Elementary Principles of HPSG*

Georgia M. Green

January 27, 2000

Introduction

This chapter describes the theoretical foundations and descriptive mechanisms of Head-Driven Phrase Structure Grammar (HPSG), as well as proposed treatments for a number of familiar grammatical phenomena. The anticipated reader has some familiarity with syntactic phenomena and the function of a theory of syntax, but not necessarily any expertise with modern theories of phrase-structure grammar. The goal of this chapter is not so much to provide a tutorial in some consistent (and inevitably dated) version of HPSG as it is to explicate the philosophy and techniques of HPSG grammars. In my opinion, the best way to fully understand this approach, to be able to write and read HPSG grammars, is to build an HPSG grammar from scratch, inventing and revising the details as one goes along, in accordance with the constraints imposed by the formal model (but not necessarily by every constraint stipulated in the language of that model).

Section 1 of this chapter describes the character of HPSG grammars, and the elements and axioms of the system. Section 2 describes how linguistic entities are modelled, and how grammars describe the modelled entities. The third section describes the ontology of feature-structure descriptions in HPSG, and Section 4 deals with the expression of constraints, especially those involving the notion 'same' or 'matching'. Section 5 describes the compositional treatment of semantics in HPSG. Section 6 discusses the representation of constituent structure, and Section 7 addresses the treatment of the order of elements within constituents. HPSG is very much a lexicon-driven theory, and Section 8 describes the organization of the lexicon, relations among lexical items, and the nature of lexical rules relating them. Section 9 describes treatments of complementation, including the treatment of Equi and Raising constructions, and their interaction with expletive noun phrases. Section 10 describes variations on the treatment of so-called extraction constructions and other unbounded dependencies (e.g., pied piping), with some attention to multiple extractions and so-called parasitic gaps, as well as the nature of alleged empty categories like traces and zero pronouns. It concludes with a discussion of constraints on where extraction gaps can occur. The last section describes the HPSG account of the binding of pronouns and anaphors. Two appendices summarize salient aspects of the sort inheritance hierarchies discussed, and the constraints embedded within them.

1 Grammars, types, constraints

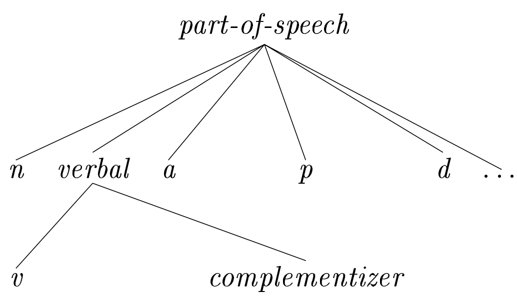
Two assumptions underlie the theory defining head-driven phrase structure grammars. The first is that languages are SYSTEMS of sorts of linguistic objects at a variety of levels of abstraction, not just

*This work was supported in part by the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. Some parts are reworked versions of material that appears in Green & Levine 1999. I am grateful to Ash Asudeh, Bob Borsley, Bob Levine, Carl Pollard and Ivan Sag for comments on earlier versions, and useful advice about consistency and clarity in describing a theory that (like all theories, and maybe all organisms) evolves piecemeal, a few systems at a time.

collections of sentence(-type)s. Thus, the goal of the theory is to be able to define the grammars (or I-languages) that generate the sets of sentences (E-languages) that represent the set of natural human languages, assigning empirically satisfactory structural descriptions and semantic interpretations, in a way that is responsible to what is known about human sentence processing.[1] The other is that grammars are best represented as process-neutral systems of DECLARATIVE CONSTRAINTS (as opposed to constraints defined in terms of operations on objects, as in transformational grammar). Thus, a grammar (and for that matter, a theory of universal grammar) is seen as consisting of an inheritance hierarchy[2] of sorts (an is-a hierarchy), with constraints of various kinds on the sorts of linguistic object in the hierarchy.

A simple sort hierarchy can be represented as a taxonomic tree representing the sort to which belong all the linguistic entities with which the grammar deals. For each local tree in the hierarchy, the sort names which label the daughter nodes partition the sort which labels the mother; that is, they are necessarily disjoint subsorts which exhaust the sort of the mother. For example, subsorts of the sort *head* can be 'parts of speech' (not words!) of various kinds, and some of those sorts are further partitioned, as illustrated in (1).

(1)



A partial inheritance hierarchy for 'parts of speech'

A multiple-inheritance hierarchy is an interlocking set of simple hierarchies, each representing a dimension of analysis that intersects with other dimensions. This is particularly striking in the lexicon: verbs are usefully classified by the number and syntactic characteristics of the arguments they require, but they may also need to be classified according to inflectional class (conjugation), and by semantic properties of the relations they describe (e.g., whether they represent states or properties or events; whether their subjects represent agents or experiencers, and so on (Green 1974, Levin 1993).

A grammar is thus a system of constraints, both unique and inherited, on sorts of linguistic objects. It would be naive to assume that all grammars have the same sorts or the same constraints on whatever sorts they might have in common. Nevertheless, all grammars are hierarchies of sorts of phrases and words and the abstract linguistic entities that need to be invoked to define them. All grammars constrain these various sorts in terms of properties of their component parts. One may speculate that grammars are as alike as they are as a result of there being only a small number of economical solutions to the problems posed by competing forces present in languages generally. For example, languages with free word order enable subtle (non-truthconditional) distinctions to be expressed by variation in phrase order, while languages with fixed word order simplify the task of parsing by limiting the possibilities for subsequent phrases. An elaborate inflectional system reduces ambiguity (especially temporary ambiguity), while relatively uninflected languages simplify

[1]See Chomsky (1986).
[2]More exactly, it is a multiple-inheritance hierarchy; see Section 2.

the choices that have to be made in speech production. At the same time, whatever psychological properties and processes guide the incremental learning about the world that is universal among human beings in their first years of life must contribute to constraining grammars to be systems of information that can be learned incrementally.[3]

Sorts can be atomic (unanalyzed) like *masc, fem, +*, and *sg*, or they can be complex. Complex sorts of linguistic objects are defined in terms of the attributes they have (represented as features), and by the value-types of those features. In HPSG, a feature's value may be defined to be one of four possible types:

- an atomic sort (like +, or *finite*)
- a feature structure of a particular sort
- a set of feature structures[4]
- list of feature structures[5]

If a value is not specified in a feature-structure description, the value is still constrained by the sort-declarations to be one of the possible values for that feature. That is, it amounts to specifying a disjunction of the possible values. Thus, if the possible values for the feature NUM are the atomic sorts *sg* and *pl*, then specifying either NP[NUM] or NP amounts to specifying NP[NUM $sg \vee pl$], and similarly for all the other possible attributes of NPs (i.e., all the features they can have).

Sort declarations are expressed in formulae of a logic for linguistic representations (King 1989, Pollard 1999), and can be perspicuously abbreviated in labelled attribute-value matrices (AVMs) as in (2), where F1, ..., Fn are feature names and $sort_i$, ..., $sort_k$ are sort names.

(2) $\begin{bmatrix} \textit{sort0} & \\ \text{F1} & \textit{sort}_i \\ \text{F2} & \textit{sort}_j \\ \ldots & \end{bmatrix}$

Sort definitions thus specify what attributes an instance of the sort has, and what kinds of things the values of those attributes can be, and sometimes what particular value an attribute must have (either absolutely, or relative to the value of some other attribute[6]). Sorts inherit all of the attributes of their supersorts and all of the restrictions on the values of those attributes. The set of feature-structures defined by a grammar is a partial subsumption ordering, i.e., a transitive, reflexive, and anti-symmetric relation on the subsumption relation. Thus, linguistic expressions, or *signs*, are words or phrases, and this is reflected as the fact that the sort *sign* subsumes both *phrase* and *word* and no other sort. In fact, since the specifications for *phrase* and *word* are mutually exclusive (phrases have attributes which specify their immediate constituents, and words don't), the sorts *phrase* and *word* PARTITION the sort *sign*. Sorts which have no subsorts are termed 'maximal sorts' because they are maximally informative or specific.

[3]See Green 1997, in prep., for some discussion.

[4]Set values are represented as sequences within curly brackets: SLASH $\{\boxed{1}, \boxed{2}\}$. The empty set is denoted: { }, while {[ ]} denotes a singleton set.

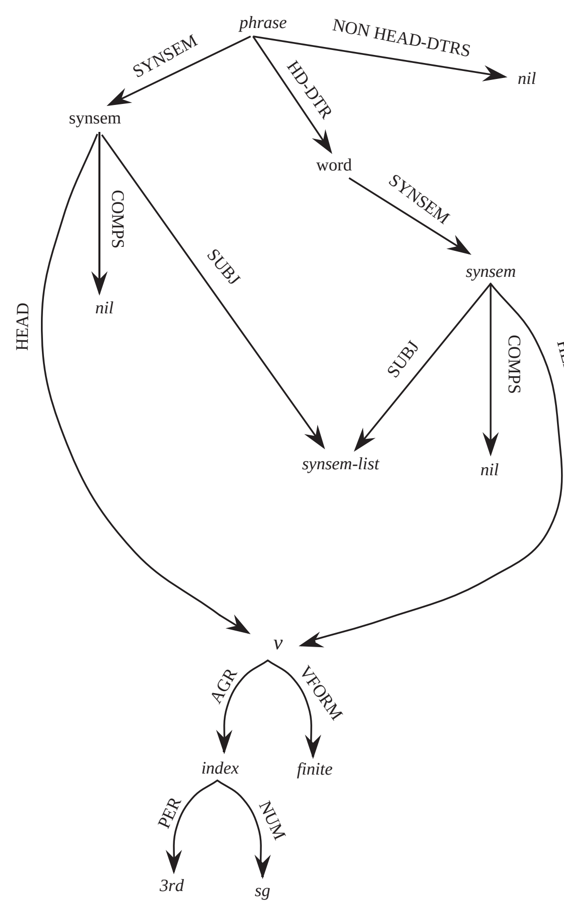
[5]List values are represented as sequences within angled brackets: SUBCAT ⟨ NP, VP[*inf*] ⟩. The empty list is denoted: ⟨ ⟩, while ⟨[ ]⟩ denotes a singleton list.

[6]See Section 4 below.

2 Feature structures, feature structure descriptions

All linguistic entities (including both expression types, and the abstract objects that are invoked to describe them) are modelled in HPSG as **feature structures**.[7] A feature structure is a complete specification of all the properties of the object it models.

To keep the distinction clear between a feature structure, which models a maximal sort, and the feature structure descriptions which are used to describe grammatically–relevant classes of feature structures in the generalizations that constitute the statements of the grammar, feature structures themselves are represented as directed graphs. A feature structure for a simplified account of the English verb phrase *sleeps* is given below:



(3)

The feature structure in (3) reflects the following information: the phrase in question has "internal" syntactic and semantic properties represented by the feature SYNSEM, as well as the property of having a head daughter (HEAD-DTR) but no other subconstituents; its NH-DTRS (non-head daughters) attribute has the value *nil*. Its "part of speech" (HEAD) value is of subsort *verb*, of finite inflectional form, and required to agree with something whose agreement (AGR) value is 3rd person and singular, and its head daughter's part of speech (HEAD) value is exactly the same. In addition, the phrase subcategorizes for (i.e., requires) a subject, but no complements, and the phrase it sub-

[7]For discussion see Pollard and Sag 1994: 8, 17–18, Pollard 1999, and for background, Shieber 1986, Pollard and Moshier 1990.

categorizes for is precisely the phrase its head daughter subcategorizes for.

