS} & \textit{head-marker-structure}\left[\begin{array}{ll} \texttt{H-DTR} & \textit{sign}[\,] \\ \texttt{M-DTR} & \left[\texttt{CAT}\,|\,\texttt{SYN}\,|\,\texttt{LOC}\left[\texttt{MARKING}\ \boxed{1}\right]\right] \end{array}\right] \end{array}\right]$$

Basically, (20) licenses the phrasal combination of a marker with a head daughter enforcing that the MARKING value on the marker daughter be identified with that of the mother. Through the feature MARKING (on the *local* level because otherwise it would be percolated up from the head instead of the marker daughter by the head feature principle) the marker (though not the head) will leave its mark on the mother node in a form indicating the type of marker involved. [Pollard and Sag 1994] suggest to account for this process involving a non-head daughter directly contributing to the category of its embedding constituent by means of a dedicated wellformedness principle that at the same time is going to account for non-marking structures in that it requires that the MARKING value of the head daughter (typically *unmarked* or similar) be passed up in case there is no marker daughter involved.

> MARKING PRINCIPLE (MP)
>
> In a headed structure, the MARKING value concides [*sic*] with that of the marker daughter if there is one, and with that of the head daughter otherwise. [Pollard and Sag 1994, 40]

Likewise, to allow the marker daughter in (20) to select for properties of the head daughter, the identification of its selector feature SPEC with the *category* of the governed element is enforced by a wellformedness principle that will be of further relevance in the [Pollard and Sag 1994] analysis of nominal structures again, viz. the

> SPEC PRINCIPLE
>
> If a non head daughter in a headed structure bears a SPEC value, it is token identical to the CAT value of the head daughter. [Pollard and Sag 1994, 40][25]

[25] As for the subcategorization principle (see 2.3), CAT has been substituted for the original SYNSEM for compatibility to the Saarbrücken top-level *sign* geometry.

Again, the government relation is carried out by the unification of two *category* type objects, very much like we saw it in the discussion of the SUBCAT mechanism. Still, of course, the subcategorization principle is going to obtain (as *head-marker-structure* is a subtype to *headed-structure*) maintaining that the unchanged SUBCAT value from the head daughter percolates up to the mother (not binding any arguments).

Assuming a lexical entry like (21) for the English complementizer *that*[26], [Pollard and Sag 1994] derive phrase structures similar to the tree sketched in (22). — In a head plus marker combination the head daughter comprises the HEAD and SUBCAT information for the mother node while the MARKING feature is percolated up from the marker daughter; furthermore the *category* of the head is identified with the SPEC value on the marker.

(21)
$$\left[\begin{array}{l}\text{CAT}\left[\begin{array}{ll}\text{PHON} & \langle \textit{that} \rangle \\ \text{SYN} & \left[\text{LOC}\left[\begin{array}{ll}\text{MARKING} & \textit{that} \\ \text{SUBCAT} & \langle\rangle \\ \text{HEAD} & \textit{marker}\left[\text{SPEC S[tensed]}\right]\end{array}\right]\right]\end{array}\right]\end{array}\right]$$

(22)
$$\left[\text{SYN}\,|\,\text{LOC}\left[\begin{array}{ll}\text{MARKING} & \boxed{1} \\ \text{SUBCAT} & \boxed{2} \\ \text{HEAD} & \boxed{3}\end{array}\right]\right]$$

M

$$\left[\text{SYN}\,|\,\text{LOC}\left[\begin{array}{ll}\text{MARKING} & \boxed{1} \\ \text{SUBCAT} & \langle\rangle \\ \text{HEAD} & \left[\text{SPEC } \boxed{4}\right]\end{array}\right]\right]$$

H

$$\boxed{4}\left[\text{SYN}\,|\,\text{LOC}\left[\begin{array}{ll}\text{SUBCAT} & \boxed{2} \\ \text{HEAD} & \boxed{3}\end{array}\right]\right]$$

3.4 Agreement in HPSG: Covariation and Structure Sharing

It has already been mentioned that [Pollard and Sag 1994] settle their account of agreement phenomena on the semantic (and, as well, pragmatic) level of linguistic signs. Hence, in reviewing the original HPSG conception of agreement we will first briefly introduce some basics of the overall geometry of feature structures of type *semantics*, the value restriction of the *category* level attribute SEM.

Originally, the HPSG account of semantics was based on recent theories of situation semantics (see [Devlin 1991] for a reasonable introduction into situation theory), so that

[26] Here, similar to earlier examples, the symbol 'S[tensed]' is to be read as an abbreviation for a feature structure description, viz. a saturated verbal projection (a sentence) of a finitely tensed (including the base form) verb.

(although a number of diversions from situation semantics together with additional descriptive devices have meanwhile been incorporated into HPSG) the structure in (23) can be read as a feature structure representation of a situation theoretic concept, viz. a restricted parameter:

(23)
$$\left[\begin{array}{ll} \texttt{INDEX} & \boxed{1}\left[\begin{array}{l}\texttt{PERSON}\ \textit{3rd}\\ \texttt{NUMBER}\ \textit{sg}\\ \texttt{GENDER}\ \textit{fem}\end{array}\right]_{\textit{index}} \\ \texttt{RESTRICTION} & \left[\begin{array}{l}\texttt{RELATION}\ \textit{girl}\\ \texttt{INSTANCE}\ \boxed{1}\end{array}\right]_{\textit{psoa}} \end{array}\right]$$

Grossly simplifying for the sake of presentation[27], (23) is taken to be the semantic contribution of the lexicon entry for the English noun *girl*. The key notion in the treatment of agreement relations (as well as in the assignment of thematic roles and coindexation in the HPSG binding theory) is the value of the attribute `INDEX`, comprising the properties that may get involved in agreement. HPSG indices, in a sense, act as logical variables (or the equivalents of discourse referents) in the semantic interpretation. To account for expletive pronouns, *index* is partitioned into subtypes *referential* and *non-referential*, so that all nominal objects are taken to introduce an index that, in the case of pronouns, may be semantically unrestricted (i.e. not incorporated into the scope of the `RESTRICTION` feature)[28].

Turning to subject verb agreement in English, [Pollard and Sag 1994] take the verbal head to employ its subcategorization frame in requiring that the subject NP bear appropriate agreement properties. Hence, using the abbreviatory notation of subscripted tags on *category* type objects to refer to the `INDEX` value (see (24)), the lexicon entry for a third person singular verb comes out as

[27] We, among others, will be omitting the `CONTEXT` anchoring conditions, quantifier storage and the fact that [Pollard and Sag 1994] assume the `RESTRICTION` attribute to be set-valued (but see [Kasper 1993] for a feature structure encoding of complex restrictions on a variable that gets by without recourse to set-valued features).

[28] Nevertheless, indices introduced by pronouns typically are contextually bound in anchoring conditions, e.g. the English personal pronoun *I* to the speaker of an utterance; *you*, in turn, to the hearer and *she* to a female animate.

(24)
$$\left[\begin{array}{l}\texttt{CAT}\left[\begin{array}{ll}\texttt{PHON} & \langle walks\rangle \\ \texttt{SYN}\,|\,\texttt{LOC} & \left[\begin{array}{ll}\texttt{HEAD} & verb\left[\texttt{VFORM}\ fin\right] \\ \texttt{SUBCAT} & \left\langle \mathrm{NP}[\mathrm{nom}]\boxed{1}\right\rangle\end{array}\right] \\ \texttt{SEM} & \left[\texttt{RESTRICTION}\left[\begin{array}{ll}\texttt{RELATION} & walk \\ \texttt{AGENT} & \boxed{1}\left[\begin{array}{ll}\texttt{PERSON} & 3rd \\ \texttt{NUMBER} & sg\end{array}\right]\end{array}\right]\right]\end{array}\right]\end{array}\right]$$

Presumably most peculiar about (24) is the fact that the verb itself is NOT marked for the properties that it — by 'agreement' — shares with its subject. Instead, it is only through the specification of a selectional restriction in the verbal subcategorization frame that eventually the identification of the subject index with the index description on the verb will enforce agreement in person and number between the verb and its subject NP. Taken to its full extent, the [Pollard and Sag 1994] notion of index agreement exhibits hardly any technical difference to what we saw in section 3.3 and accordingly, is best understood as an instance of government[29].

Obviously, loosing the distinction between government and agreement has at least two undesirable consequences: (i) as the verb form itself is unmarked with respect to person and number, there is no ground for generalizations over the systematic covariation between the morphological form and its agreement properties; and (ii) the state of affairs that in English (as well as in German) only person and number are relevant to the subject verb agreement pattern has no overt reflection in a structure like (24), simply because the *index* type statically incorporates gender too, so that, after combining a verb with its subject, by unification of the indices the AGENT variable in structure (24) will be assigned a gender property.

Furthermore, having the functor impose agreement properties as selectional restrictions onto its argument, again leads to a genuinely asymmetric (directional) explication of the agreement relation. Even if this directionality was linguistically motivated (which we doubt), the direction of — in the terminology of [Zwicky 1985] — agreement determination resulting from the [Pollard and Sag 1994] approach is just the reverse of what has been postulated in section 3.2.5; while Zwicky (not necessarily better founded) assumes the NP to be the determinant of concord in both V + NP and NP + VP combinations, the index agreement account treats the verbal head (and its projections) as inducing the covariation of morphosyntactic properties.

[29] Roughly the same characterization holds for what in [Pollard and Sag 1994] is called CASE CONCORD ("a kind of syntactic agreement that must be sharply distinguished from index agreement" [Pollard and Sag 1994, 91]) between nouns, determiners and adjectives; again, the selector (i.e. the noun for the core nominal group, the adjective for a modification of $\overline{\mathrm{N}}$) is assumed to identify its own case marking with that of the element it combines with by means of the selection mechanism (SUBCAT or MOD, respectively) — see section 4.1 for details.

Summing up our short review of the HPSG agreement fundamentals, we note that the [Pollard and Sag 1994] idea of treating agreement phenomena basically as instances of government can in no sense reflect the fundamental distinction between the two syntagmatic relations that has been drawn in sections 3.2.2 to 3.2.5. Instead, we feel that turning inherently non-directional covariation patterns into a directed selectional mechanism is more of a technical solution than a theoretically motivated decision. Accordingly, in the review of the [Pollard and Sag 1994] analyis of (German) nominal groups and the complex syntagmatic relations holding therein it will be argued that the approach fails in both (i) its prediction of the linguistic facts (i.e. the combinatoric potential and wellformedness constraints within German nominals); as well as in (ii) reflecting commonly held intuitions about how the distribution of morphosyntactic marks in determiner (adjective) noun combinations is carried out.