As is clear from this example, the directed graphs that represent feature structures differ from the directed graphs conventionally used to represent constituent structure diagrams, in that distinct nodes can be the source for paths to (i.e. to "dominate") a single node. This situation, as indicated by the convergence of the arrows in (3), represents the fact that the *part-of-speech* (of subtype *v*) of the head daughter the same feature-structure as the *part-of-speech* of the phrase itself.

Graph representations of feature structures are awkward both to display and to read, so descriptions of feature structures in the form of attribute value matrices (AVMs) are commonly used instead. Attribute or feature names are typically written in upper case in AVMs, and values are written to the right of the feature name, in lower case italics if they are atomic, as in (4).

(4) $\begin{bmatrix} \text{PER } \textit{3rd} \\ \text{NUM } \textit{sg} \\ \text{GEN } \textit{fem} \end{bmatrix}$

Feature structures are the entities constrained by the grammar. It is crucially important to distinguish between feature structures (fully specified objects that model linguistic expressions) and **feature structure descriptions**, representations (usually underspecified) that (partially) describe feature structures, and which feature structures allowed by a grammar must **satisfy**. Feature structure descriptions characterize classes of objects. For example, The NP *she* could be represented by a fully specified feature structure (representable as a directed graph), but "NP" is (an abbreviation for) a feature structure description, and could not be so represented. Put another way, a partial description such as a feature structure description represented by an AVM is a constraint on members of a class of feature structures, while a total description is a constraint which limits the class to a single member. For the most part, grammar specification deals with generalizations over classes of words and phrases, and therefore with (partial) feature structure descriptions.

3 Signs and their attributes

Head-driven phrase structure grammars describe languages in terms of the constraints on linguistic expressions (*signs*) of various types. Signs are, as in the Saussurean model, associations of form and meaning, and have two basic subsorts: phrases, which have immediate constituents; and words, which don't. Signs are abstractions, of course; an act of uttering a linguistic expression that is modelled by a particular sign amounts to intentionally producing a sound, gesture, or graphical object that satisfies the phonological constraints on that sign, with the intent that the product of that act be understood as intended to have syntactic, semantic, and contextual properties that are modelled by the respective attributes of that sign.[8]

Signs have phonological, syntactico–semantic, and contextual properties, each represented by the value of a corresponding feature. Thus all signs have PHON and SYNSEM attributes, recording their phonological and syntactico-semantic structures, respectively.[9] The value of the SYNSEM attribute is a feature structure which represents the constellation of properties that can be grammatically selected for. It has a LOC(AL) attribute, whose value (of type *local*) has CAT(EGORY), CONT(ENT), and CONTEXT attributes. *Local* values are what is shared by filler and gap in so-called extraction constructions. The SYNSEM value also has a NONLOC(AL) attribute, which in effect encodes information about all types of unbounded dependency constructions (UDCs).

[8]For more on the nature of this modelling, see Pollard & Sag 1994: 6-10, 58.

[9]PHON values are usually represented in standard orthography, solely for the sake of convenience and readability.

The CATEGORY attribute takes as its value an entity of the sort *category*, whose attribute HEAD has a *part-of-speech* as its value and whose SUBJ, COMPS, and SPR attributes have lists of synsems as their values. These type declarations, and others discussed in this chapter, are summarized in Appendix A.

Within words, categories also have an argument structure (ARG-ST) feature whose value is a list of the synsems which denote the sign's arguments. They are ordered by the obliqueness of the grammatical relations they represent,[10] and the ARG-ST list represents the obliqueness record that is invoked in constraining binding relations (cf. Pollard & Sag 1994: Ch. 6, or Sag & Wasow 1999: Ch. 7).[11]

The valence attributes of a sign (SUBJ, SPR, COMPS) record what the subcategorization requirements of the sign are. These attributes take lists of synsems as their values; in S's and referential NPs, all the lists are empty lists. In most cases, the ARG-ST list is the concatenation of the contents of the subject, specifier and complements lists, in that order. The exceptions are that so-called null pronouns and the *gap-synsems*[12] representing "extracted" elements are on ARG-ST lists but not on valence lists.

In Pollard & Sag 1994, the value of the CONTENT attribute is a *nominal-object* if the sign is a referring expression, but a parameterized state-of-affairs (or *psoa*) if it is a predicative expression, or a quantifier. *Psoas* are representations of propositional content as feature-structures. They are sub-typed by the relation they express, and have attributes for the roles of their arguments, while *nominal-objects* have index-valued INDEX attributes and *psoa*-set valued RESTR(ICTION) attributes. As illustrated in (5),[13] more current versions of the theory[14] include an INDEX and a RESTR attribute for predicative as well as referring expressions.

(5) a. $\begin{bmatrix} \text{INDEX} & \textit{index} \\ \text{RESTR} & \left\{ \begin{bmatrix} \textit{nominate-rel} & \\ \text{NOMINATOR} & \textit{index} \\ \text{NOMINEE} & \textit{index} \end{bmatrix} \right\} \end{bmatrix}$

[10]Arguments are ordered from least oblique to most oblique on the ranking familiar since Keenan & Comrie (1977): subject > direct object > secondary object > oblique argument.

[11]N.b.: in certain other respects the binding theory presented in Sag & Wasow reflects its character as an introductory textbook, and does not correspond to the binding theory in Pollard & Sag 1994.

[12]Gap-synsems and PRO-synsems, which describe "extracted" elements and implicit Equi (or PRO) subjects, respectively, are never the SYNSEM value of a syntactic constituent. Syntactic constituents are signs, and signs are constrained to have SYNSEM values of the subsort *canonical*, with a nonempty value for PHON. Null pronouns also have a non-canonical synsem type, and thus can be present in ARG-ST lists, but can never satisfy a SUBJ or COMPS requirement, since only signs can do that.

[13]This representation of propositional content does not reflect an essential property of HPSG. It would make no difference if some other kind of coherent representation of a semantic analysis was substituted, as long as it provided a way of indicating what properties can be predicated of which arguments, how arguments are linked to individuals in a model of the universe of discourse, and how the meaning of each constituent is a function of the meaning of its parts. In other words, the exact form of the representation is not crucial as long as it provides a compositional semantics.

[14]For example, Copestake et al., To appear, Sag & Wasow 1999, Ginzburg & Sag To appear.

$$\text{b.}\ \left[\begin{array}{ll} \text{INDEX} & \textit{index} \\ \text{RESTR} & \left\{ \left[\begin{array}{ll} \textit{persuade-rel} & \\ \text{PERSUADER} & \textit{index} \\ \text{PERSUADEE} & \textit{index} \\ \text{STATE-OF-AFFAIRS} & \textit{psoa} \end{array}\right] \right\} \end{array}\right]$$

$$\text{c.}\ \left[\begin{array}{ll} \text{INDEX} & \boxed{1} \\ \text{RESTR} & \left\{ \left[\begin{array}{ll} \textit{coconut-rel} & \\ \text{SITUATION} & \textit{index} \\ \text{INSTANCE} & \boxed{1} \end{array}\right] \right\} \end{array}\right]$$

Indices for expressions that denote individuals are of the subsort *indiv-ind*, while indices for expressions that denote properties or propositions (situations) are of the sort *sit-ind*.[15] Indices for nominal entities in turn have attributes for PER(SON), NUM(BER), and GEN(DER). For perspicuity, in abbreviated AVMs, index values are often represented as subscripts on category designations: NP_i, for example. The CONTENT specification is abbreviated as a tag following a colon after a category designation; VP:$\boxed{i}$ represents a VP with the CONTENT value $\boxed{i}$.

Finally, the CONTEXT attribute records indexical information (in the values of the SPEAKER, ADDRESSEE, and LOCATION features), and is supposed to represent, in the value of its BACKGROUND attribute, linguistically relevant information that is generally considered pragmatic. For some discussion, see Green (1995).

4 Constraints, Structure-sharing, and Subcategorization

As indicated in Section 2, constraints on feature structures are expressed in terms of feature-structure descriptions. The more underspecified a description is, the larger the class of objects that satisfy it, and the greater the generalization it expresses. Anything that is entailed in sort definitions (including lexical representations) or in universal or language-specific constraints[16] does not have to be explicitly mentioned in the constraints on (i.e., descriptions of) classes of linguistic objects. For example, since the Head Feature Principle requires that the HEAD value of the head-daughter of a phrase be the same as the HEAD value of the phrase itself, the details of this value only need to be indicated once in each representation of a phrase.

The notion of the values of two attributes being the same is modelled in feature structures as the sharing of structure. This is represented in feature structures by means of distinct paths of arcs terminating at the same node as in (3),[17] and in descriptions by means of identical boxed integers (TAGS like $\boxed{3}$) prefixed to feature-structure descriptions, denoting that they are constrained to describe the same structure, as illustrated above in (5c).[18] STRUCTURE–SHARING is a crucial property of HPSG. Because it refers to token-identity, and not just type-matching, it does not have a direct counterpart in transformational theories. Structure-sharing amounts to the claim that the value of some instance of an attribute is the same feature-structure as the value of some other

[15] See Appendix A for more details.

[16] These may be expressed as sort definitions for higher-level sorts. Sag (1997) is an example of this approach.

[17] This property is sometimes referred to as re-entrancy for this reason.

[18] Technically, a tag refers to feature structures described by the unification of all of the feature structure descriptions o with the same tag. The unification of two feature structures descriptions is a consistent feature structure description that contains all of the information in each one.

instance of an attribute, i.e., it is the SAME THING—not something that just shares significant properties with it, or a different thing which happens to have all the same properties.

Thus, the following three AVMs are equivalent descriptions of the same feature structure (a representation of a third person singular noun phrase consisting of a single head noun).[19]

(6) a.
$$\begin{bmatrix} \textsc{synsem}|\textsc{content}|\textsc{index}\ \boxed{1}\begin{bmatrix}\textsc{per}\ \textit{3rd}\\ \textsc{num}\ \textit{sg}\\ \textsc{gen}\ \textit{fem}\end{bmatrix}\\ \textsc{head–dtr}\ |\textsc{synsem}|\textsc{cont}|\textsc{index}\ \boxed{1}\\ \textsc{non-hd–dtrs}\ \langle\rangle \end{bmatrix}$$

b.
$$\begin{bmatrix} \textsc{synsem}\ |\textsc{content}|\textsc{index}\ \boxed{1}\\ \textsc{head–dtr}|\textsc{synsem}|\textsc{cont}|\textsc{index}\ \boxed{1}\begin{bmatrix}\textsc{per}\ \textit{3rd}\\ \textsc{num}\ \textit{sg}\\ \textsc{gen}\ \textit{fem}\end{bmatrix}\\ \textsc{non-hd–dtrs}\ \langle\rangle \end{bmatrix}$$

c.
$$\begin{bmatrix} \textsc{synsem}|\textsc{content}|\textsc{index}\ \boxed{1}\begin{bmatrix}\textsc{num}\ \textit{sg}\end{bmatrix}\\ \textsc{head–dtr}|\textsc{synsem}|\textsc{cont}|\textsc{index}\ \boxed{1}\begin{bmatrix}\textsc{per}\ \textit{3rd}\\ \textsc{gen}\ \textit{fem}\end{bmatrix}\\ \textsc{non-hd–dtrs}\ \langle\ \ \rangle \end{bmatrix}$$

All three descriptions convey the same information, since there is only one way to satisfy the token–identities in the three descriptions.