However, before getting into the details of the suggested HPSG analyses for nominal structures, we will first conclude the discussion of the [Pollard and Sag 1994] index agreement approach by contrasting it with an alternate representation of agreement properties which, in a sense, partly will be taken up in the DP-style analysis of German nominals proposed by [Netter 1994][30].

For languages exhibiting so-called grammatical gender (like German) there is hardly a sense in which the morphosyntactic marking of, say, the noun *Mädchen* (girl) to have *neuter* grammatical gender can be taken to be part of the semantic contribution of the use of *Mädchen* in a token utterance. Likewise for case marking and the declension class property of German adjectives and, maybe, a subclass of de-adjectival nouns: first, neither of them is to be directly semantically interpreted; and second, case and declension class are irrelevant in pronoun antecedens relations (while admittedly gender is in most cases relevant), which in HPSG (as in GB) are assumed to involve coindexation, i.e. the unification of the *index* objects, on the pronoun and its antecedents.

Now, what if we moved gender together with other properties that are involved in syntactic agreement from the semantics part of HPSG signs (viz. the `INDEX` attribute) into the syntax and, at the same time, collapsed the set of morphosyntactic properties that take part in agreement relations into a structure of its own right, say, a `HEAD` feature `AGR`? First, in the specification of an appropriate *agreement* type with respective subtypes (e.g. *noun-agreement* vs. *verb-agreement* et al.) in the lexicon it would be possible to define for any particular class of lexical items which agreement properties will be approriate for its members. Second, the clearer distinction between morphosyntactic and semantic information would allow the semantics to deploy an independent set of relevant properties (e.g.

[30] A quite similar argumentation for a more syntactic theory of agreement and feature structure geometry is to be found in the draft version of a paper by Andreas Kathol and Bob Kasper from Ohio State University ([Kathol and Kasper 1993]); although independent in its development, their analysis too comes to suggest a revision of the original [Pollard and Sag 1994] account, roughly along the lines of section 5.

I am especially grateful to Andreas Kathol for providing me with an early insight into this ongoing research.

real world sex instead of grammatical gender; cardinality instead of number) that of course could, but need not in all cases, be linked to corresponding properties in the syntax. Third, in having functors be marked for morphosyntactic agreement properties as well as their arguments, the morphology would be free to systematically relate agreement properties to inflectional variation and the resulting surface form within a sign (see presently). Fourth and finally, in the grammar there would now be a place to spell out the generalizations over common agreement patterns (and the domain of categories involved)[31], viz. as type specifications that will be inherited into the lexicon and, where necessary, relevant phrase structure schemata.

Such an approach would technically lead to a revision of the lexicon entry (24) (see above), roughly along the following line (since there will be no further reference to it, we omit the semantics in (25)):

(25)
$$\left[\begin{array}{l}\textsc{cat}\ \left[\begin{array}{ll}\textsc{phon} & \left\langle \boxed{3} \bullet \underset{\boxed{1}}{\left\{\begin{array}{l}\emptyset\\ s\\ \ldots\end{array}\right\}} \right\rangle \\ \textsc{morph} & \left[\begin{array}{ll}\textsc{stem} & \boxed{3}\langle walk\rangle \\ \textsc{infl} & \boxed{4}\left[\begin{array}{l}\textsc{person}\ \underset{\boxed{1}}{\left\{\begin{array}{l}\underset{\boxed{2}}{\left\{\left\{\begin{array}{l}1st\\ 2nd\end{array}\right\}\right\}}\\ 3rd\\ 3rd\\ \ldots\end{array}\right\}} \\ \textsc{number}\ \underset{\boxed{1}}{\left\{\begin{array}{l}\underset{\boxed{2}}{\left\{\left\{\begin{array}{l}sg\\ pl\end{array}\right\}\right\}}\\ pl\\ sg\\ \ldots\end{array}\right\}}\end{array}\right]\end{array}\right] \\ \textsc{syn}\,|\,\textsc{loc} & \left[\begin{array}{ll}\textsc{head} & \underset{verb}{\left[\begin{array}{ll}\textsc{agr} & \boxed{4} \\ \textsc{vform} & fin \\ \textsc{tense} & pres\end{array}\right]} \\ \textsc{subcat} & \left\langle \textsc{np}\left[\textsc{agr}\ \boxed{4}\left[\textsc{case}\ nom\right]\right]\right\rangle\end{array}\right]\end{array}\right]\end{array}\right]$$

[31] The concept of agreement patterns and their domains of locality (e.g. covariation in the nominal group can be captured on the level of HEAD features, whereas subject verb agreement involves the larger domain of *local* properties) is originally due to (and in fact the focus of) [Kathol and Kasper 1993].

The fundamental difference between (25) and the structure cited earlier from [Pollard and Sag 1994] is that now the verb itself is marked in its HEAD properties for number and person, rather than merely governing specific number and person values on its subject. Furthermore, the agreement properties of (25) are related to the morphological marking (the MORPH | INFL path) through the coreference $\boxed{4}$; likewise the surface string (the PHON value) is — by means of the named disjunction $\boxed{1}$[32] — linked to the morphological STEM and its inflectional properties (for the sake of readability we ignore the majority of slots in the English verbal paradigm).

Still, in (25) the verbal head technically employs the SUBCAT mechanism in identifying its own AGR value with that of its complement (at the same time governing *nominative* case), but it is now an easy task to factor the subject verb agreement pattern out of of the lexical entry into a type of its own right, thus, in a sense, really establishing agreement as a theoretically discernible concept in the HPSG ontology:

$$\textit{subject-verb-agreement} \equiv \textit{local} \sqcap \left[\begin{array}{l} \mathtt{HEAD} \quad \left[\mathtt{AGR}\ \boxed{1}_{\textit{verb-agreement}}\left[\ \right]\right] \\ \mathtt{SUBCAT}\ \left\langle \left[\mathtt{SYN} \mid \mathtt{LOC} \mid \mathtt{HEAD}\ \left[\mathtt{AGR}\ \boxed{1}_{\textit{noun-agreement}}\left[\ \right]\right]\right] \ldots \right\rangle \end{array}\right] \tag{26}$$

[32] In [Kathol and Kasper 1993], lacking distributed disjunctions, the dependency of PHON, the STEM feature (which they place on the *sign* top-level, i.e. next to PHON, SYNSEM and DTRS) and the morphosyntactic marks on the level of HEAD features is spelled out in a separate relational constraint holding between the three attributes.

As it has already been argued in section 2.4, we consider the approach using named disjunctions both linguistically more attractive (as it resembles the concept of morphosyntactic paradigms) and computationally more tractable.

4 Existing HPSG-Style Analyses of German Nominals

In the following sections we will contrast the HPSG analysis for German nominals put forth by [Pollard and Sag 1994] with the very different approach suggested in [Netter 1994]. In line with the set of data addressed in the two accounts, the range of phenomena considered will be restricted to the core nominal group, i.e. determiner plus noun combinations optionally incorporating premodifying attributive adjectives[1]. Though we will not attempt to give a formal characterization of any of the three categories involved, they will roughly be taken in the (intuitive) sense of [Eisenberg 1989]; e.g. the set of determiners in German is assumed to contain exactly[2] (the syntactic paradigms of) *der* (the), *ein* (a), *mein* (my) and *kein* (no), but, among others, NO pronouns (e.g. *dieser* (this) or *einer* (one)), numerals (*zwei* (two), *beide* (both) etc.) and adjectives (*viel* (much), *zahlreich* (plenty) etc.)[3].

For the set of morphosyntactic properties under inspection, in the following we will exclusively be concerned with (grammatical) gender, case, number and the adjectival declension class. — In being restrictive about the range of linguistic phenomena, we hope to be able to meaningfully apply some of the terminology and abstract concepts introduced in the previous sections. Still, as we will see, the syntagmatic relations that will be considered are sufficiently complex and, here and there, pose severe problems to a well formalized analysis. Finally for phrase-structural matters, especially in the review of [Netter 1994] there will be some emphasis on determinerless constructions and so-called bare plurals.

Before coming to the point, we will first have to introduce a final piece of the HPSG type lattice, viz. the [Pollard and Sag 1994] approach to modification. Very similar to the treatment of markers we saw in section 3.3, modifiers (or adjuncts) are not licensed by the constituent they combine with, but instead select the element they modify by means of the *category* valued attribute `MOD`. Accordingly, for configurations combining an adjunct with another constituent we need to define a new subtype to *headed-structure* ((1) introducing the *sign* valued feature `A-DTR`) and a suitable immediate dominace schema:

---

[1] For a broader range of linguistic data see among others [Haider 1988], [Maier and Steffens 1989], [Vater 1985] and [Vater 1986]. — However, none of them really gets the basic distribution of morphosyntactic properties in the core nominal group formally right.

[2] See [Eisenberg 1989, 160] for the motivations underlying this classification. As Walter Kasper brought to our attention, [Eisenberg 1989, 201] (in the section on pronouns) additionally introduces (the paradigms) of *dein* (yours) and *sein* (his) as what is called 'possessive determiners'.

[3] Although nothing will really hinge on this classification (and in fact, many of the generalizations over determiners in our sense can easily be applied to a number of pronouns, all numerals and a set of words on the borderline to adjectives), we take it to be given, simply not to loose the thread of the argument in classificatory side-issues.

However, in principle we acknowledge that there really seem to be elements that in some context act like a determiner but in another configuration, say, exhibit adjectival properties. Here, as often, drawing the borderline between categories may be a tricky problem.