Structure-sharing is a key descriptive mechanism in HPSG. For example, the structure–sharing required by the description of a topicalization structure, given in (7), requires the SYNSEM|LOCAL value of the filler daughter to be the same as the single member of the SLASH value of the head daughter, as explained in Section 10.

(7)
$$\begin{bmatrix} \textit{hd-filler-ph}\\ \textsc{head–dtr}\begin{bmatrix}\textit{phrase}\\ \textsc{synsem}\begin{bmatrix}\textsc{local}\ |\ \textsc{cat}\begin{bmatrix}\textsc{head} & \textit{verb}\\ \textsc{subj} & \langle\ \ \rangle\\ \textsc{comps} & \langle\ \ \rangle\end{bmatrix}\\ \textsc{non–local}\ |\ \textsc{slash}\ \left\{\boxed{1}\right\}\end{bmatrix}\end{bmatrix}\\ \textsc{non-hd–dtrs}\left\langle\begin{bmatrix}\textit{phrase}\\ \textsc{synsem}\begin{bmatrix}\textsc{local}\ \boxed{1}\\ \textsc{nonlocal}\ |\ \textsc{slash}\end{bmatrix}\end{bmatrix}\right\rangle \end{bmatrix}$$

[19]In the following AVMs, sort annotations are omitted, and feature–name pathways like [A [B [C x]]] are represented as A|B|C x, as is conventional. For perspicuity, sometimes values are labelled with the name (in *italics*) of the sort that structures their content, but such information is usually omitted when predictable. The attributes HEAD-DTR and NON-HEAD-DTRS organize information about the constituent structure of phrases. Their properties are described in subsequent paragraphs and elaborated in Section 6.

There are a variety of restrictions that generalize across various subtypes of phrases. These are, in general, constraints on the highest type they apply to in the phrase-type hierarchy. Several depend on the notion of structure-sharing to constrain feature-value correspondences between sisters, or between mother and some daughter, for particular features. These include familiar principles like the HEAD-Feature Principle (which constrains the HEAD value of a phrase to be the same as the HEAD value of its head-daughter) and the Valence Principle, as well as a principle (e.g., the Nonlocal Feature Principle of Pollard & Sag 1994) which governs the projection of the unbounded dependency features (SLASH, REL, and QUE). Principles which constrain the CONTENT value of a phrase to have a certain relation to the CONTENT values of its daughters, depending on what subtype of phrase it is, are specified in the sort declarations for particular subsorts of *phrase*.

An example of a principle which is represented as part of the description of a subsort of *phrase* is the Valence Principle.[20] It constrains subcategorization relations of every object of the sort *headed-phrase* so that the value of each valence feature corresponds to the respective valence value of its head daughter, minus elements which correspond to elements with the same SYNSEM values in the NON-HD–DTRS list for that phrase. In other words, the Valence Principle says that the SUBJ, COMPS and SPR values of a phrase correspond to the respective SUBJ, COMPS and SPR values of its head daughter except that the synsems on those lists that correspond to phrases constituting any non-head daughters are absent from the valence attributes of the mother. In versions of HPSG with defaults,[21] the Valence Principle can be formulated as a constraint on headed phrases to the effect that the values of the valence features of a phrase are the same as those of its head daughter, except where specified to be different, as in (8).[22]

(8) $$\begin{bmatrix} \textit{headed-phrase} & \\ \text{SYNSEM}\,|\,\text{LOCAL}\,|\,\text{CAT} & \begin{bmatrix} \text{SUBJ} & /\,\boxed{1} \\ \text{SPR} & /\,\boxed{2} \\ \text{COMPS} & /\,\boxed{3} \end{bmatrix} \\ \text{HD–DTR}\,|\,\text{SYNSEM}\,|\,\text{LOCAL}\,|\,\text{CAT} & \begin{bmatrix} \text{SUBJ} & /\,\boxed{1} \\ \text{SPR} & /\,\boxed{2} \\ \text{COMPS} & /\,\boxed{3} \end{bmatrix} \end{bmatrix}$$

Valence features of the phrase would only be specified to be different in sort declarations for particular headed-phrase types where the synsems of the signs that are sisters to the head are absent from the appropriate valence feature on the phrase, as discussed in Section 6.

Other aspects of selection also rely on the notion of structure-sharing. Adjuncts select heads via a HEAD feature MOD, and determiners (and markers) select heads via a HEAD feature SPEC in very much the same way that heads select arguments by valence features. Structure-sharing is the essence of the HEAD-Feature Principle. The HFP is described in Pollard & Sag 1994 as an independent

[20] A reformulation of the Subcategorization Principle of Pollard & Sag (1994).

[21] I.e., in nonmonotonic formulations such as Sag 1997 or Ginzburg & Sag (To appear).

[22] The right-leaning slash in a default specification has the interpretation 'unless otherwise specified, has the following value.' The logic of default inheritance and the notation are described in Lascarides & Copestake (1999). If VALENCE is an attribute of categories and its *valence* value is a feature-structure with SUBJ, COMPS and SPR values as described, then the Valence Principle can be very simply represented as:

$$\begin{bmatrix} \textit{headed-phrase} & \\ \text{SYNSEM}\,|\,\text{LOCAL}\,|\,\text{CAT}\,|\,\text{VALENCE} & /\,\boxed{1} \\ \text{HD–DTR}\,|\,\text{SYNSEM}\,|\,\text{LOCAL}\,|\,\text{CAT}\,|\,\text{VALENCE} & /\,\boxed{1} \end{bmatrix}$$

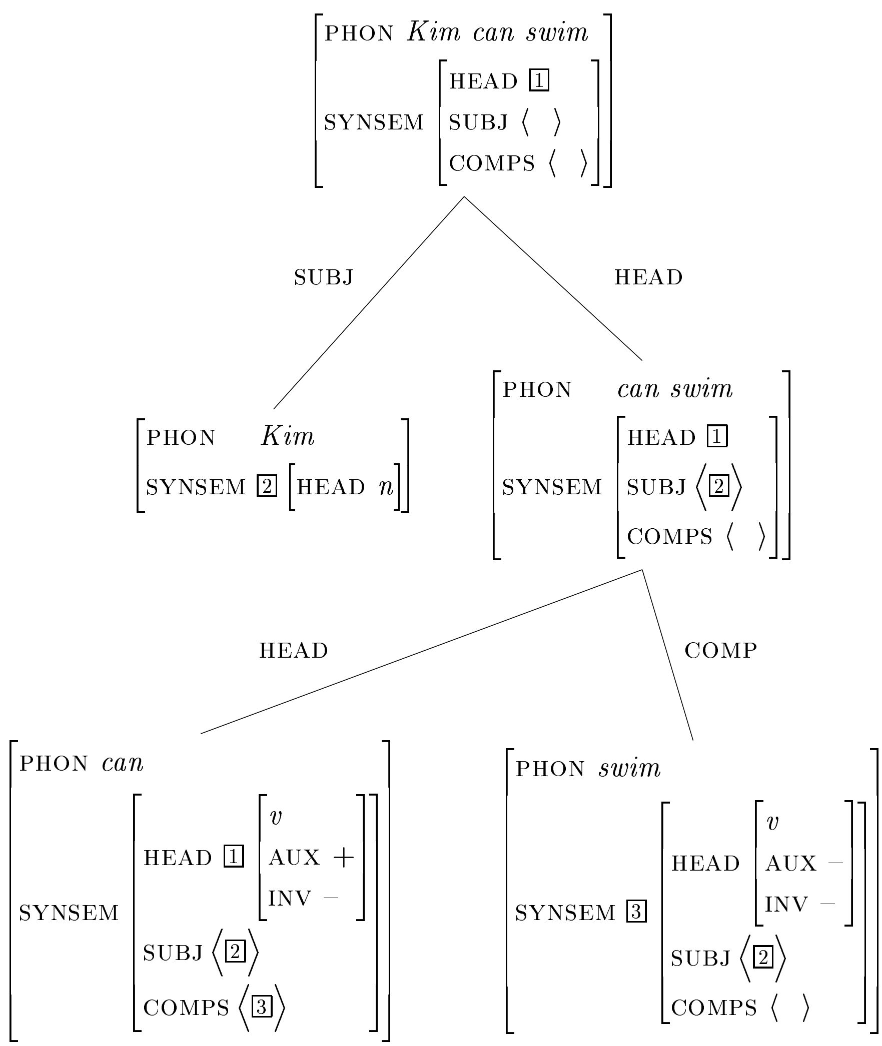
constraint, but is perhaps more perspicuously represented as part of the sort declaration for the sort *headed-phrase*. as shown in (9).

(9) $$\begin{bmatrix} \textit{headed-phrase} \\ \text{SYNSEM} \mid \text{LOCAL} \mid \text{CATEGORY} \mid \text{HEAD } \boxed{1} \\ \text{HEAD–DTR} \begin{bmatrix} \text{SYNSEM} \mid \text{LOCAL} \mid \text{CATEGORY} \mid \text{HEAD } \boxed{1} \end{bmatrix} \end{bmatrix}$$

HPSG licenses phrase types either through Immediate Dominance Schemata (Pollard & Sag 1994), or through definitions of particular sorts of phrasal constructions (Sag 1997). It does not treat constituent-structure trees as formal objects, although constituent structure is represented by the various daughters attributes of phrasal signs, and trees are used as a convenient graphic representation of the immediate constituents and linear order properties of phrasal signs. In informal arboreal representations, nodes are labelled by analyzable category names in the form of AVMs,[23] linear order is imposed, and branches may be annotated to indicate the relation (e.g., head, adjunct, complement) of daughter to mother. The AVMs are usually abbreviated, with predictable parts of paths suppressed as in (10).

[23] Or very underspecified abbreviations for them, like NP or NP[3SG].

(10) *Kim can swim.*



5 Semantics

As mentioned in Section 3, every sign has a CONTENT feature which represents its semantics, and as HPSG has been committed from the start to providing an account of how the meaning of each phrase is a function of the meanings of its parts (compositionality), principles defining this function have been a part of every version of HPSG. Naturally, what this function is depends on what the representations of the semantics of the different types of phrases look like.

In P&S (1994), the sort that is the value of the CONTENT feature varies according to whether the sign's HEAD value is of type *noun, quantifier*, or something else. For most parts of speech, the CONTENT value is a *parameterized-state-of-affairs (psoa)*, a structure representing a particular kind of relation, with attributes for each of the roles defined for it. The values of those attributes are indexes or psoas that must be structure-shared with particular items on the ARG-ST list, as in the abbreviated representation in (11).