(1) *head-adjunct-structure*
$\equiv$ *headed-structure* $\sqcap$ $\left[\text{A-DTR}\ \textit{sign}\left[\ \right]\right]$

SCHEMA 5 (HEAD + ADJUNCT)

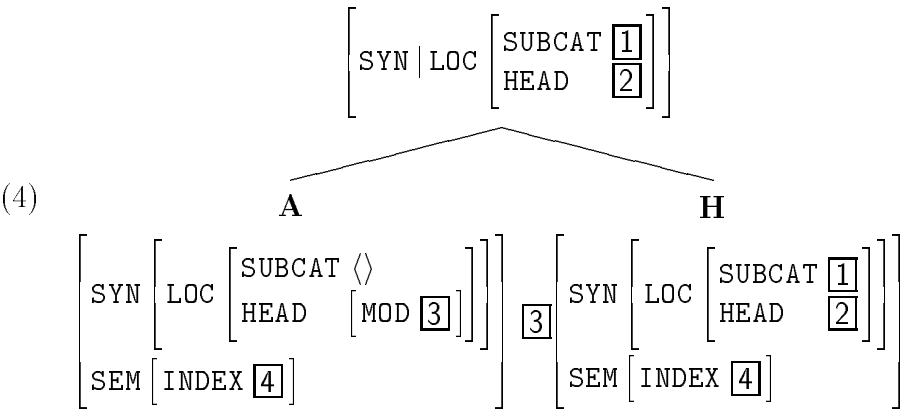
(2) $\left[\begin{array}{l}\text{DTRS} \\ \textit{head-adjunct-structure}\end{array}\left[\begin{array}{l}\text{H-DTR}\ \textit{sign}\left[\text{CAT}\ \boxed{1}\right] \\ \text{A-DTR}\left[\text{CAT}\,|\ldots|\,\text{HEAD}\left[\text{MOD}\ \boxed{1}\right]\right]\end{array}\right]\right]$

As for the marker daughter, the non-head is given access to the *category* of the head daughter in order to be able to impose selectional restrictions on it. This way the adjunct is enabled to specify through the feature structure description in its MOD feature what constituents it can modify and, once again, by unification of the two structures the compatibility of properties will be guaranteeed. According to the commonly held assumption that adjunction does NOT affect the categorial status of the element being modified (except of course for its semantics), unlike the marker daughter in Schema 5 the adjunct daughter in (2) is not allowed to determine properties of the mother node. Instead, the head feature and subcategorization principle will percolate the HEAD features and (unchanged) SUBCAT list from the head daughter up to the mother.

Leaving aside the question on how exactly to characterize the (abbreviated) *category* $\overline{\text{N}}$ in (3) (i.e. the type of constituents that can be modified by attributive adjectives) which, it will be argued, turns out to be rather problematic in the [Pollard and Sag 1994] approach, it is obvious that through the instantiation of Schema 5 with an adjectival lexicon entry like (3) as the adjunct daughter, trees along the lines of (4) will be derived. Additionally, we note that (i) the *index* objects of the head and adjunct have been identified (enforcing semantic agreement); and that (ii) Schema 5 can be applied recursively to the mother node of (4) because its syntactic properties are essentially those of the head daughter.

(3) $\left[\begin{array}{l}\text{CAT}\left[\begin{array}{l}\text{PHON}\ \langle \textit{red} \rangle \\ \text{SYN}\left[\begin{array}{l}\text{LOC}\left[\begin{array}{l}\text{HEAD}\ \textit{adjective}\left[\begin{array}{l}\text{MOD}\ \overline{\text{N}}\boxed{1} \\ \text{PRD}\ -\end{array}\right] \\ \text{SUBCAT}\ \langle\rangle\end{array}\right] \\ \text{SEM}\ \left[\text{INDEX}\ \boxed{1}\right]\end{array}\right]\end{array}\right]\end{array}\right]$

(4)



4.1 [Pollard and Sag 1994] (P&S2)

In [Pollard and Sag 1994, chapter 1] it is acknowlegded that in determiner noun combinations it is desirable to allow the determiner to select properties of the nominal that it combines with. Although in English the main motivation for this assumption is on semantic grounds (roughly speaking, determiners typically incorporate the nominal semantics into the scope of a quantifier or operator), with reference to the [Pollard and Sag 1994, chapter 2] analysis of German noun phrase agreement (see presently) it is argued that the German data provide additional arguments for giving the determiner access to the *category* of its sister constituent. However, in [Pollard and Sag 1994, 46] it is explicitly rejected to account for this state of affairs by making the determiner the head of the nominal group (treating it as a determiner phrase), but instead it is suggested to stay with a more 'conservative' analysis that is characterized as follows:

> While continuing to assume that $\overline{\text{N}}$s are the heads of NPs and subcategorize for their determiners, we will also assume that determiners reciprocally select their $\overline{\text{N}}$ sisters. We effect this selection by means of a mechanism introduced in the preceding section, viz. the SPEC feature. [Pollard and Sag 1994, 46]

What is intended, is that a noun employs its SUBCAT list to license the determiner (possibly governing specific properties on it) as a complement daughter while, at the same time, the determiner selects its head sister through the *category* valued SPEC attribute already introduced in the context of head marker configurations. Determiners according to this analysis behave very similar to markers (i.e. as functional elements; see section 3.3) in that they use the SPEC mechanism to access the CAT value of the *sign* they combine with, thus being enabled to (i) incorporate the semantics of the sister constituent into their own semantic contribution (hence, the determiner acts as the semantic functor); and (ii) select, if desirable, particular syntactic or semantic properties of the head sister.

Still, there are two major differences between the [Pollard and Sag 1994] analysis of determiner noun configurations and an instance of the head marker schema (Schema 4)

as we saw it on page 48. First, a noun binding the determiner as its argument yields a phrasal sign with a DTRS value of type *head-complement-structure*[4], whereas for head marker configurations a separate subtype of *headed-structure* has been assumed. Second, the mutual selection on the *category* level and the identification of the SUBCAT and SPEC feature structure descriptions with the CAT values of the determiner and noun, respectively, yields a CYCLIC (or infinite) feature structure (see (7)) which at least increases the computational complexity of the underlying feature logic and furthermore as a linguistic sign has a questionable ontological status. Third, other than the marker daughter in, say, a complementized sentence, the determiner in the nominal group according to the proposed analysis has no access to the (syntactic) properties of the mother node (the noun phrase as a whole) because, in that the Marking Principle does not apply, it is not allowed to leave its mark on the embedding category.

Given the (underspecified) lexicon entries for forms of the German determiner *der* (the) in (5)[5] and the common noun *Mann* (man) in (6), it should become clear how through the interaction of the head feature, subcategorization and SPEC principles the (cyclic) structure in (7) is obtained.

---

[4] Which, we note in passing, [Pollard and Sag 1994, 45] assume to be licensed by (the head + subject configuration) Schema 1. However, Schema 1 would require the head daughter to be PHRASAL (see page 16) so that, technically speaking, one would have to promote an unmodified lexical noun to a phrasal level before combining it with a determiner. As it stands, this move would presumably involve a unary (non-branching) projection transforming the head noun into what [Pollard and Sag 1994] call $\overline{\text{N}}$.

[5] We have choosen to sketch parts of the semantics of (5) as an illustration of how (i) the embedding of the nominal semantics into the scope of the *iota* operator — which is to be interpreted as a definite existential quantification — is carried out (including the identification of the respective *index* objects); and (ii) the resulting semantic contribution is put into a feature structure encoding of a Cooper or quantifier storage (which, in [Pollard and Sag 1994] actually is assumed to be set-valued); roughly speaking, the distribution of the QSTORE feature is maintained by the QUANTIFIER INHERITANCE PRINCIPLE (QIC) that for phrasal nodes requires the quantifier storage on the mother to be the union of the QSTORE values of the daughters less those operators that are locally retrieved at a node.

In the following there will be no further reference to the QSTORE attribute and to how determiners contribute to the semantics of HPSG signs.

(5)
$$\left[\begin{array}{ll}
\multicolumn{2}{l}{\left[\begin{array}{ll}
\textsc{phon} & \langle der \rangle \\
\textsc{syn}\,|\,\textsc{loc} & \left[\begin{array}{ll}
\textsc{head} & \left[\begin{array}{l}\textsc{case}\ \boxed{1}\left\{\begin{array}{c} nom \\ \{gen\ dat\} \\ gen \end{array}\right\}\end{array}\right] \\
determiner & \left[\textsc{spec}\ \overline{\mathrm{N}}\left[\textsc{sem}\ \boxed{2}\right]\right]
\end{array}\right] \\
\textsc{sem} & \boxed{3}\left[\begin{array}{ll}
\textsc{det} & iota \\
\textsc{scope}\ \boxed{2} & \left[\textsc{index}\ \boxed{1}\left[\begin{array}{l}\textsc{number}\left\{\begin{array}{c} sg \\ sg \\ pl \end{array}\right\} \\ \textsc{gender}\left\{\begin{array}{c} mas \\ fem \\ \top \end{array}\right\}\end{array}\right]\right]
\end{array}\right]
\end{array}\right]} \\
\textsc{qstore}\ \boxed{3} &
\end{array}\right]$$

(6)
$$\left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}
\textsc{phon} & \langle Mann \rangle \\
\textsc{syn}\,|\,\textsc{loc} & \left[\begin{array}{ll}
\textsc{head} & \left[\textsc{case}\ \boxed{1}\left\{\begin{array}{c} nom \\ dat \\ acc \end{array}\right\}\right] \\
noun & \\
\textsc{subcat} & \left\langle \mathrm{Det}\left[\textsc{case}\ \boxed{1}\right]\boxed{2}\right\rangle
\end{array}\right] \\
\textsc{sem} & \left[\textsc{index}\ \boxed{2}\left[\begin{array}{l}\textsc{number}\ sg \\ \textsc{gender}\ mas\end{array}\right]\right]
\end{array}\right]
\end{array}\right]$$

(7)
$$\left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}\textsc{phon} & \langle \textit{der Mann}\rangle \\ \textsc{syn}\,|\,\textsc{loc} & \left[\begin{array}{ll}\textsc{head} & \boxed{1}\left[\textsc{case}\ \textit{nom}\right] \\ \textsc{subcat} & \langle\rangle\end{array}\right]\end{array}\right] \\
\textsc{dtrs} & \left[\begin{array}{ll}\textsc{h-dtr} & \left[\begin{array}{ll}\textsc{cat}\ \boxed{2} & \left[\begin{array}{ll}\textsc{phon} & \langle \textit{Mann}\rangle \\ \textsc{syn} & \left[\textsc{loc}\left[\begin{array}{ll}\textsc{head} & \boxed{1} \\ \textsc{subcat} & \langle\boxed{3}\rangle\end{array}\right]\right] \\ \textsc{sem} & \left[\textsc{index}\ \boxed{4}\left[\begin{array}{ll}\textsc{number} & \textit{sg} \\ \textsc{gender} & \textit{mas}\end{array}\right]\right]\end{array}\right]\end{array}\right] \\
\textsc{c-dtr} & \left[\begin{array}{ll}\textsc{cat}\ \boxed{3} & \left[\begin{array}{ll}\textsc{phon} & \langle \textit{der}\rangle \\ \textsc{syn} & \left[\textsc{loc}\,|\,\textsc{head}\left[\textsc{spec}\ \boxed{2}\right]\right] \\ \textsc{sem} & \left[\textsc{scope}\,|\,\textsc{index}\ \boxed{4}\right]\end{array}\right]\end{array}\right]\end{array}\right]\end{array}\right]$$

Example (7) in fact nicely illustrates the monotonic cumulation of information, in that although both (5) and (6) are ambiguous with respect to case, number or gender, the unification of the disjunctive specifications (i.e. the intersection of the respective properties) yields a fully specified result. As, in a sense, the noun as well as the determiner are marked for the morphosyntactic properties involved, we conceive of them being shared in the noun phrase as an instance of agreement. Looking at the CASE values of (5) and (6), apparently it is impossible to think of either of them as determining the others effective case marking; — therefore, what [Pollard and Sag 1994] call case concord is a genuinely symmetric relation that, acting as a wellformedness constraint, has to hold in any determiner plus noun combination.