(11)
$$\begin{bmatrix} \textit{word} & \\ \text{HEAD } \textit{v} & \\ \text{SUBJ } \left\langle \boxed{3} \right\rangle & \\ \text{COMPS } \left\langle \boxed{4} \right\rangle & \\ \text{ARG-ST } \left\langle \boxed{3}\text{NP}\left[\text{CONT}|\text{INDEX } \boxed{1}\right], \boxed{4}\text{NP}\left[\text{CONT}|\text{INDEX } \boxed{2}\right] \right\rangle & \\ \text{CONT}|\text{RESTR } \begin{bmatrix} \textit{choose-relation} \\ \text{AGENT } \boxed{1} \\ \text{UNDERGOER } \boxed{2} \end{bmatrix} & \end{bmatrix}$$

In (11) *choose-relation* is the name of a subsort of *psoa* that represents the relation that the verb *choose* refers to. The immense and elaborate cross-classifying taxonomy of psoa-sorts is in effect a representation of encyclopedic (world) knowledge, as seen through the eyes of a language.[24] As illustrated in (12) the value of CONTENT for a [HEAD *noun*] sign is a feature-structure of sort *nominal-object*.

(12)
$$\begin{bmatrix} \textit{sign} & \\ \text{HEAD} & \begin{bmatrix} \text{CAT}\,|\,\text{HEAD } \textit{n} & \\ \text{CONT} & \begin{bmatrix} \textit{nom-object} & \\ \text{INDEX} & \textit{index} \\ \text{RESTR} & \textit{set(psoa)} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

As mentioned, indexes have PERSON, NUMBER and GENDER attributes. NUMBER and GENDER values may be a function of the object in the world that is referenced by the utterance of the nominal expression (the object it is **anchored to**), or reflect an arbitrary property of the word, or both, depending on the language. Nominal-objects also have a RESTRICTION attribute, whose value is a set of psoas restricting the referent of the nominal-object to have certain properties.[25]

Quantifiers have a third kind of feature structure as their CONTENT value. As illustrated in (13), they have an attribute DET, whose value is an object of sort *semantic-determiner*, and an attribute RESTIND (for Restricted Index), which is a nominal-object of the subsort *nonpronoun*, whose INDEX value is always a referential index (as opposed to an expletive index).

(13)
$$\begin{bmatrix} \textit{word} & \\ \text{HEAD} & \begin{bmatrix} \text{CAT} & \textit{det} \\ \text{CONT} & \begin{bmatrix} \textit{quant} & \\ \text{DET } \textit{semdet} & \\ \text{RESTIND} & \begin{bmatrix} \textit{npro} & \\ \text{INDEX} & \textit{ref} \\ \text{RESTR} & \textit{set(psoa)} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

[24] See Davis 1995 for a cogent discussion of the complexity of the factors entering into the classification, and an illustration of how their interaction can be represented to reflect the correct generalizations about how semantic roles are linked to syntactic argument indexes. One major achievement of this work is the multiple-inheritance hierarchy of relations according to entailments and role-types (reminiscent of Dowty's (1991) proto-roles, but reified).

[25] For some discussion, see Green 1995.

The first formulation of the compositionality principle that constrains the CONTENT of a phrase in terms of the CONTENT values of the phrase's immediate constituents is relatively simple: the CONTENT of a phrase is structure-shared with the CONTENT of its head-daughter. Because of the nature of psoa repesentations for all predicative expression types (phrases headed by verbs, and predicative prepositions, adjectives, and nouns), this works fine for phrases consisting of a head and complements, or a head and a subject. It doesn't work at all for phrases consisting of a head and a modifier, like *[[eats peas] slowly]*. This necessitates a second, disjunctive formulation which adds the condition that if the non-head daughter is an adjunct, then the CONTENT of the phrase is structure-shared with the CONTENT of the adjunct daughter. This appears to give exactly the right results for modifiers of VPs, as long as those modifiers are analyzed like functions which take their heads as arguments, as is familiar from many semantic traditions. However, it necessitates a somewhat strained analysis of attributive adjectives.

Because the semantics of a head-adjunct phrase is the same as that of the adjunct, the CONTENT value of attributive adjectives has to have the same semantic type as nouns and NPs have (nominal objects). That means they have an INDEX attribute (whose value has to be stipulated to be the same as that of the nominal head they modify), and a RESTRICTION attribute (whose value has to be stipulated to be whatever properties are contributed by the adjective, unioned with whatever properties are contributed by the nominal expression that is modified).

In any case, that clause of the P&S (1994) Semantics Principle has been convincingly shown to make incorrect predictions (Kasper 1995). The clause defining the contribution of quantifiers and attempting to account for the ambiguity of scope is so complex (P&S: 323) that the informal version of it requires four to six clauses to state, and I will not try to summarize it.

At the same time as the empirical problems involved in adjunct semantics were being brought to light, work on computational implementations showed a need for semantic representations that minimized recursion. One motivation for minimizing recursion relates to the fact that in many of the natural language processing applications which utilize implementations of HPSG, it is unnecessary to resolve ambiguities of quantifier scope, as in (14).

(14) **A** \$3000 investment is enough to become a shareholder in **thousands** of mutual funds.

Minimal Recursion Semantics (referred to as MRS—see Copestake, Flickinger, Pollard and Sag (To appear)) enables semantic representations which are underspecified with respect to quantifier scope.

Another motivation relates to the fact that computing with semantic representations with unrestricted recursion consumes inordinate quantities of computational resources. As a consequence of minimizing recursion, the CONTENT values in Sag & Wasow (1999) are more uniform, and allow a simpler compositionality statement, though at the cost of additional typed indices.

In addition, the factoring entailed by the minimal recursion approach to semantics enables a feature geometry which enforces the Locality Constraint (that only immediate complements can be selected, not arguments embedded within complements); the list of semantic objects involved in a representation just has to be a sign-level attribute, rather than an attribute within CONTENT.

In MRS-style analyses, CONTENT values have three attributes: MODE, INDEX and RESTRICTION.[26] The possible values of MODE are the atomic modes *proposition, directive, interrogative*, and *reference.* References have INDEX values which are indices to either entities (in the case that the expression is a referring expression) or situations (in the case that the expression is predicative (e.g., a verbal, adjectival, or predicative prepositional or nominal expression) like the italicized expressions in (15)).

[26]Sag & Wasow (1999) do not attempt an analysis of quantifiers.

(15) a. Kim *laughed.*

b. Kim is *funny.*

c. Kim is *under the table.*

d. Kim is a *pediatrician.*

Proposition-, directive-, and interrogative-valued *content*s always have a situation-valued index. RESTRICTION values are sets of typed *predications*, similar in structure to *psoas* in P&S (1994), except that each one has a SITUATION attribute with a *sit-ind* value.[27] To illustrate, *Kim, pediatrician* and *Kim is a pediatrician* in (15d) would have the CONTENT values in (16a, b, c), respectively.

(16) a.
$$\begin{bmatrix} \text{MODE} & \textit{ref} \\ \text{INDEX} & \boxed{1}\textit{indiv-ind} \\ \text{RESTR} & \left\{ \boxed{4} \begin{bmatrix} \textit{called} \\ \text{SIT } \boxed{3} \\ \text{ENTITY } \boxed{1} \\ \text{NAME } \textit{Kim} \end{bmatrix} \right\} \end{bmatrix}$$

b.
$$\begin{bmatrix} \text{MODE} & \textit{prop} \\ \text{INDEX} & \boxed{2} \\ \text{RESTR} & \left\{ \boxed{5} \begin{bmatrix} \textit{pediatrician} \\ \text{SIT } \boxed{2} \\ \text{INSTANCE } \boxed{1} \end{bmatrix} \right\} \end{bmatrix}$$

c.
$$\begin{bmatrix} \text{MODE} & \textit{prop} \\ \text{INDEX} & \boxed{2} \\ \text{RESTR} & \left\{ \boxed{4}, \boxed{5} \right\} \end{bmatrix}$$

The theory as sketched and the representation in (15b) entail that it is a representation of a proposition whose index is of type *situation-index*, and the situation indexed is required to be described by the one-place predication that something satisfies the predicate of pediatricianhood. In the representation of *Kim is a pediatrician* in (16c), that something is required to be whatever satisfies the predication in (16a) that something bears the name *Kim*.

The CONTENT value in (16c) illustrates conformity to the principles of Semantic Compositionality and Semantic Inheritance:

- **Semantic Compositionality:** A phrase's RESTR value is the union of the RESTR values of the daughters.

- **Semantic Inheritance:** A headed-phrase's MODE and INDEX values are structure shared with those of the head daughter.

[27] This analysis is a synthesis of the analysis of Pollard & Sag (1994) as refined in Section 8.5.3 (pp. 342-343), and the Minimal Recursion Semantics analysis as simplified in Sag & Wasow's (1999) introductory-level textbook. The terminology of Minimal Recursion Semantics is explained in detail in Copestake, Flickinger, Pollard & Sag (To appear).

These amount to additional constraints in the definition of *phrase* and *headed-phrase*, respectively.[28]

(17)
$$\left[\begin{array}{l} \textit{phrase} \\ \text{SYNSEM} \mid \text{LOCAL} \mid \text{CONT} \mid \text{RESTR}\ \boxed{0} \cup \ldots\ \boxed{n} \\ \text{HEAD–DTR} \left[\text{SYNSEM} \mid \text{LOCAL} \mid \text{CONT} \mid \text{RESTR}\ \boxed{0}\right] \\ \text{NON-HD–DTRS} \left\langle \left[\ldots\text{RESTR}\ \boxed{1}\right], \ \ldots\ \left[\ldots\text{RESTR}\ \boxed{n}\right] \right\rangle \end{array}\right]$$

(18)
$$\left[\begin{array}{l} \textit{headed-phrase} \\ \text{SYNSEM} \mid \text{LOCAL} \mid \text{CONT} \left[\begin{array}{l}\text{MODE}\ \boxed{1} \\ \text{INDEX}\ \boxed{2}\end{array}\right] \\ \text{HEAD–DTR} \left[\text{SYNSEM} \mid \text{LOCAL} \mid \text{CONT} \left[\begin{array}{l}\text{MODE}\ \boxed{1} \\ \text{INDEX}\ \boxed{2}\end{array}\right]\right] \end{array}\right]$$

6 Constituent structure

As with many other aspects of grammar, HPSG allows both monotonic and default-logic accounts of constituent structure. In the monotonic account of Pollard & Sag (1994), information about the constituent structure of phrases (as well as information about the relation of the constituent parts to each other) is recorded in the various daughters attributes (HEAD-DTR, COMPS-DTRS, SUBJ-DTR, FILLER-DTR, ADJUNCT-DTR, SPR-DTR (SPECIFIER-DTR)) of particular phrase types. These features are all list-valued, enabling them to be present but empty, though HEAD-DTR, SUBJ-DTR, and SPR-DTR seem to have to be limited to being no longer than singleton lists. Thus, a description like (19) indicates a tree with three daughters: a verb head daughter, and two complement daughters (an NP and a PP).

[28] In a path-description (i.e., in a description of a path in a feature-structure from a node through attributes in complex values), "...fsd..." denotes any valid path through a feature-structure satisfying the constraint *fsd*. In a set-description, "..., fsd, ..." denotes a feature-structure-description that is satisfied by any feature structure which contains an arc to a node that satisfies the description *fsd*.