Turning to nominal modification and attributive adjectives, [Pollard and Sag 1994] introduce a further attribute DECL that they use in encoding declension class information. With reference to [Wunderlich 1988] and [Fenchel 1987][6], DECL is assumed to range over only the two values strong and weak, so that the traditional postulation of a third (mixed) declension class is abandoned.

In loosing the mixed declension, [Pollard and Sag 1994, 64 – 65] build on an observation that can be found in [Zwicky 1986] already, viz. the fact that all but three forms in the mixed paradigm are equivalent to the corresponding slots of the weak declension class. Furthermore, the three positions that do not correspond to the weak forms (i.e. *nominative masculine singular* and *nominative* or *accusative neuter singular*), instead can be analyzed as being taken from the strong paradigm. Accordingly, it is realized that the determiner

[6] Actually, the reference to the manuscript by Fenchel is only to be found in [Pollard and Sag 1988] which, more or less, is a prepublication of the second chapter of [Pollard and Sag 1994].

strong	*masculine*	*feminine*	*neuter*	*plural*
nominative	starker	starke	starkes	starke
genitive	starken	starker	starken	starker
dative	starkem	starker	starkem	starken
accusative	starken	starke	starkes	starke

weak	*masculine*	*feminine*	*neuter*	*plural*
nominative	schwache	schwache	schwache	schwachen
genitive	schwachen	schwachen	schwachen	schwacher
dative	schwachen	schwachen	schwachen	schwachen
accusative	schwachen	schwache	schwache	schwachen

mixed	*masculine*	*feminine*	*neuter*	*plural*
nominative	*gemischter*	gemischte	*gemischtes*	gemischten
genitive	gemischten	gemischten	gemischten	gemischten
dative	gemischten	gemischten	gemischten	gemischten
accusative	gemischten	gemischte	*gemischtes*	gemischten

Figure 3: Adjectival declension classes in the traditional three-valued distinction. The *nominative masculine singular* and *nominative* or *accusative neuter singular* forms in the mixed pattern are taken from the strong, all others from the weak paradigm.

ein (a) that is traditionally taken to establish the context for mixed adjectival inflection, in the uninflected form coincides with elements from the strong declension paradigm and in its inflected forms with those from the weak declension class.

Yet, [Pollard and Sag 1994, 91 – 95] do not understand the distribution of declension class information as a property that is lexically governed by the determiner, but instead assume uninflected forms of *ein* to be marked $\left[\texttt{DECL}\ we\right]$ and inflected forms — together with all forms of *der* — to be marked $\left[\texttt{DECL}\ st\right]$. The former class of elements they call 'weak', the latter 'strong' determiners. In this respect, given that the declension class property is assumed to be relevant to both determiners and adjectives, [Pollard and Sag 1994] make a move very similar to the approach of [Maier and Steffens 1989, 19] in that they account for the inflectional covariation between the determiner and prenominal adjectives as a sort of 'disagreement' that, as it stands, actually amounts to a government relation.

Weak adjectives, it is suggested, require the determiner to be marked strong and likewise a strong adjective is only allowed to combine with either a weak or no determiner. However,

as in the proposed constituent structure for the nominal group a lexical adjective will never be a direct sister to the determiner (but only to the head noun and its $\overline{\text{N}}$ projections), the declension class determination induced by the adjective has to follow a rather complicated path. By use of the MOD attribute, the line of argument goes, the adjective will require the noun (or $\overline{\text{N}}$) it combines with to select the appropriate declension class on the determiner. Technically, [Pollard and Sag 1994] have the noun license its determiner through the SUBCAT mechanism, so that the adjective employing its own MOD feature and the subcategorization frame of its sister constituent is given way to govern properties of the determiner that the $\overline{\text{N}}$ projection will eventually take as a complement (thus forming a noun phrase).

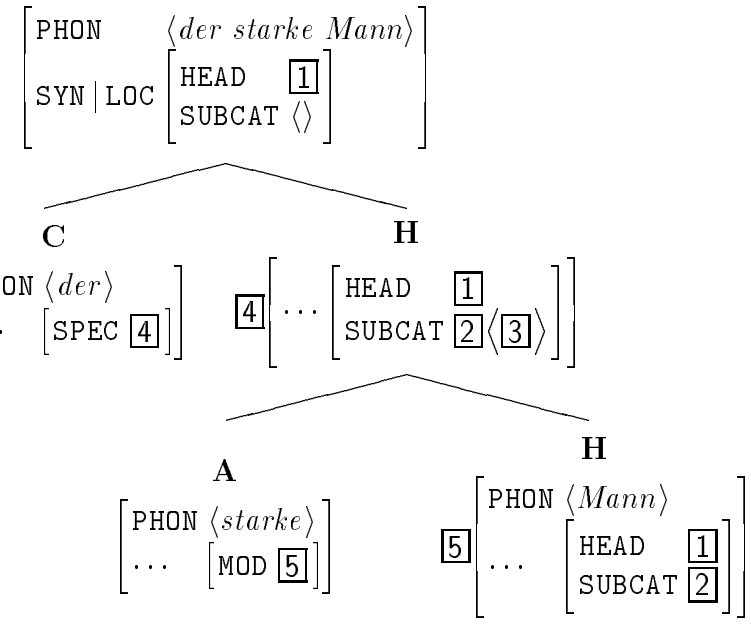
Looking at structure (8) for instance, we see how the suggested selection of strong declension on the determiner is carried out by a weak adjective. Likewise, an adjective form of the strong paradigm[7] would instead require the head noun to either subcategorize for a strong determiner or no determiner at all (i.e. constrain it to have an empty SUBCAT list).

(8)
$$\left[\begin{array}{l}\text{CAT}\left[\begin{array}{l}\text{PHON}\ \langle starke\rangle\\ \text{SYN}\ \left[\cdots\left[\text{MOD}\left[\cdots\left[\begin{array}{l}\text{HEAD}\ noun\left[\text{CASE}\ nom\right]\\ \text{SUBCAT}\ \left\langle \text{Det}\left[\text{DECL}\ st\right]\right\rangle\end{array}\right]\right]\right]\right]\end{array}\right]\end{array}\right]$$

Instantiating the head adjunct schema with, say, a weak adjective and a head noun and combining the resulting phrasal sign with a form of the strong determiner *der* (see (5)) according to Schema 1, [Pollard and Sag 1994] derive the tree sketched in (9); as we have outlined, (i) number and gender 'agreement' are achieved through the identification of INDEX values at all levels of the constituent structure; (ii) case concord is enforced between the determiner and the head noun along the lines of (6) and (7), and between the adjective and the head noun by means of the MOD mechanism (as in (8)); finally, (iii) the declension class covariation is accounted for by an (indirect) selectional restriction imposed on the determiner by the adjective.

[7] In fact in the [Pollard and Sag 1994] analysis, adjectives are only strong in the sense that they govern a weak or null determiner; the adjectival form itself however, is not assumed to be marked for DECL at all. The earlier remarks on the lack of morphosyntactic marking on verbal functors and its consequences for the syntax morphology interface (see section 3.4) therefore apply to adjectives as well.

(9)



Applying the terminology and insights from sections 3.1 and 3.2, we note that the analyis of German nominals proposed in [Pollard and Sag 1994] bears a number of limitations both in its prediction of the actual distributional data as well as in the underlying linguistic stipulations.

First, we feel that the assignment of morphosyntactic properties to the three major categories involved in the core nominal group hardly is intuitively or systematically motivated. Besides the fact that morphosyntactic properties that are inflectionally relevant for a particular category, e.g. CASE on the adjective, are often not represented as a feature of the token lexicon entry itself but rather as a selectional restriction on a sister constituent, especially the distribution of DECL marks appears to be at odds with what we find in the syntactic paradigms of determiners and adjectives. Whereas the adjectival paradigm clearly is inflectionally structured along the declension class dimension, we find little sense in treating DECL as a property of determiners (or accordingly, the classification into 'strong' vs. 'weak' determiners). Obviously, in the determiner paradigm the declension class is not a morphologically relevant property, so that, if we ever wanted to associate determiners with a specific adjectival declension class, it had to be either in the sense that (i) its inflectional material morphologically corresponds to what is found in the adjectival paradigm (or possibly nominal paradigms as well, for these traditionally have often been characterized as being strongly or weakly inflected too); or (ii) any particular form of a determiner governs a specific declension class on attributive adjectives in the nominal group as a lexical property.

Second, in the [Pollard and Sag 1994] analysis (as in the overall HPSG conception) it is often difficult to distinguish government from agreement relations. Technically speaking, all systematic covariation in the nominal group except for the semantic index agreement and case concord between the head noun and its determiner is accounted for in asymmetric

government relations. Furthermore, in how the declension class distribution between the determiner and prenominal adjectives is obtained, the governing and governed elements are actually reverted.

Third, adjectives governing the DECL property of the determiner employ a selectional mechanism that is of an inherently non-local nature. In selecting a particular declension class on the determiner complement to the head noun in the nominal group, the (non-head) adjective by means of its MOD feature governs properties in the subcategorization (i.e. government) frame of the head daughter in head adjunct structures.

Fourth, as a technical matter, adjectival forms of the strong paradigm have to impose disjunctive subcategorization information onto their sister constituent in that they require the head noun to either select for a 'weak' determiner or simply no determiner at all (but see the remarks on null determiners). Similarly, nouns that may, but need not combine with a determiner (e.g. most mass nouns and all nominal plurals) bear a disjunctive specification of their subcategorization frame in the lexicon; as we conceive of the determiner in these cases as an optional complement and an encoding of optionality in the SUBCAT list clearly is desirable outside the context of noun phrases anyhow, as opposed to [Netter 1994] (see section 4.2) we do not take the disjunctive specification in this case as a real argument against the [Pollard and Sag 1994] analysis.

Fifth, the control of adjunction to nominal projections turns out to be a testing problem. As in determinerless constructions (e.g. mass nouns and so-called bare plurals) the head noun would be lexically specified to have an empty SUBCAT list, prenominal adjectives can not require the $\overline{\text{N}}$ projection they modify to subcategorize for a determiner anymore. Accordingly, to block adjunction to a nominal group that has already combined with a determiner (as in e.g. **starke der Mann* (*strong the man)), [Pollard and Sag 1994, 94] in a footnote speculate about the postulation of phonologically empty or null determiners. Allowing for empty determiners would, of course, allow to characterize the $\overline{\text{N}}$ projection level that can be modified by attributive adjectives as comprising a non-empty SUBCAT list (and additionally would eliminate the disjunctive specification of the determiner governed by strong adjective forms). However, phonologically empty elements are rarely employed elsewhere in the HPSG theory of grammar and above all, have an ontologically questionable status and appear to be fairly controversial in contemporary linguistic theory.