(19)
$$\begin{bmatrix}
\textit{phrase} & \\
\text{PHON} & \left\langle \textit{gives, a, book, to, Sandy} \right\rangle \\
\text{SYNSEM} \mid \text{LOCAL} \mid \text{CAT} & \begin{bmatrix} \text{HEAD} & v \\ \text{SUBJ } \boxed{4} & \left\langle \begin{bmatrix} \textit{synsem} & \\ \text{HEAD} & n \\ \text{SPR } \langle\ \rangle & \end{bmatrix} \right\rangle \\ \text{COMPS } \boxed{5} \ \langle\ \rangle & \\ \text{SPR } \boxed{6} \ \langle\ \rangle & \end{bmatrix} \\
\text{HEAD-DTR} \mid \text{SYNSEM} \mid \text{LOCAL} \mid \text{CAT} & \begin{bmatrix} \text{SUBJ} & \textit{App-synsems } (\boxed{1}, \boxed{4}) \\ \text{COMPS} & \textit{App-synsems } (\boxed{2}, \boxed{5}) \\ \text{SPR} & \textit{App-synsems } (\boxed{3}, \boxed{6}) \end{bmatrix} \\
\text{SUBJ-DTRS } \boxed{1} \ \langle\ \rangle & \\
\text{COMP-DTRS } \boxed{2} \left\langle \text{NP, PP} \right\rangle & \\
\text{SPR-DTRS } \boxed{3} \ \langle\ \rangle &
\end{bmatrix}$$

This analysis employs an *App(end)-synsems* function which appends its second argument (a list of *synsems*) to a list of the SYNSEM values of its first argument (which is a list of phrases). In the case of (19), appending the list of synsems $\boxed{4}$ to the list of the synsems of the elements of the list $\boxed{1}$ yields the list of synsems $\boxed{4}$, because the list $\boxed{1}$ is the empty list. Appending the list of synsems $\boxed{5}$ to the list of the synsems of the elements of the list $\boxed{2}$ yields the list of synsems $\boxed{2}$, because the list $\boxed{5}$ is the empty list. Appending the list of synsems $\boxed{6}$ to the list of the synsems of the elements of the list $\boxed{3}$ yields an empty list, because it amounts to appending the empty list $\boxed{6}$ to the empty list of synsems of the elements of the empty list $\boxed{3}$. It is important to note that "NP" and "PP" are abbreviations for phrases, not synsems, since the values of SUBJ-DTRS, COMPS-DTRS, etc. are lists of phrases, while the values of SUBJ, COMPS etc. are synsems.

Sag (1997) offers an alternative representation which eliminates the redundancy of daughters-features with the valence features by distinguishing subtypes of phrases in terms of relations between the values of their valence features and a NON-HEAD–DAUGHTERS list. Considering the Valence Principle to constrain the values of the valence features of the phrases relative to their values on the head daughter and to the SYNSEM values of the non-head daughters, as described in Section 4, a head-subject phrase (e.g., a finite declarative clause) is defined as in (20a), and a head-complement phrase as in (20b).[29]

(20) a.
$$\begin{bmatrix}
\textit{hd-subj-ph} & \\
\text{SYNSEM} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{SUBJ } \langle\ \rangle \end{bmatrix} \\
\text{HEAD-DTR} \mid \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{SUBJ } \left\langle \boxed{1} \right\rangle \\ \text{SPR } \langle\ \rangle \\ \text{COMPS } \langle\ \rangle \end{bmatrix} \\
\text{NON-HD-DTRS} \left\langle \begin{bmatrix} \text{SYNSEM } \boxed{1} \end{bmatrix} \right\rangle &
\end{bmatrix}$$

[29] The sequence union of lists $\boxed{m}$ and $\boxed{n}$, $\boxed{m} \oplus \boxed{n}$, is the list consisting of $\boxed{n}$ appended to $\boxed{m}$.

b.
$$\begin{bmatrix} \textit{hd-comps-ph} \\ \text{SYNSEM}\,|\,\text{LOC}\,|\,\text{CAT}\ \begin{bmatrix}\text{COMPS}\ \langle\ \rangle\end{bmatrix} \\ \text{HEAD-DTR}\,|\,\text{SYNSEM}\,|\,\text{LOC}\,|\,\text{CAT}\ \begin{bmatrix} \textit{word} \\ \text{SPR}\ \langle\ \rangle \\ \text{COMPS}\ \langle\boxed{1}\rangle \oplus \ldots \oplus \langle\boxed{n}\rangle \end{bmatrix} \\ \text{NON-HD-DTRS}\ \left\langle \begin{bmatrix}\text{SYNSEM}\ \boxed{1}\end{bmatrix}, \ldots\ , \begin{bmatrix}\text{SYNSEM}\ \boxed{n}\end{bmatrix} \right\rangle \end{bmatrix}$$

As shown in (20), some valence values of the phrase and the head daughter are required to be different, but the valence constraint in (8) ensures that all the valence values not specified to be different will be the same. As a consequence, in this approach, the analysis of the verb phrase *gives a book to Sandy* is as in (21).

(21)
$$\begin{bmatrix} \textit{head-comps-ph} \\ \text{PHON}\quad \left\langle \textit{gives, a, book, to, Sandy} \right\rangle \\ \text{SYNSEM}\,|\,\text{LOCAL}\,|\,\text{CAT}\ \begin{bmatrix} \text{HEAD} & v \\ \text{SUBJ}\ \boxed{1} & \\ \text{COMPS}\ \langle\ \rangle & \end{bmatrix} \\ \text{HEAD-DTR}\,|\ \ \text{SYNSEM}|\text{LOCAL}|\text{CAT}\ \begin{bmatrix} \text{SUBJ}\ \boxed{1}\ \left\langle \begin{bmatrix} \textit{synsem} & \\ \text{HEAD} & n \\ \text{SPR}\ \langle\ \rangle & \end{bmatrix} \right\rangle \\ \text{SPR}\ \langle\ \rangle \\ \text{COMPS}\ \left\langle \boxed{2}, \boxed{3} \right\rangle \end{bmatrix} \\ \text{NON-HEAD-DTRS}\ \left\langle \text{NP}\begin{bmatrix}\text{SYNSEM}\ \boxed{2}\end{bmatrix}, \text{PP}\begin{bmatrix}\text{SYNSEM}\ \boxed{3}\end{bmatrix} \right\rangle \end{bmatrix}$$

7 Constituent order

The general outlines of the HPSG approach to constituent order derive from the theory of linear precedence rules developed within the tradition of Generalized Phrase Structure Grammar, as motivated in Gazdar & Pullum (1981), and summarized in Gazdar, Klein, Pullum, & Sag (1985). There are generalizations about the order of constituents in a phrase relative to one another that standard versions of X-Bar theory (cf. Pullum 1985, Kornai & Pullum 1990) are too restrictive to capture without positing a multitude of empty nodes and forced movement chains. The theory of Linear Precedence (LP) rules allows ordering constraints to be stated in terms of any attributes of constituents, as long as the ordering relations hold of EVERY set of sister constituents licensed by the grammar, and this proviso of Exhaustive Constant Partial Ordering (ECPO) imposes a very restrictive constraint on possible grammars. Because these ordering constraints are partial, they allow unconstrained pairs of elements to be ordered freely relative to each other. Thus, as in GPSG, so-called "free word order" (i.e., free phrase order) is a consequence of not constraining the order of constituents at all. (Genuinely free word order, where (any) words of one phrase can precede (any) words of any other phrase requires a word-order function that allows constituents of one phrase to be interleaved with constituents of a sister phrase (Pullum (1982a), Pollard & Sag (1987)).

LP rules for HPSG were discussed at some length in Pollard & Sag (1987, Ch. 7). It was envisioned that linear precedence constraints would be constraints on the PHON[30] values of phrases with content along the following lines:

- A lexical head precedes all of its sisters (or follows, depending on the language).
- Fillers precede phrasal heads.
- Less oblique complements not headed by V precede more oblique phrasal complements.

Presumably an *Order-phon* function would apply to the list consisting of the (append of the) head-daughter and the list of the non-head-daughters to constrain the ordering in the PHON value of the phrase in terms of the relevant properties of the various daughter phrases. Such a function might amount to something paraphrasable as:

> The PHON value of the filler-daughter precedes the PHON value of the head-daughter, and the PHON values of daughters that are words precede those of daughters that are phrases, etc.

As serious grammar development for a number of languages (especially notably, German and French) has made clear, word order constraints are not always compatible with the semantic and syntactic evidence for constituency. German is a "configurational" language, verb-second in main clauses, verb-final in subordinate clauses. However, the constituents of certain types of phrases may be ordered discontintuously–interleaved rather than concatenated–with sister constituents so that the position of complements (and parts of complements!) is remarkably (to English speakers, anyway) free. The German sentences glossed in (22) provide representative examples of the problem.

(22) a. Kaufen glaube ich nicht, dass Maria das Auto will.
buy believe I not that Maria the car wants
I don't believe Maria wants to buy the car.

b. [Das Manuskript gezeigt] hat Maria dem Studenten.
the manuscript shown has Maria the student-DAT
Maria has shown the manuscript to the student.

c. Ich glaube dass der Mann das Lied hat singen koennen.
I believe that the man the song has sing-INF can-INF.
I believe that the man has been able to sing the song.

Thus, in (22a), the head of an embedded complement appears in initial position in the main clause, with its arguments and the finite verb of the complement in their canonical places.[31]

[30] Order is represented in PHON values because PHON is the feature that represents the physical reflections of the constituents, which are abstract, postulated objects of type *sign*.

[31] In (22b), the head of the complement and one of its arguments appears in the initial position of the main clause, with the other two arguments in their normal position after the verb. In (22c), the head (*können*) of the complement of the main verb of the highest embedded clause follows the finite (main) verb *hat*, while the complement *das Lied* of the more deeply embedded verb *singen* precedes *hat*.

The resolution to this dilemma constitutes a lively topic in current research. Kathol (1995), Nakazawa & Hinrichs (1999) and Reape (1994, 1996) explore these issues in more detail (see also Dowty (1996)).

8 The lexicon, lexical relations, lexical rules

As in Lexical Functional Grammar, most of the detail in individual Head-Driven Phrase-Structure grammars is encoded in the lexical entries for particular lexical elements—everything that isn't in the (mostly) universal definitions of phrase-types (which include most of the various named Principles (e.g., the HEAD Feature Principle, the Valence Principle)). But if the specification of phrase types is hierarchical and multi-dimensional, the lexicon is hierarchical and multi-dimensional with a vengeance.

8.1 Organization of the lexicon

What kind of phrase will license a particular lexical head is, as described in Sections 3–6, a function of the argument structure of that head (literally, of the ARG-ST value of the lexical head): what sort of arguments it needs as its complements, subject and/or specifier, whether any of them has a gap in it (see Sec. 10), whether the subject of an infinitive complement must be the same as, or the same in reference as some other argument of the predicate (see Sec. 9). This information is to a large extent predictable: verbs of a certain class require at least an NP direct object; verbs of a certain subclass of that class require that NP object to be identified with the unexpressed subject of an infinitive VP complement. Third-person singular present tense verbs differ systematically from other present tense verbs, and past tense and participial forms of verbs differ systematically from stem (base) forms. In addition, auxiliary verbs differ systematically from main verbs, but this distinction cross-cuts several other ones. Similarly, nouns are classified by whether or not they allow or require a determiner when they are singular (pronouns and proper nouns don't allow a determiner), and what sort of determiner it can be. For example, the quantifier *much* goes only with a mass noun, *many* goes only with a plural count noun, *a/an* requires a singular count noun, and *the* and *some* are not selective.