Sixth, we recall that [Pollard and Sag 1994] suggest to license the combination of an unmodified noun with its determiner as an instance of Schema 1 so that they, in principle, would have to promote the lexical noun to a phrasal projection first. Nevertheless, we conceive of this as a minor unobservance in the HPSG noun phrase analysis that would presumably be addressed in a modification to Schema 1 rather than by means of a unary rule.

Finally, in a concluding remark [Pollard and Sag 1994, 94] claim that their analysis of the German noun phrase in its generative power is superior to

> [...] a generalization that is traditionally taught about so-called mixed declension patterns in German. That generalization runs something like this: A strong inflection must be realized exactly once within the NP — if the determiner is weak then the adjective must be strong; if the determiner is strong, then the adjective must be weak. [Pollard and Sag 1994, 94]

The alleged generalization (that the authors attribute to personal communication with Thomas G. Bever), it is argued, would lead to the prediction that in a series of multiple attribute adjectives within the same nominal group only the first was taken from the strong declension class, while all subsequent adjective forms were to be weakly inflected. Nevertheless, [Pollard and Sag 1994] note that in (10) actually all adjectival forms are strongly inflected, which they consider a natural consequence of the proposed analysis, in that in the recursive application of Schema 5 the selectional restrictions imposed on the determiner (i.e. the implicit encoding of the adjectival declension class) are unified in the subcategorization frame of the head noun.

(10) ein kleines$_{st}$ kluges$_{st}$ Mädchen.

(11) [mit] gutem$_{st}$ alten$_{we}$ Wein.

(12) [mit] gutem$_{st}$ altem$_{st}$ Wein.

Still, examples like (11) actually exhibit the grammatical alternation of strong and weak inflection that [Pollard and Sag 1994] question; however, we take it that in contemporary German both (11) and (12) are equally common[8].

4.2 [Netter 1994]

The analysis of agreement phenomena in the German nominal group set out in [Netter 1994] is part of a general theory of FUNCTIONAL HEADS in the framework of HPSG (see [Netter 1995]). The fundamental idea is to draw a clear distinction between the class of so-called MAJOR (substantive or lexical) categories — like verbs, nouns, adjectives and prepositions — and MINOR (or functional) categories such as determiners, numerals and complementizers (and, maybe, case marking prepositions; see page 47).

The class of functional elements in [Netter 1994] is characterized as follows:

[8] There is a substantial degree of disagreement on this issue in the literature on German grammar; while [Drosdowski 1984, 298] generally allows the alternating pattern for *dative singular masculine* or *neuter* forms, [Helbig and Buscha 1991, 303] restrict it to a limited class of quantificational adjectives (*wenig* (few), *sämtlich* (all) et al.).

As additional evidence for the uncertainty in this respect, we note that in the first edition of a textbook on German grammar (viz. [Eisenberg 1989]) one could read "beim gemischt deklinierendem unbestimmtem Artikel [...]" (page 224) whereas in the current edition the phrase has been changed to "beim gemischt deklinierenden unbestimmten Artikel [...]" (page 236).

> At least in terms of their syntactic properties and with respect to the set of lexical items realizing them, functional categories typically form close classes of elements. Although they cannot be regarded as entirely void of semantic information, their function is primarily of a syntactic nature. Functional categories typically do not occur on their own but must combine with a substantive category. However, in contrast with major categories, which may take any number of complements, functional categories combine with only one major category, which has to be fully saturated. [Netter 1994, 13]

As opposed to the [Pollard and Sag 1994] explication of markers (see section 3.3), Netter assumes that elements from the class of minor categories select their major complement by means of the standard HPSG `SUBCAT` mechanism, thus maintaining a uniform selectional mechanism and phrase structure. Additionally, the relation between a functional category and its substantive complement is defined to be of a 'parasitic' nature, in the sense that certain properties of the complement are incorporated into the set of properties comprised by the functional head itself. In encoding this kind of property raising from a complement daughter to its head sister, [Netter 1994, 11] partitions the HPSG domain of `HEAD` features into what he calls major (`MAJ`) and minor (`MIN`) properties. However, for the time being it is not in principle ruled out that (i) there be attributes that do NOT in general fall into one or the other class (e.g. the `MOD` feature); and (ii) that (lexical) categories will be allowed to be marked for both major and minor properties.

Technically, the division of major and minor properties is carried out in two complex-valued `HEAD` features that embed whatever attributes will be appropriate for a particular category. The part of speech marking, for instance[9], and the nominal agreement properties case, number and gender are assumed to be `MAJ` properties, whereas the declension class feature `DECL` is subsumed under the `MIN` attributes. An additional binary minor feature `FCOMPL` is used to encode what [Netter 1994, 14 – 15] defines as FUNCTIONAL COMPLETENESS: the indication of whether a HPSG sign has already combined with a functional head (or is functionally complete in its own right; see presently) or yet has to be bound by a functional category. Functional completeness in the explication of Netter is a primitive theoretical notion that complements the concept of saturation introduced in section 2.3; accordingly, HPSG signs need to (i) have satisfied all their subcategorization requirements; and (ii) be marked as functionally complete in order to qualify as a maximal projection.

The revised feature structure geometry at the `HEAD` feature level for, say, a noun then comes out as follows:

[9] The part of speech in [Netter 1994] is originally given in two binary attributes `N` and `V`, but for the sake of notational uniformity presently will be recoded in the style of [Pollard and Sag 1994].

(13)
$$\left[\begin{array}{l} \texttt{HEAD}\left[\begin{array}{ll} \texttt{MAJ} & \textit{nominal}\left[\begin{array}{ll} \texttt{AGR} & \textit{agreement}\left[\begin{array}{ll} \texttt{CASE} & \textit{case}[\,] \\ \texttt{NUMBER} & \textit{number}[\,] \\ \texttt{GENDER} & \textit{gender}[\,] \end{array}\right] \end{array}\right] \\ \textit{head} & \texttt{MIN}\ \textit{minor}\left[\texttt{FCOMPL}\left\{+\ -\right\}\right] \end{array}\right] \end{array}\right]$$

Building on two fundamental principles that characterize the interaction of functional and substantive categories —

> Functional Complementation
>
> In a functional category the value of its MAJ attribute is token identical with the MAJ value of its complement.
>
> Functional Completeness Constraint
>
> Every maximal projection is marked as functionally complete in its MIN feature.

— [Netter 1994, 15] suggests to treat determiners as functional heads that use the SUBCAT mechanism to select a nominal complement, adopt its major properties and mark the resulting phrase as functionally complete. Because in this setup the determiner technically licenses the presence of the noun, the state of affairs that particular nouns require a determiner to form a nominal group (e.g. *singular* forms of what is usually referred to as count nouns), while others need not (*plural* forms) or must not (so-called proper nouns) combine with a determiner has to be encoded separately.

Obviously, the binary FCOMPL attribute of Netter is suitable for exactly this purpose: as a lexical property a noun can be marked to be either (i) functionally complete on its own, thus prohibiting the combination with a functional category; (ii) functionally incomplete, i.e. requiring that it be bound by a determiner to form a full nominal group; or (iii) simply underspecified with respect to FCOMPL allowing it to stand with or without the determiner. Likewise functional completeness prevents a sign from further combination with a functional element, so that in having the determiner bear the specification $\left[\texttt{FCOMPL}\ +\right]$ sequences of multiple determiners within one nominal group are ruled out.

Common to all determiners is a structure like (14); we note that in their major properties determiners are indiscernible from nouns (especially, they share the same part of speech marking), but typically differ in their MIN values. Syntactic agreement between the determiner and its complement in the case, number and gender properties is (implicitly) enforced through the identification of the MAJ values; yet, in (14) there is no information on how the minor feature DECL is distributed.

(14)
$$\left[\begin{array}{l} \cdots \left[\begin{array}{ll} \texttt{HEAD} & \left[\begin{array}{l} \texttt{MAJ}\ \boxed{1}_{nominal}\left[\ \right] \\ \texttt{MIN}\left[\texttt{FCOMPL}\ +\right] \end{array}\right] \\ \texttt{SUBCAT} & \left\langle \left[\texttt{SYN}\,|\,\texttt{LOC}\left[\begin{array}{ll} \texttt{HEAD} & \left[\begin{array}{l} \texttt{MAJ}\ \boxed{1} \\ \texttt{MIN}\left[\texttt{FCOMPL}\ -\right] \end{array}\right] \\ \texttt{SUBCAT} & \langle\rangle \end{array}\right]\right] \right\rangle \end{array}\right] \end{array}\right]$$

In addressing the declension class covariation between the determiner and premodifying adjectives, [Netter 1994, 27 – 33] similar to [Pollard and Sag 1994] assumes DECL to be a property of both adjectives and determiners. Although in treating the determiner as the head in the nominal group, Netter would be free to account for the declension class of attributive adjectives as being lexically governed by their (functional) head sister, the assumption that the determiner too be marked for DECL in his analysis results from the stipulation that the declension class property is relevant to the maximal projection as a whole. This might appear at first glance to be an odd idea, but is motivated in the proposed account of determinerless constructions.

Strictly rejecting the postulation of phonologically empty determiners in mass or proper noun constructions and bare plurals, [Netter 1994, 20] suggests to account for optional determiners simply by underspecification of the FCOMPL attribute in the lexicon. Thus, a noun that was lexically not specified for its functional completeness property would be free to either combine with a determiner or form a maximal saturated projection on its own. A determinerless nominal could still incorporate inflected adjectival forms, but then there would be no functional head to govern their declension class. In accounting for these configuration, [Netter 1994, 29] requires the maximal projection itself to be marked for strong inflection, so that either the determiner is in a position to mediate between the declension class property of the whole phrase and its $\overline{\text{N}}$ complement, or — in a configuration lacking the determiner — the requirement for strong inflection is percolated down the projection line of the lexical noun[10].

Reconsidering the examples introduced in section 4.1, in the [Netter 1994] approach we get the structures (15) to (17)[11]

[10] In a sense the stipulation that the noun phrase as a whole be marked for strong inflection can be taken as a gross feature structure encoding of what in [Zwicky 1986] is called the characteristic exponent approach (section 3.2.4), the determinating adjectival inflection in [Eisenberg 1989] or the alleged generalization of [Pollard and Sag 1994] quoted earlier. Clearly the strong inflection pattern in comparison to the weak paradigm has fewer syncretism and, hence, more overtly contributes to the morphosyntactic marking of the nominal group.