Facts such as these motivate classifying the elements of the lexicon in multiple dimensions. Argument-taking predicates are classified by transitivity, by argument-coreference with a VP complement's subject (Equi-predicates), by argument-identity with a VP complement's subject (Raising predicates). Nominal expressions are classified by properness, pronominality, and so on. The fact that verbs of EVERY class have inflectional forms drawn from the same set, the fact that different forms of individual verbs encode the same semantic roles, whether or not any argument is unexpressed, and the fact that, in more-inflected languages, nouns (and adjectives) of every class have the same sets of morphological cases motivate lexical representations which link lexemes with their regular and idiosyncratic inflectional characteristics (Miller & Sag (1997), Abeillé, Godard & Sag (1999) Sag & Wasow (1999)). The lexeme dimension of the lexical hierarchy encodes syntactic and semantic information that distinguishes the lexical element from others, including information inherited from higher sorts, and information specifying how semantic roles are linked to grammatical relations and morphological cases (see Davis 1996), as well as any idiosyncratic syntactic or semantic information. The inflectional dimension relates to information that might be reflected in morphological inflection: for example, on a verb, the person, number, and pronominality of its arguments, as well as the presence of all arguments; on a determiner, whether it is count or mass, singular or plural, or indifferent to either distinction, and so on. Thus, there is a lexeme sort *give*, and whole families of words of the form *give, gives, giving, gave, given.*

Finally, there is no pretense that the lexicon is quite as systematic as the foregoing description makes it sound. There is no denying that some properties of lexemes can be described at the most general level only if provision is made for them to have occasional exceptions—either individual lexemes, or particular subclasses whose requirements are contrary to what is true of the class as a whole. Consequently, it is assumed that at least some specifications in the hierarchical lexicon should be represented as having default values, which is to say that inheritance within the lexicon is not strictly monotonic; values specified for a supersort can be contradicted in specifications for a particular subsort (including an individual lexeme) of the sort bearing the default value. For example, the overwhelming majority of nominal lexemes are non-reflexive ([ANA –]), but reflexive pronouns have to be represented as such ([ANA +]) so that the binding theory (see Sec. 11) can refer to them and constrain their distribution. Thus, the specifications for the sort *noun-lexeme* indicate that the property of being non-reflexive is a default: the specification [ANA / –] can be overridden in the specification for the reflexive pronoun subsort of *pronoun-lexeme*.

8.2 Relations among lexical entries

Three kinds of relations among lexemes motivate lexical rules. First, as anyone who has ever done an electronic search, compiled a concordance, or dealt with word frequency lists knows, words that have the same stem and differ only in their inflection (e.g., for number, tense, agreement) count in one sense as "instances of the same word," since they have the same meaning, require the same arguments filling the same roles, and so on. In another sense, of course, they are clearly not instances of "the same word," since they have different inflections. Lexical rules allow the shared characteristics to be stated once in a relatively underspecified lexeme, with the non-shared characteristics specified by lexical rules that depend on the class (or classes) to which that lexeme belongs.

Lexemes can be (somewhat less freely) related derivationally, as well as by inflection. Thus, languages typically have means of making nouns that correspond in regular ways to verbs (e.g., agent nominalizations (*do-er*), zero-affix deverbal result nouns (*kick, spit*), as well as deverbal nominalizations in *-tion*. Languages also have means of making verbs that correspond in regular ways to adjectives (e.g., de-adjectival causative verbs in *-ify* and *-ize*), and to nouns (e.g., *en*-prefixed denominal verbs (*enact, empower, emplane, engulf*)), and to zero-affixed instrumental verbs (*hammer*) and change-of-state verbs (*tile, bone, bottle*). These relations typically embed the meaning of a lexeme with one part of speech within the meaning of a lexeme with a very general meaning, often with a different part of speech. The derived lexeme has a phonology that is a function of both factors.[32]

A third, much more restricted kind of relation involves valence alternations, where two verbs with the same phonology and roughly the same meaning map the same semantic relation to different lists of syntactic categories. Some familiar examples are:

(23)	dative alternation:	*Kim sent Lee a letter. / Kim sent a letter to Lee.*
	causative alternation:	*Fido walked. / Kim walked Fido.*
	telic alternations:	*Dale walked. / Dale walked a mile.*
		Dale ate. / Dale ate lunch.

Levin (1993) offers the most complete description of English valence alternations. In the case of verb alternations, one of the alternants may have more semantic restrictions in addition to having

[32] In zero-affixation cases, the matrix lexeme contributes no phonology at all. When denominal verbs get formed from nouns with irregular plurals, or deverbal nouns get formed from verbs with irregular third singular present tense forms, their categorial properties are determined by the matrix lexeme, while their inflections are completely regular:

one leaf, two leaves but *to leaf, it leafs*
to do, he [dʌz] but *a do, two [duz]*

a different mapping of arguments to roles.

8.3 Lexical rules

Lexical rules express the correspondences among lexemes and words of various sorts. The basic idea of a lexical rule is to define a class of words or lexemes in terms of how its members correspond to and differ from a class of lexemes or another class of words, usually a class of the same size or smaller. Thus, in general, a lexical rule says: for every lexeme meeting [such-and-such] specifications, there is a word (or a lexeme) with some particular additional constraints that satisfies all of those specifications, except for specifications that are directly or indirectly contradicted by the additional constraints.

The most obviously necessary classes of cases are the lexical rules that define inflected words on the basis of lexemes unspecified for inflections. For example, in the grammar of English, there is a lexical rule that defines present participles in terms of verb-lexemes unspecified for any inflection, and another one for third-singular present tense finite verbs, and one for non-third-singular verbs, and so on.

Other classes of lexical rules define classes of lexemes in terms of correspondences involving their valence specifications and semantic roles. For example, a causative lexical rule might define lexemes that systematically had an additional, less oblique argument with an agentive role in its semantics.

In the earlier literature (cf. Pollard & Sag 1994, Chapter 9), an "extraction" lexical rule defines lexemes with a non-empty SLASH value and an ARG-ST list that systematically has one more element than the list-append of its valence-feature lists, where the LOCAL value of the additional element is the same as the SLASH value. Similarly, a null-pronoun lexical rule would define verb-lexemes with a pronominal element in their ARG-ST list which again is greater than the list-append of its valence-feature lists by exactly that element. An alternative, explored in more recent work (e.g., Bouma, Malouf, & Sag 1998) and described in Section 10.3, builds these correspondences into constraints on lexical types.

More specifically, in the analysis outlined in the last chapter of Pollard & Sag (1994) and refined in Sag (1997), a Complement Extraction lexical rule defines a lexical item that is identical to a similar lexical item, except that it systematically has a synsem of type *gap* (as defined in (24)) on its ARG-ST list in some noninitial (i.e., non-subject) position, and its COMPS list has no corresponding synsem of any sort on it.

(24) $$\begin{bmatrix} \textit{gap-synsem} \\ \text{LOCAL } \boxed{1} \\ \text{SLASH } \left\{ \boxed{1} \right\} \end{bmatrix}$$

Gap-synsems never describe the syntactico-semantic properties of actual constituents, because actual constituents must be *signs*, and the synsem of a sign must be of the sort *canonical-synsem*. The constraint that a sign's synsem must be *canonical* means that there are no "empty" constituents.

In both analyses the theory defines words whose ARG-ST lists are more than the *Append* of their SUBJ and COMPS lists, by means of implicit or explicit reference to words that are identical except for the specified properties. Other constraints (see Section 10) require gap-synsems to share their LOCAL value with a non-local ("extracted") dependent.

Finally, lexical rules can define both an inflection and a derived subcategorization. For example, Sag & Wasow's (1999) passive lexical rule states a correspondence between verb lexemes whose COMPS list has at least one argument-saturated nominal *synsem* on it, and words where an argument-saturated nominal *synsem* with those specifications is the the sole member of the SUBJ list, and is absent from the COMPS list, which may have a *synsem* for an (oblique) *by*-phrase on it whose NP object has the same index as the argument-saturated nominal *synsem* on the SUBJ list of the source lexeme, as shown in (25).[33]

(25) $$\left\langle \begin{bmatrix} \textit{verb-lexeme} & \\ \textrm{SYNSEM}|\textrm{LOC} & \begin{bmatrix} \textrm{CAT}|\textrm{ARG-ST} & \langle\, \boxed{2}_{\boxed{4}},\ \boxed{3},\ \ldots\, \rangle \\ \textrm{CONT} & \boxed{1} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \textit{verb-word} & \\ \textrm{SYNSEM}|\textrm{LOC}|\textrm{CAT} & \begin{bmatrix} \textrm{HEAD} & \begin{bmatrix} \textrm{VFORM} & \textit{psv} \end{bmatrix} \\ \textrm{ARG-ST} & \langle\, \boxed{3},\ \ldots,\ (\textrm{PP}_{\boxed{4}})\, \rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

The informal notation in (25) represents lexical rules as functions from the class described by the domain description to the class defined by the range description.[34] A problem with this informal notation is that it does not make explicit that all properties of the domain class are the same in the defined range class, except where specified or entailed to be different. As Meurers and Minnen (1997) show, the task of fully specifying how to compute what is defined by a lexical rule is a non-trivial one. First of all, ensuring that properties mentioned in the domain specification but not the range specification are present in the sorts defined by the rule requires a closed world assumption (cf. Meurers & Minnen 1997). That is, it requires a conception of lexical rules as relations among already existent elements in the lexicon. It is in effect a way of describing the lexicon by organizing it, rather than, say, by using a meta-level description to generate it from "basic" elements. It is a similarly non-trivial problem to determine what properties that are mentioned in neither the domain specification nor the range specification are required (by type constraints) to be present in the defined sort. Furthermore, sometimes the properties of the resultant sort have no relation to properties of the domain sort—for example, when the values for certain features shared by both sorts are independent of each other, and when the sorts differ in such a way that features are defined for one of the sorts but not for the other. Providing a real syntax and semantics for a convenient representation for lexical rules is a non-trivial problem because the range specifications are not always consistent with the domain specifications. For instance, in the passive lexical rule in (25), constraints on *word* or *lexeme* entail that feature structures denoted by the domain description have a COMPS list that begins with whatever $\boxed{3}$ denotes; in the range description, those same constraints entail that $\boxed{3}$ be the sole member of the SUBJ list, and that it not be on the COMPS list at all.

Meurers 1995 and Meurers & Minnen 1997 provide an eye-opening discussion of just what is necessary to ensure that lexical rules entail exactly what linguists expect them to entail in a fully explicit system. The latter work explains what it takes to accomplish this in a computational implementation without invoking non-branching syntactic rules in the grammar. They treat lexical rules as descriptions of parts of the sort hierarchy, licensed only if each of their subdescriptions is licensed (i.e., is defined by the grammar). At the same time, they allow for "phantom" lexical rules which are defined so that the resultant sorts cannot participate in any constructions, but can be the domain for other lexical rules. Interestingly, their computational implementation (which is faithful

[33]By convention, a subscripted tag denotes the INDEX value of what it is subscripted to.