In configurations where there is no determiner to mediate between the phrasal declension marking and the requirements on the inflection of adjectival forms at the $\overline{\text{N}}$ level, the propagation of the [DECL *st*] postulation from the maximal nominal projection down to the lexical adjective — in the terminology of [Zwicky 1986] — turns out to be a vertical (or structural) government relation.

[11] For representational simplicity presently we give only the *nominative singular masculine* forms. In

(15)
$$\left[\begin{array}{l}\texttt{CAT}\left[\begin{array}{ll}\texttt{PHON} & \langle der\rangle\\ \texttt{SYN}\,|\,\texttt{LOC} & \left[\begin{array}{ll}\texttt{HEAD} & \left[\begin{array}{ll}\texttt{MAJ} & \boxed{1}\\ & nominal\end{array}\right]\left[\texttt{AGR}\left[\begin{array}{ll}\texttt{CASE} & nom\\ \texttt{NUMBER} & sg\\ \texttt{GENDER} & mas\end{array}\right]\right]\\ & \texttt{MIN}\left[\begin{array}{ll}\texttt{DECL} & st\\ \texttt{SPEC} & +\\ \texttt{FCOMPL} & +\end{array}\right]\end{array}\right]\\ & \texttt{SUBCAT}\left\langle\left[\cdots\left[\begin{array}{l}\texttt{MAJ}\ \boxed{1}\\ \texttt{MIN}\left[\begin{array}{ll}\texttt{DECL} & we\\ \texttt{FCOMPL} & -\end{array}\right]\end{array}\right]\right]\right\rangle\end{array}\right]\end{array}\right]$$

(16)
$$\left[\begin{array}{l}\texttt{CAT}\left[\begin{array}{ll}\texttt{PHON} & \langle starke\rangle\\ \texttt{SYN}\,|\,\texttt{LOC} & \left[\begin{array}{ll}\texttt{HEAD} & \left[\begin{array}{ll}\texttt{MAJ} & \left[\texttt{AGR}\ \boxed{1}\left[\begin{array}{ll}\texttt{CASE} & nom\\ \texttt{NUMBER} & sg\\ \texttt{GENDER} & mas\end{array}\right]\right]\\ & adjective\end{array}\right]\\ & \texttt{MIN}\left[\texttt{DECL}\ \boxed{2}\,we\right]\\ & \texttt{MOD}\left[\cdots\left[\begin{array}{l}\texttt{MAJ}\left[\texttt{AGR}\ \boxed{1}\right]\\ \texttt{MIN}\left[\begin{array}{ll}\texttt{DECL} & \boxed{2}\\ \texttt{SPEC} & -\end{array}\right]\end{array}\right]\right]\end{array}\right]\end{array}\right]\end{array}\right]$$

(17)
$$\left[\begin{array}{l}\texttt{CAT}\left[\begin{array}{ll}\texttt{PHON} & \langle Mann\rangle\\ \texttt{SYN}\,|\,\texttt{LOC} & \left[\begin{array}{ll}\texttt{HEAD} & \left[\begin{array}{ll}\texttt{MAJ} & \boxed{1}\\ & nominal\end{array}\right]\left[\texttt{AGR}\left[\begin{array}{ll}\texttt{CASE} & nom\\ \texttt{NUMBER} & sg\\ \texttt{GENDER} & mas\end{array}\right]\right]\\ & \texttt{MIN}\left[\texttt{FCOMPL}\ -\right]\end{array}\right]\end{array}\right]\end{array}\right]$$

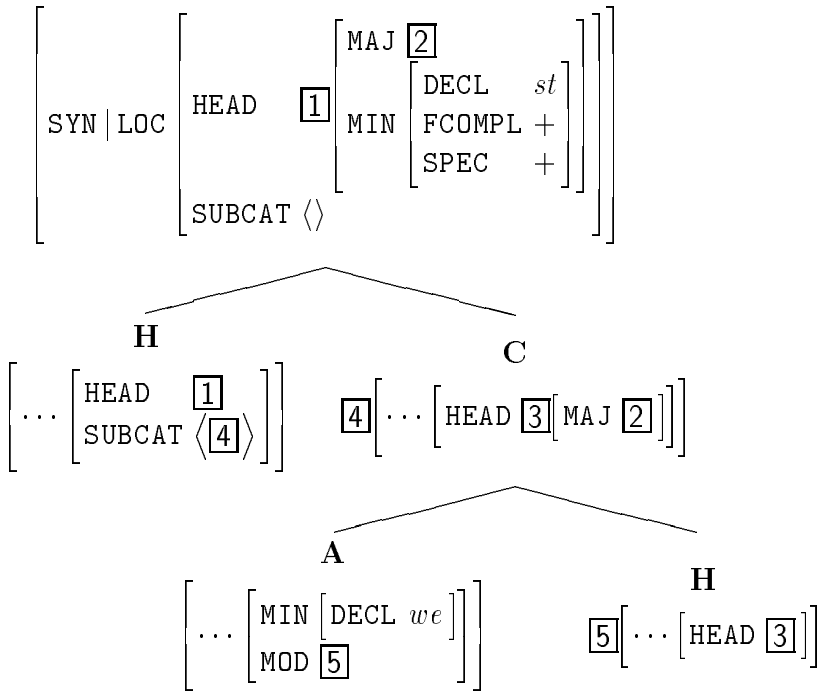
Looking at the declension class distribution, we find that in (15) the determiner is itself marked for $\left[\texttt{DECL}\ st\right]$ but in its SUBCAT list governs the $\overline{\text{N}}$ complement to be weakly inflected. The adjective (16) accordingly is marked for weak inflection and employing the

section 4.1 we have already seen how in the unification of morphologically ambiguous categories a fully specified nominal group is obtained.

MOD mechanism (see above) propagates its declension class (via the HEAD features of the constituent modified and the head feature principle) to the $\overline{\text{N}}$ node.

Assuming a slightly modified head-initial version of Schema 1 (or possibly a new phrase structure schema accounting for functional complementation), in combining (17) with (16) along the lines of the head adjunct schema and having (15) bind the resulting $\overline{\text{N}}$ node as a functional complement we get the tree:

(18)



Yet, a comment is necessay on the attribute SPEC incorporated as a minor feature in both the determiner and adjective forms (15) and (16). In characterizing the $\overline{\text{N}}$ level that prenominal adjectives can modify, Netter neither can rely on a non-empty SUBCAT list (as [Pollard and Sag 1994] do in postulating a null determiner in what Netter analyses as a truely determinerless construction), nor can the concept of functional completeness be practically employed. The first alternative is blocked simply because the determiner is not subcategorized for at all. Therefore, all nouns comprise a saturated subcategorization frame at the $\overline{\text{N}}$ level already. Complementary, if the adjective in its MOD feature would require the constituent it modifies to be functionally incomplete, first, adjunction to nouns that are marked $\left[\text{FCOMPL} +\right]$ in the lexicon (e.g. proper nouns if we were to assume that phrases like *der Peter* are ungrammatical) would be ruled out; and, second, a category that was lexically underspecified with respect to FCOMPL by unification with the MOD value of an adjective would be constrained to $\left[\text{FCOMPL} -\right]$ so that it could no longer form a noun phrase without combining with a determiner first.

Hence, in order to prevent adjectival modification of a phrase that has already been

bound by a determiner (e.g. *starke der Mann*), [Netter 1994, 25] introduces the binary feature SPEC (not to be confused with the *category* valued attribute of the same name employed in the analysis of [Pollard and Sag 1994]) that records whether a sign yet has been SPECIFIED by a determiner or not. Since determiners and nouns in the analysis of Netter share the same set of major features and part of speech, it follows that SPEC has to be at the level of MIN features, so that the determiner is free to introduce it in the process of functional complementation (see (15)). Since adjectival forms require that the constituent they modify be marked $\left[\text{SPEC} -\right]$ in their MOD feature (see (16)), only adjunction to nouns or nominal projections that have not been combined with a determiner before is allowed.

Summing up the short review of the analysis put forth by Netter, we recall the specification of maximal nominal projections given in the type definition (19)[12]. Actually, we consider it a misnomer that [Netter 1994, 19] associates the type name *dp* (suggesting to think of a determiner phrase) with (19), because (i) through the categorial identification of determiners and nouns at the level of major properties *dp* equally well subsumes 'pure' noun phrase configurations as the one in (21) (see presently); and (ii) in a phrase functionally headed by a determiner the MAJ properties of the maximal projection still are percolated up from the lexical noun.

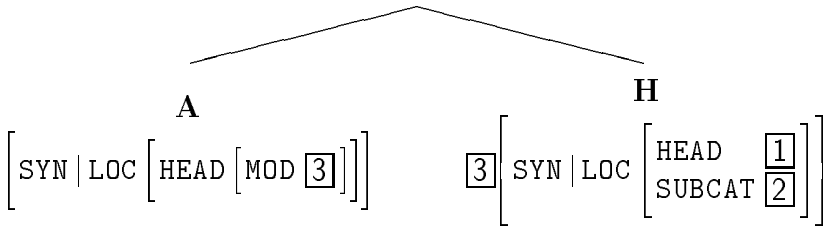
(19)
$$dp\left[\text{CAT}\left[\text{SYN}\mid\text{LOC}\left[\begin{array}{l}\text{HEAD}\left[\begin{array}{ll}\text{MAJ} & nominal\left[\ \right]\\ \text{MIN} & \left[\begin{array}{ll}\text{DECL} & st\\ \text{FCOMPL} & +\end{array}\right]\end{array}\right]\\ \text{SUBCAT}\ \langle\rangle\end{array}\right]\right]\right]$$

It should be fairly clear now, how [Netter 1994, 33] derives a determinerless construction headed by a plural form of the noun *Mann* that in the lexicon is underspecified regarding its functional completeness (see (20)). As a matter of fact, the mother node in (21) without a further rule application is directly compatible with the requirements on maximal nominal projections as they have been specified in the *dp* type definition; since by application of the head feature principle the HEAD features of the phrasal node in (21) are token identical to those of the lexical head, the (plural) form *Männer* on its own would qualify as a maximal 'projection' as well.

[12] In passing, we note that the specification of properties that have to be satisfied in the maximal projection of a particular category constitutes a descriptive device that is an extension to the standard HPSG inventory. However, the *dp* type of Netter could either by type inheritance be easily incorporated into the subcategorization requirements of heads that take nominal complements, or — presumably preferably — be intepreted corresponding to the start symbol (e.g. 'S') in traditional phrase structure grammar at the phrasal level.

(20)
$$\left[\begin{array}{l}\text{CAT}\left[\begin{array}{ll}\text{PHON} & \langle \textit{Männer} \rangle \\ \text{SYN}\,|\,\text{LOC} & \left[\begin{array}{ll}\text{HEAD} & \left[\begin{array}{ll}\text{MAJ} & \boxed{1}\left[\begin{array}{ll}\text{AGR} & \left[\begin{array}{ll}\text{NUMBER} & \textit{pl} \\ \text{GENDER} & \textit{mas}\end{array}\right]\end{array}\right] \\ & \textit{nominal} \\ \text{MIN} & \left[\text{FCOMPL}\left\{+\ -\right\}\right]\end{array}\right]\end{array}\right]\end{array}\right]\end{array}\right]$$

(21)
$$\left[\begin{array}{ll}\text{SYN}\,|\,\text{LOC} & \left[\begin{array}{ll}\text{HEAD} & \boxed{1}\left[\begin{array}{ll}\text{MAJ} & \textit{nominal}\left[\ \right] \\ \text{MIN} & \left[\begin{array}{ll}\text{DECL} & \textit{st} \\ \text{FCOMPL} & +\end{array}\right]\end{array}\right] \\ \text{SUBCAT} & \boxed{2}\langle\rangle\end{array}\right]\end{array}\right]$$



In contrasting, we find that the analyis of [Netter 1994] overcomes most of the problematic issues noted in the review of the [Pollard and Sag 1994] approach. Above all, the declension class distribution is linguistically adequately accounted for as being lexically governed by determiners and not vice versa. Furthermore, the account of Netter is free of the 'non-local government' relation induced by adjectives in the [Pollard and Sag 1994] analysis, disjunctive specifications in the nominal subcategorization frame, phonologically empty elements (null determiners) and a unary branching constituent structure.

The employment of the attribute SPEC to block adjunction to nominal groups that have already combined with a functional head may at first glance look like a rather technical solution that partly duplicates the FCOMPL mechanism. However, (i) as [Netter 1994, 41] notes, the introduction of $\left[\text{SPEC}\ +\right]$ by the determiner in a way resembles the HPSG MARKING principle postulated for functional elements (see section 3.3) except that — the determiner being the head of the construction — there is no need for a specialized immediate dominace schema and structural percolation principle; and (ii) the encoding of the three relevant projection levels in the nominal group in a pair of two binary properties directly relates to the suggestion in [Haider 1988] to represent the (GB-style) nominal bar level in the combination of binary features MIN and MAX.

The only remaining point that appears to us to be questionable in the analysis suggested by Netter is how it accounts for the licensing of determiners. In the functional approach

to the determiner plus noun combination we find that the HPSG SUBCAT list actually no longer is employed to license the $\overline{\text{N}}$ complement, but rather is reduced to a technical device enforcing the raising of the complement MAJ properties and compatibility of declension class and functional completeness information. In fact, it is apparent that it is still the noun licensing the determiner and not vice versa, viz. by means of the FCOMPL feature, and since in the nominal group there are no more than three alternatives in combining with a determiner (it can be either obligatory, optional or prohibited), its licensing can indeed be encoded in a binary property on the lexical noun (using underspecification for optionality).

Nevertheless, as in the [Netter 1994] analysis the determiner properly governs the declension class of its $\overline{\text{N}}$ complement, the SUBCAT specification still acts as a selectional restriction on the constituent that the determiner combines with; accordingly, we see no major reservations to the employment of the subcategorization frame on the determiner in the approach of Netter, but note that with regard to licensing phenomena in the German nominal group it is not that fundamentally different to the [Pollard and Sag 1994] mutual selection account as it superficially appears.

5 Conclusion: a Weak NP-Analysis

Summing up our review of analyses for the core German nominal group that have recently been suggested in the HPSG literature, we note that the substantial number of morphosyntactic properties involved in both government and agreement relations poses a number of serious questions to a well formalized theory of grammar. Still, we think that — abstracting away from the details of the feature structure geometry — the approaches of [Pollard and Sag 1994] and [Netter 1994] agree in two of the very fundamental assumptions. This observation, given that the former account claims to be a 'conservative' NP analysis while the latter, in a sense, can be characterized as a DP-style approach, may at first glance look rather surprising.

First, we note that both analyses assign the lexical noun a superior status in the nominal group in that it (i) is assumed to be the licensor of the determiner (where it is appropriate); and (ii) is taken to comprise the substantial characterization of the whole phrase (either — in the [Pollard and Sag 1994] analysis — as the set of HEAD features or — according to [Netter 1994] — in the cluster of MAJ properties), i.e. the part of speech and morphosyntactic marking. In this sense, a noun phrase is a noun phrase, is a noun phrase.

Second, in both [Pollard and Sag 1994] (or a slightly streamlined version of their analysis; see presently) and [Netter 1994], intuitively speaking, there are two head domains to be distinguished: the level of $\overline{\mathrm{N}}$ and the maximal nominal projection. Assuming the complete nominal group to incorporate one or more prenominal adjectives, in the approach of Netter the two distinct domains will be overtly realized in that the intermediate phrasal projection is headed by the noun, whereas for the whole the determiner will serve as the head. In the [Pollard and Sag 1994] mutual selection analysis, the lexical noun technically will be the head of both the $\overline{\mathrm{N}}$ and the nominal group as a whole. Still, obviously the determiner is (through the SPEC feature and principle) given way to (i) govern properties of the nominal projection that it combines with; and (ii) — at least in the modification of the SPEC principle suggested presently — determine properties of the resulting noun phrase.

Presumably this cross-dependency within the German nominal group, we suggest, is among the fundamental reasons for the traditional lack of consensus on whether it is appropriate to treat the noun or the determiner as the head in the nominal group (i.e. the NP vs DP opposition). Therefore, we claim, in a reasonably formal analysis of the German data neither of the two alternatives will account for the complex distributional facts exclusively. Instead, certainly the division of head domains indicated above will in one way or the other have to be taken into account, e.g. as for [Netter 1994] and [Pollard and Sag 1994] in partioning the set of HEAD features or assuming mutual selection, respectively.

We argue that the two HPSG analyses reviewed in sections 4.1 and 4.2 are weakly equivalent in their generative capacity and with respect to the range of data they can been applied to (although the differences are clearly more than just technical ones). In further pursuing both of them — i.e. in the application to more complex phenomena within the nominal group (determiner-like elements, numerals, possessives, adjunction to the specifier

position et al.) — one would expect to gain further insight into which of the underlying stipulations can be proven to be linguistically most adequate.

For the time being, let us conclude the study of German nominal structures in the light of syntactic categories and syntagmatic relations with a very brief outlook on how the [Pollard and Sag 1994] noun phrase analysis could be rephrased to incorporate the suggestions for a more explicit HPSG account of syntactic agreement made earlier in the review of the *index* agreement approach.

In section 3.4 it was suggested to integrate the morphosyntactic properties that for a particular category take part in agreement relations in a feature AGR at the level of HEAD features. Common agreement patterns, it was argued, could then be encoded in type specifications that for a given category identify the relevant features and — by type inheritance — become incorporated into the actual lexicon entries.

Thus, assuming that determiners, adjectives and nouns were all independently marked for case, number and gender (and adjectives and certain nouns additionally for the declension class property), syntactic agreement[1] with the head noun could be obtained using the SPEC mechanism on the determiner (see (1)). Likewise would adjective noun agreement be enforced by means of the adjectival MOD feature along the lines of (2).

determiner-noun-agreement ≡ *head* ⊓

(1)
$$\left[\begin{array}{l}\mathtt{AGR}\ \boxed{1}_{\textit{determiner-agreement}}[\,] \\ \mathtt{SPEC}\left[\mathtt{SYN}\,|\,\mathtt{LOC}\,|\,\mathtt{HEAD}\left[\mathtt{AGR}\ \boxed{1}_{\textit{noun-agreement}}[\,]\right]\right]\end{array}\right]$$

adjective-noun-agreement ≡ *head* ⊓

(2)
$$\left[\begin{array}{l}\mathtt{AGR}\ \boxed{1}_{\textit{adjective-agreement}}[\,] \\ \mathtt{MOD}\left[\mathtt{SYN}\,|\,\mathtt{LOC}\,|\,\mathtt{HEAD}\left[\mathtt{AGR}\ \boxed{1}_{\textit{noun-agreement}}[\,]\right]\right]\end{array}\right]$$

In order to explain the strong adjectival inflection in determinerless constructions (while in the presence of a determiner the inflection of premodifying adjective forms would simply be lexically governed in the SPEC attribute) without the recourse to phonologically empty elements, we could make a stipulation similar to that of [Netter 1994], viz. that the nominal

[1] As do the analyses of [Pollard and Sag 1994] and [Netter 1994], we leave aside the question on how to treat gender in this respect. If it was really desirable, in making reference to the individual AGR features rather than to the whole cluster and have the head noun employ its subcategorization frame to impose its gender onto the determiner, one could of course account for the gender covariation as a government relation. However, as both agreement and government are carried out through unification, in the resulting nominal group there would hardly be a visible difference.

group as a whole is required to be marked for strong inflection. But because in the mutual selection analysis the determiner is not in a position to mediate between the declension class properties of the $\overline{\text{N}}$ level and the maximal projection, respectively, in our setup we would have to impose strong inflection in a more bottom-up fashion, i.e. depending on the subcategorization frame of the head noun. In encoding that in a determinerless configuration the $\overline{\text{N}}$ phrase has to be strongly inflected, we could either link the declension class property to the licensing of the determiner in the lexicon entry for nouns (in a type like (3)) or postulate an implicational constraint that would require weakly inflected nominal projections (that have not yet combined with a determiner; see presently for the MARKING value) to bear a non-empty SUBCAT list (see (4)). Since the former is theoretically and computationally simpler, it would presumably be preferable.

(3)
$$\left[\text{CAT}\,|\,\text{SYN}\,|\,\text{LOC}\left[\begin{array}{ll}\text{HEAD} & \left[\text{DECL}\left\{\begin{array}{c}st\\we\end{array}\right\}\right]\boxed{1}\\ \text{SUBCAT} & \left\{\begin{array}{c}\{\langle\text{Det}\rangle\ \langle\rangle\}\\ \langle\text{Det}\rangle\end{array}\right\}\boxed{1}\end{array}\right]\right]$$

(4)
$$\mathit{local}\left[\begin{array}{l}\text{HEAD}\ \mathit{noun}\left[\text{DECL}\ we\right]\\ \text{MARKING}\ -\end{array}\right] \Longrightarrow \left[\text{SUBCAT}\ \langle\text{Det}\rangle\right]$$

Finally, to block adjunction to phrases that have already combined with a determiner (since we analyse determinerless constructions not as comprising a null determiner, the $\overline{\text{N}}$ projection level can no longer be characterized in terms of a non-empty SUBCAT list), we suggest to allow the determiner to mark the noun phrase as a whole as an instance of an extended version of the HPSG SPEC principle (see section 3.3 for the original wording):

> SPEC PRINCIPLE (REVISED VERSION)
>
> If a non head daughter in a headed structure bears a SPEC value, it is token identical to the CAT value of the head daughter and the MARKING value is token identical to that of the mother[2].