As is well-known, non-predicative PPs like the passive *by*-phrase have the properties of the NPs that are their objects for many syntactic phenomena, including binding. For discussion, see Sag & Wasow, 1999 (Chapter 7). This fact is represented in the informal notation of (25) by the coindexing of the subject NP in the domain lexeme's valence structure with the optional prepositional phrase in the range word's valence structure.

[34]In many computational implementations, lexical rules are represented as non-branching (*head-only* or *non-headed*) phrase-structure rules. Representing lexical rules as derivational chains in syntactic structures, with one node for each class of words involved, seems to distort the insight that lexical rules are intended to characterize implicational relations among classes of words in the lexicon.

to the central ideas of HPSG) involves maneuvers which mimic the non-branching phrase-structure rule approach, but all of that is part of the computational implementation (the compiler for the grammar), not part of the grammar.

9 Complementation

9.1 Complementizers

On the assumption that clauses and verb phrases with and without complementizers are subsorts of the same sort, (i.e., that examples like (26a,b) belong to the same (super)sort, as do examples like (27a,b,c)), HPSG analyses (e.g., Pollard & Sag 1994, Sag 1997) treat them as differing in only one or two features.

(26) a. Sandy went home

b. that Sandy went home

(27) a. sign the form

b. to sign the form

c. for Lee to sign the form

The analysis of Sag (1997) integrates Pullum's (1982b) insight that the English infinitival complementizer *to* behaves in many ways like a subject-to-subject raising verb like *seem*[35] without claiming that *to* is a verb. Instead it treats verbs and complementizers generally as distinct subsorts of a the part-of-speech subsort *verbal*. The infinitival complementizers *to* and *for* turn out to be much like raising verbs in that they have arguments which are the same as their verbal complement's subject. The main property that distinguishes complementizers from verbs is that the VFORM value of any complementizer is the same as that of its complement; since VFORM is a HEAD feature, the VFORM value of the complementizer-headed phrase will have to be the same as the VFORM value of the complement also. Verbs (like *know*) that are indifferent to whether their finite complement has a complementizer or not ("*that*-deletion" verbs) are simply unspecified for the category of their complement, requiring only that it have a CONTENT value of sort *proposition*.

Thus in Sag (1997), *to* is a complementizer which is like a subject-to-subject raising verb in subcategorizing for a *base* form[36] VP complement whose subject is the same as the complementizer's subject, while *for* is like a raising-to-object verb in subcategorizing for an NP complement and a *base*-form CP complement whose subject is the same as the NP complement, but no subject at all.[37]

[35]This insight was incorporated in GPSG as the claim that *to* WAS an auxiliary verb (the unique [VFORM *inf*] member of a subclass of [VFORM *aux*] verbs).

[36]The present reanalysis of complementizers eliminates the need for two "infinitival" values (*base* and *inf*). Sag 1997 and Sag & Wasow 1999 call the remaining value *inf*; I prefer to call it *base* precisely because of the frequent uncertainty or confusion as to whether "infinitive" refers to complementized verbs or bare verbs.

[37]Inverted auxiliary verbs are similarly analyzed as having no subject, but an NP complement and a VP complement whose subject shares structure with the NP complement, following analyses suggested by Borsley (1986, 1989), as discussed in Pollard & Sag 1994: 351-352.

9.2 Infinitival complements: Equi and raising

Thus, infinitive complements are treated as projections of verbal heads. Equi and Raising structures both are projections of heads which subcategorize for an unsaturated predicative complement, and indeed, have the same possibilities for constituent structure—either (28a) or (28b), depending on the verb.

(28) a. [$_{S}$ NP [$_{VP}$ V CP[*base*]]]

b. [$_{S}$ NP [$_{VP}$ V NP CP[*base*]]]

Pretheoretically, the argument structure difference between Raising verbs and Equi verbs is that Raising verbs have an argument to which they don't assign a semantic role, while Equi verbs assign roles to all their arguments. Pollard & Sag (1994) represent this difference by saying that an Equi verb subcategorizes for an NP with an index (of sort *ref*, i.e., not an expletive) which is the same as the index of the SUBJ specification of its complement, and assigns a semantic role to the index of the coindexed NP, as indicated in (29a), while a Raising verb requires its subject to share the LOCAL value of the feature structure of the synsem that its complement VP selects as subject, but assigns no semantic role to the index of that element, as indicated in (29b).[38]

(29) a.
$$\begin{bmatrix} \textit{intr–Equi–verb} \\ \text{CAT} \mid \text{ARG-ST} \left\langle \text{NP}_{\boxed{1}},\ \text{CP} \left[\textit{base},\ \text{SUBJ} \left\langle \text{NP}_{\boxed{1}} \right\rangle \right]_{\boxed{2}} \right\rangle \\ \text{CONTENT} \begin{bmatrix} \textit{try} \\ \text{TRIER}\ \boxed{1}\textit{ref} \\ \text{SITUATION}\ \boxed{2} \end{bmatrix} \end{bmatrix}$$

b.
$$\begin{bmatrix} \textit{intr–raising–verb} \\ \text{CAT} \mid \text{ARG-ST} \left\langle \left[\text{LOC}\ \boxed{1}\right],\ \text{CP} \left[\textit{base},\ \text{SUBJ} \left\langle \left[\text{LOC}\ \boxed{1}\right] \right\rangle \right]_{\boxed{2}} \right\rangle \\ \text{CONTENT} \begin{bmatrix} \textit{tend} \\ \text{SITUATION}\ \boxed{2} \end{bmatrix} \end{bmatrix}$$

The absence of a role assignment for one subcategorized element for raising verbs entails that the content of that argument has no semantic relation to the raising verb. Thus, there is no reference to an index for *Pat* in the semantic representation of *tend* in (30).

[38] Most of the HPSG literature treats Raising in terms of shared SYNSEM values. Ginzburg & Sag (To appear) treat Raising in terms of shared LOCAL values, because the shared SYNSEM value analysis incorrectly predicts that nonlocal features such as SLASH (or AFF in French) will be represented on both the raising verb and the verb whose logical subject is the raised NP. In addition, it appears (Sag, personal communication) that in languages where extractions are reflected in verb morphology, that morphology appears on the raising verb, but not on the head of the complement verb.

(30) Pat tends to like jazzy arrangements.

Assignment of a role to the index of an Equi verb's subject entails that sentences like (31a) with active Equi complements will have different truth-conditional semantics from ones like (31b) with passive complements.

(31) a. Sandy persuaded the doctor to examine Kim.

b. Sandy persuaded Kim to be examined by the doctor.

By the same logic, sentences with active and passive raising complements will have the same truth-conditional semantics, as in (32).

(32) a. Sandy allowed the doctor to examine Kim.

b. Sandy allowed Kim to be examined by the doctor.

The restriction that the Equi controller have an index of sort *ref* follows from the assignment of a semantic role (because arguments of roles in relations have to be of sort *ref* or sort *situation* (Pollard & Sag 1994: 397). This precludes the possibility of expletive Equi controllers, which indeed do not exist, although Raising controllers can have expletive subjects and complements, as illustrated in (34) and (33).

(33) a. There tried to be a protest against garment manufacturers with plants abroad.

b. There seemed to be a protest against garment manufacturers with plants abroad.

(34) a. *Sandy persuaded there to be a party after the first week of classes.

b. Sandy allowed there to be a party after the first week of classes.

Structure–sharing between the valence values in raising constructions predicts the possibility of 'quirky' case on Raising controllers as in Icelandic (Andrews (1982); Sag, Karttunen, & Goldberg (1992)), and the existence of non-NP raising controllers. Non-nominal phrases that occur as the subject of *be* also occur as subjects of subject-to-subject raising verbs, as shown in (35).

(35) a. Here and not earlier seemed to be the best place to introduce that approach to extraction constructions.

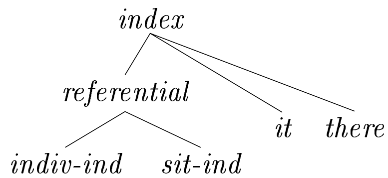
b. Grilled or baked is how they prefer their fish.

c. Very carefully is the best way to approach a 600-pound gorilla.

Semantic roles are assigned only to situational and individual indexes. Consequently, roles are never assigned to expletives, and role-assigned arguments are never expletives. But some predicates do subcategorize for expletive subjects, for example:

- "weather" expressions (*it*): *rain, late, Tuesday...*
- existential verbs (*there*): *be, arise, occur, ...*
- extraposition verbs and adjectives (*it*): *seem, bother, obvious...*

In fact, as demonstrated by Postal & Pullum (1988), some predicative expressions subcategorize for expletive objects. For example, transitive idioms like *wing, go at, out of...* require an expletive object, as do extraposition predicates like *resent, take, depend upon...*, which require a sentential complement in addition. The HPSG analysis is that the expletive *it* has a [PER *3rd*, NUM *sg*] index of sort *it*, the expletive *there* has a [PER *3rd*] index of sort *there*, and both index sorts are subsorts of the sort *index*, along with the subsort *ref*.



The appearance in *there*-constructions of agreement between the verb and its first object, as in (36) comes from the fact that the verb subcategorizes for a direct object whose NUM value is shared with that of its *there* subject.

(36) a. There are two rabbits in the garden.

b. *There are a rabbit in the garden.

Agreement is, as usual, a linking of a morphological form of the verb to the value of the index of the subject it subcategorizes for. The Valence Principle interacts with the lexical specifications of raising verbs to allow the subcategorization requirements of verbs recursively embedded in raising structures, as in (37), to be satisfied by an indefinitely higher NP.

(37) a. There seem to have to be two defenders in the backfield at all times.

b. *There seem to have to be a keeper near the goal at all times.

Note that structure–sharing is again critical for expressing correspondences within Equi and Raising constructions. Thus, the valence specifications of the raising verb *tend* are represented as in (38).

(38)
$$\begin{bmatrix} \textsc{subj} \left\langle \begin{bmatrix} \textsc{loc}\ \boxed{1} \end{bmatrix} \right\rangle \\ \textsc{comps} \left\langle \mathrm{VP} \begin{bmatrix} \textsc{vform}\ \textit{inf} \\ \textsc{subj} \left\langle \begin{bmatrix} \textsc{loc}\ \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

This constraint says that *tend* needs as a subject whatever its VP complement needs as ITS subject. It specifies *tend*'s SUBJ value as identical to the SUBJ value of the VP which *tend* selects as its complement. Similarly, (39) represents a description of the valence value of a raising verb in a structure where it happens to have a quirky-case infinitive complement.

(39)
$$\begin{bmatrix} \textsc{subj} \left\langle \begin{bmatrix} \textsc{loc}\ \boxed{1}\mathrm{NP} \end{bmatrix} \right\rangle \\ \textsc{comps} \left\langle \mathrm{VP} \begin{bmatrix} \textsc{vform}\ \textit{inf} \\ \textsc{subj} \left\langle \begin{bmatrix} \textsc{loc}\ \boxed{1} \begin{bmatrix} \textsc{case}\ \textit{gen} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

The structure–shared SUBJ values entail that the case of the subject selected by the VP complement must be realized on the subject of the raising verb taking that complement.