We leave it as an exercise to the untired reader to see how in this reformulation of the [Pollard and Sag 1994] mutual selection account the problems noted in section 4.1 are adequately addressed, while the same weak generative capacity is preserved.

[2] The proposal to extend the MARKING mechanism to determiner noun configurations is due to [Netter 1994, 41]. Looking at how [Pollard and Sag 1994] characterize the class of markers (page 47) together with the fact that determiners share the SPEC mechanism with markers, it indeed is a tempting move to generalize over the class of functional elements.

References

[Abney 1987] Steven Paul Abney. The English Noun Phrase in its Sentential Aspect. Mit working papers in linguistics, Massachusetts Institute of Technology (Cambridge, Massachusetts), 1987.

[ACL 1993] *Proceedings of the 31th Meeting of the Association for Computational Linguistics (ACL)*, Columbus, Ohio, 1993.

[Barlow and Ferguson 1988] Michael Barlow and Charles A. Ferguson (editors). *Agreement in Natural Languages. Approaches, Theories, Descriptions.* Stanford, California (Center for the Study of Language and Information), 1988.

[Barlow 1988] Michael Barlow. Unification and Agreement. CSLI Report # 120, Center for the Study of Language and Information (Stanford, California), 1988.

[Bouma 1988] Gosse Bouma. Modifiers and Specifiers in Categorial Unification Grammar. *Linguistics*, 26 : 21–46, 1988.

[Bresnan 1982] Joan Bresnan (editor). *The Mental Representation of Grammatical Relations.* Cambridge, Massachusetts (MIT Press), 1982.

[Brettschneider and Lehmann 1980] Gunter Brettschneider and Christian Lehmann (editors). *Wege zur Universalienforschung. Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hans-Jakob Seiler.* Tübinger Beiträge zur Linguistik # 145. Tübingen (Gunter Narr Verlag), 1980.

[Busemann and Harbusch 1993] Stephan Busemann and Karin Harbusch. Proceedings of the DFKI Workshop on Natural Language Systems: Reusability and Modularity. DFKI Document # 107, Deutsches Forschungszentrum für Künstliche Intelligenz (Saarbrücken), 1993.

[Carpenter 1992] Bob Carpenter. *The Logic of Typed Feature Structures. With Applications to Unification Grammars, Logic Programs and Constraint Resolution.* Cambridge Tracts in Theoretical Computer Science # 24. Cambridge, New York, Port Cheaster, Melbourne, Sidney (Cambridge University Press), 1992.

[COLING 1984] *Proceedings of the 10th Conference on Computational Linguistics (COLING)*, Stanford, California, 1984.

[COLING 1986] *Proceedings of the 11th Conference on Computational Linguistics (COLING)*, Bonn, 1986.

[Cooper 1990] Richard Cooper. Specifiers, Complements and Adjuncts in HPSG. EUCCS RP # 42, Centre for Cognitive Science, University of Edinburgh (Edinburgh), 1990.

[Devlin 1991] Keith Devlin. *Logic and Information.* Cambridge, New York, Port Cheaster, Melbourne, Sidney (Cambridge University Press), 1991.

[Drosdowski 1984] Günther Drosdowski (editor). *DUDEN. Grammatik der deutschen Gegenwartssprache.* Mannheim, Wien, Zürich (Bibliographisches Institut), 4. völlig neu bearbeitete und erweiterte Auflage, 1984.

[Eisenberg 1989] Peter Eisenberg. *Grundriß der deutschen Grammatik.* Stuttgart (Metzlersche Verlagsbuchhandlung), 2. überarbeitete und erweiterte Auflage, 1989.

[Fanselow and Felix 1990] Gisbert Fanselow and Sascha W. Felix (editors). *Strukturen und Merkmale syntaktischer Kategorien.* Studien zur deutschen Grammatik # 39. Tübingen (Gunter Narr Verlag), 1990.

[Fenchel 1987] Klaus Fenchel. Agreement in the German Noun Phrase. (Unpublished Manuscript, Linguistic Institute, Stanford University), 1987.

[Gazdar *et al.* 1985] Gerald Gazdar, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag. *Generalized Phrase Structure Grammar.* Oxford (Basil Blackwell), 1985.

[Görz 1992] Günther Görz (editor). *Tagungsband KONVENS 92. 1. Konferenz "Verarbeitung Natürlicher Sprache"*, Nürnberg, 1992. Springer Verlag.

[Grosz and Stickel 1983] Barbara J. Grosz and Mark E. Stickel (editors). *Research on Interactive Acquisition and Use of Knowledge.* Menlo Park, California (Artificial Intelligence Center, SRI International), 1983.

[Haider 1988] Hubert Haider. Die Struktur der deutschen Nominalphrase. *Zeitschrift für Sprachwissenschaft*, 7(1):32–59, 1988.

[Helbig and Buscha 1991] Gerhard Helbig and Joachim Buscha. *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht.* Leipzig, Berlin, München (Verlag Enzyklopädie — Langenscheidt), 13. durchgesehene Auflage, 1991.

[Jacobs *et al.* 1993] Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann (editors). *Syntax. Ein Handbuch internationaler Forschung.* Berlin, New Yorck (Walter de Gruyter), 1993.

[Kaplan and Bresnan 1982] Ronald Kaplan and Joan Bresnan. Lexical Functional Grammar: A Formal System for Grammatical Representation. In: [Bresnan 1982].

[Kasper 1993] Walter Kasper. Integration of Syntax and Semantics in Feature Structures. [Busemann and Harbusch 1993].

[Kathol and Kasper 1993] Andreas Kathol and Robert Kasper. Agreement and the Syntax-Morphology Interface in HPSG. Unpublished Manuscript. Draft of February 7, 1993.

[Kay 1986] Martin Kay. Functional Unification Grammar. A Formalism for Machine Translation. In: [COLING 1986].

[Krieger *et al.* 1993] Hans-Ulrich Krieger, John Nerbonne, and Hannes Pirker. Feature-Based Allomorphy. In: [ACL 1993], pages 140–147.

[Lieb 1980] Hans-Heinrich Lieb. Words as Syntactic Paradigms. In: [Brettschneider and Lehmann 1980], pages 115–123.

[Lieb 1983] Hans-Heinrich Lieb. *Integrational Linguistics. Volume I: General Outline.* Current Issues in Linguistic Theory # 17. Amsterdam, Philadelphia (John Benjamins Publishing Company), 1983.

[Lieb 1993] Hans-Heinrich Lieb. Integrational Linguistics. (Manuskript. Nur für den Privatgebrauch. To appear in [Jacobs *et al.* 1993]), 1993.

[Maier and Steffens 1989] Petra Maier and Petra Steffens. Zur Syntax pränominaler Elemente in einer kategorialen Unifikationsgrammatik. IWBS Report # 70, Institut für Wissensbasierte Systeme (Stuttgart), 1989.

[Nerbonne *et al.* 1994] John Nerbonne, Klaus Netter, and Carl Pollard (editors). *German in Head-Driven Phrase Structure Grammar.* CSLI Lecture Notes # 46. Stanford, California (Center for the Study of Language and Information), 1994.

[Netter 1992] Klaus Netter. On Non-Head Non-Movement. An HPSG Treatment of Finite Verb Position in German. In: [Görz 1992], pages 218–227.

[Netter 1993] Klaus Netter. Architecture and Coverage of the DISCO Grammar. [Busemann and Harbusch 1993].

[Netter 1994] Klaus Netter. Towards a Theory of Functional Heads. German Nominal Phrases. In: [Nerbonne *et al.* 1994].

[Netter 1995] Klaus Netter. A Theory of Functional Heads in HPSG (Working title). Doctoral Dissertation, Computational Linguistics Department, Universität des Saarlandes (forthcoming), 1995.

[Oepen 1991] Stephan Oepen. Unifikation und Grammatik. Beispiele moderner Beschreibungsformalismen. Unpublished Course Paper, School for German Studies, Freie Universität Berlin, 1991.

[Olsen 1990] Susan Olsen. AGR(eement) und Flexion in der deutschen Nominalphrase. In: [Fanselow and Felix 1990], pages 51–69.

[Pollard and Sag 1987] Carl Pollard and Ivan A. Sag. *Information-Based Syntax and Semantics. Volume 1: Fundamentals.* CSLI Lecture Notes # 13. Stanford, California (Center for the Study of Language and Information), 1987. (Distributed by University of Chicago Press).

[Pollard and Sag 1988] Carl Pollard and Ivan A. Sag. An Information-Based Theory of Agreement. CSLI Report # 132, Center for the Study of Language and Information (Stanford, California), 1988.

[Pollard and Sag 1994] Carl Pollard and Ivan A. Sag. *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. Stanford, California (The Univeristy of Chicago Press), 1994.

[Shieber *et al.* 1983] Stuart Shieber, Hans Uszkoreit, Fernando Pereira, Jane Robinson, and Mary Tyson. The Formalism and Implementation of PATR-II. In: [Grosz and Stickel 1983], pages 39–79.

[Shieber 1986] Stuart M. Shieber. *An Introduction to Unification-Based Approaches to Grammar*. CSLI Lecture Notes # 4. Stanford, California (Center for the Study of Language and Information), 1986.

[Uszkoreit 1986] Hans Uszkoreit. Categorial Unification Grammars. In: [COLING 1986].

[Vater 1985] Heinz Vater. Einführung in die Nominalphrasensyntax des Deutschen. KLAGE (Kölner Linguistische Arbeiten Germanistik) # 10, Universität zu Köln (Köln), 1985.

[Vater 1986] Heinz Vater (editor). *Zur Syntax der Determinantien*. Studien zur deutschen Grammatik # 31. Tübingen (Gunter Narr Verlag), 1986.

[Wunderlich 1988] Dieter Wunderlich. Einige Bemerkungen über Kongruenz. Unpublished Manuscript, August 1988.

[Zwicky 1985] Arnold M. Zwicky. Heads. *Journal of Linguistics*, 21 : 1–29, 1985.

[Zwicky 1986] Arnold M. Zwicky. German Adjective Agreement in GPSG. *Linguistics*, 24 : 957–990, 1986.