10 Extraction

The general outline of the HPSG treatment of unbounded extractions follows the three-part strategy developed in GPSG (Gazdar 1981, GKPS 1985).

- An extra constituent is constrained to match a constituent missing from its clausal sister, and what is missing is represented in the description of the clausal sister as the value of the extraction-recording feature SLASH.
- The clausal sister must be missing a constituent (not necessarily an immediate constituent).
- The correspondence between the extra constituent and the constituent that is missing is recorded by means of local (mother-daughter) correspondences over an indefinitely large array of structure.

10.1 Licensing the "extra" constituent

Following work by Hukari & Levine (1987, 1991), HPSG distinguishes between strong and weak extractions. In strong extraction constructions like (40), the extra constituent has all the categorial properties expected for the missing constituent.

(40) a. Okra, I don't think anyone will eat — .

b. The refrigerator in which everyone thinks someone left some limburger — may be hard to sell.

As shown in (41), the head-daughter's value for SLASH shares structure with the LOCAL value of a non-argument filler-daughter.[39]

(41)
$$\begin{bmatrix} \textit{head-filler phrase} \\ \text{SYNSEM}|\text{NONLOCAL}|\text{SLASH}\ \{\} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{NONLOCAL}|\text{SLASH}\ \left\{\boxed{1}\right\} \\ \text{NON-HEAD-DTRS}\ \left\langle \begin{bmatrix}\text{SYNSEM}|\text{LOCAL}\ \boxed{1}\end{bmatrix}\right\rangle \end{bmatrix}$$

In weak extraction phenomena, which are all licensed as a property of a particular lexical element, a constituent that is the argument of some predicative element must be coreferential with the missing constituent. As illustrated in (42), only coindexing is required (not full categorial identity) between some constituent and the value of SLASH on another constituent.

[39] This is slightly simplified. The more accurate representation below indicates that any elements in the head daughter's value for SLASH that are not matched by the LOCAL value of the nonhead daughter remain as the phrase's value for SLASH.

$$\begin{bmatrix} \textit{head-filler phrase} \\ \text{SYNSEM}|\text{NONLOCAL}|\text{SLASH}\ \boxed{2} \uplus \boxed{3} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{NONLOCAL}|\text{SLASH}\ \boxed{2} \uplus \left\{\boxed{1}\right\} \\ \text{NON-HEAD-DTRS}\ \left\langle \begin{bmatrix}\text{SYNSEM}\ \begin{bmatrix}\text{LOCAL}\ \boxed{1} \\ \text{NONLOCAL}\ \begin{bmatrix}\text{SLASH}\ \boxed{3}\end{bmatrix}\end{bmatrix}\end{bmatrix}\right\rangle \end{bmatrix}$$

(The symbol $\uplus$ represents disjoint set union, which is just like the familiar set union, except that it is only defined for sets with an empty intersection).

(42) Partial description of a *tough*-class predicate

$$\left[\text{ARG-ST}\ \left\langle \left[\text{LOCAL}\ \boxed{1}_{\boxed{2}}\right], \ldots, \left[\text{SLASH}\ \left\{\boxed{3}_{\boxed{2}}\right\}\right], \ldots \right\rangle\right]$$

The operative difference is that in strong cases, the CASE value of the two elements must match; in weak ones it need not (Pollard & Sag 1994: 187). That is, in weak extractions, CASE is specified on arguments independently of case specified for the missing constituents, as in phrases like those in (43), where the missing constituent and the item it is coindexed with are required to have different cases.

(43) a. He$_i$ is easy to please —$_i$. (*tough*-complements)

b. I$_i$ am available to dance with —$_i$. (purpose infinitives)

c. I gave it to the man$_i$ Dana thinks —$_i$ is French. (*that*-less relative clauses)

d. It's me$_i$ who Dana says —$_i$ is ferocious.[40] (*it*-clefts)

In weak extraction cases like *tough*-constructions, a head of the relevant class selects a complement with a non–null SLASH specification, as shown in (42); this entails that some descendent of this complement will not be lexically realized.

10.2 Licensing the absence of the "missing" constituent

The nature of the non–realization of the missing constituent is still a matter of some dispute. In early versions of modern phrase structure grammar, missing constituents were treated as traces, i.e., as lexically defined phrasal constituents of various category types which in each case were missing a constituent of exactly that type; thus an NP-trace was NP[SLASH NP], a PP-trace was PP[SLASH PP], a PP[*to*]-trace was PP[*to*, [SLASH PP[*to*]], a 3rd-singular-NP-trace was NP[NUM *sg*, PER 3, [SLASH NP[NUM *sg*, PER 3]], and so on. Traces were licensed in phrases by defining, for each lexical element that subcategorizes for one or more phrasal sisters, a corresponding item with a phrasal trace sister. In most current versions of HPSG, it is accomplished by a lexical rule or constraint which defines lexical entries which lack certain elements on valence lists, and have corresponding elements in their SLASH sets.

Missing embedded subjects, which are not sisters of lexical heads, have been licensed in HPSG (following a GPSG analysis) by a lexical rule which lets a lexical head that would ordinarily subcategorize for an S instead subcategorize for a VP (i.e., an S that lacks a subject), just in case its mother is missing an NP (main clause subject relatives and interrogatives like *Who likes M&Ms?* and *someone who likes M&Ms* were treated as not involving any SLASH dependency at all, but simply a WH-valued subject). This treatment has what was at one time regarded as the happy consequence of entailing the familiar *that*-trace facts (Gazdar et al. (1985): 57-162, Pollard & Sag 1994: 384). However, a number of facts have more recently been seen to converge in favor of treating subject extraction as simply another instance of the same filler–gap relation as is seen in complement extraction. For example, the fact that in several languages a clause whose morphology reflects the fact that it is missing an argument is marked the same way whether it is missing a subject or a complement[41] suggests that there ought to be a universally available means of treating

[40] Note that here there is no correspondence of values for either CASE or PERSON.

[41] Cf. Clements, McCloskey, Maling & Zaenen (1983).

subject and complement extraction uniformly.[42] Furthermore, sentences with parasitic gaps dependent on subject gaps in non-complement clauses such as (44a) would not be allowed if subject extractions were described by a rule which treated missing subjects as constructions which were stipulated to consist of just an S[SLASH NP] mother and an unslashed VP daughter, as the original acount maintained.[43]

(44) a. Sandy is someone who until you know == well, == can seem quite cold.

b. Sandy is someone who until you know her history, == can seem quite cold.

c. *Sandy is someone who until you know == well, her manner can seem quite cold.

However, they are predicted by treating all missing subjects the same way as missing complements. In addition, the *that*-trace effect has been shown to vanish when material bearing a phrasal stress (such as an adverbial phrase) intervenes between the complementizer and the site of the missing subject (a point noted in passing in Bresnan (1977) and more recently rediscovered and investigated by Culicover (1993)). Treating missing subjects and complements together with a single missing dependents rule gives the theory of extraction constructions in HPSG a more homogeneous appearance.[44]

10.3 Guaranteeing that the extra constituent matches what's missing

The correspondence between the extra constituent and the missing constituent, which precludes the possibility of sentences like (45) with an extra NP and a missing PP, is guaranteed by a constraint on the occurrence of unbounded dependency (NONLOCAL) features.

(45) *I wonder [what table]$_{NP}$ he will put the books [—]$_{PP}$.

Such features are constrained to appear on a phrase if they are present with the same value on at least one daughter,[45] and on a daughter constituent only if they are present with the same value on the mother. Thus, the match between the extra constituent and the missing constituent is "a global consequence of a linked series of local mother-daughter feature correspondences" (GKPS: 138). In Pollard & Sag (1994), this was achieved by the Nonlocal Feature Principle, a configurational constraint[46] on all phrases. In more recent treatments (Sag (1997), Bouma et al. (1998)), a lexical constraint requires SLASH values of all elements on an ARG-ST list to be recorded in the SLASH set of the word,[47] as shown in (46).[48]

[42]Cf. Hukari & Levine (1996), Hukari & Levine (Ms.), Bouma, Malouf & Sag (ms.) and Ginzburg & Sag (To appear, Chapter 6).

[43]Cf. Haegeman (1984), Hukari & Levine (1996), Hukari & Levine (Ms.)

[44]Cf. Bouma, Malouf & Sag (ms.) and Ginzburg & Sag (To appear, Chapter 6).

[45]In most recent formulations, on the head daughter (cf. Sag 1997).

[46]It referred to specifications within DAUGHTERS attributes requiring them to appear on a phrase if they are present with the same value on at least one daughter, and on a daughter constituent only if they are present with the same value on the mother. Specifically, the NONLOCAL Feature Principle of Pollard & Sag (1994) required that the INHERITED value of any nonlocal feature on a phrase be the union of the INHERITED values for that feature on the daughters, minus the value of the TO–BIND feature on the head daughter. The more recent treatments described just below achieve the same effect word-internally, without invoking an INHERITED feature for bookkeeping. They track SLASH- and WH-binding in head-valence phrases through heads: Inheritance Constraints ensure that the value of each NONLOCAL feature on a phrase is the same as its head daughter's value for that feature.

[47]In Bouma et al. (1998), this approach is extended to include postverbal adverbials with variable scope, by recording them as dependents in *head-arguments-phrases*.

[48]Similarly for the NONLOCAL features REL and QUE. Having the value for SLASH be the set difference of the amalgamated set and whatever is bound off by the lexical item allows weak extractions, which are lexically governed, to be included by the constraint.

(46) **SLASH Amalgamation Constraint:**

$$\begin{bmatrix} \textit{word} \\ \text{SYNSEM} \begin{bmatrix} \text{LOCAL} \begin{bmatrix} \text{CAT}|\text{ARG-ST} \left\langle \begin{bmatrix}\text{SLASH } \boxed{1}\end{bmatrix}, \ldots \begin{bmatrix}\text{SLASH } \boxed{n}\end{bmatrix} \right\rangle \\ \text{BIND } \boxed{0} \end{bmatrix} \\ \text{NONLOCAL}|\text{SLASH } (\boxed{1} \cup \ldots \cup \boxed{n}) - \boxed{0} \end{bmatrix} \end{bmatrix}$$

A constraint on *head-nexus-phrases*, shown in (47), constrains the SLASH value of a phrase to be the same as the SLASH value of the head daughter.[49]

(47) **NONLOCAL Inheritance Constraint:**

$$\begin{bmatrix} \textit{head-nexus-ph} \\ \text{SYNSEM}|\text{NONLOCAL} \ / \ \boxed{1} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{NONLOCAL} \ / \ \boxed{1} \end{bmatrix}$$

10.4 Multiple Extractions

In contrast to many other treatments of extractions and other unbounded dependencies, in HPSG, unbounded dependency features including SLASH take a **set** of feature structures as values, not a single feature structure. This allows for the possibility of sentences with multiple (non-coreferential) extractions, as in (48).

(48) This is a problem$_i$ which John$_j$ is tough to talk to —$_j$ about —$_i$.

Like GPSG, the account of HPSG in Pollard & Sag (1994) licenses multiple binding of a single extracted element as in (49) by not saying anything to prevent it.

(49) a. That was the rebel leader who$_i$ rivals of —$_i$ shot — $_i$.

b. Those reports$_i$, Kim filed —$_i$ without reading —$_i$.

c. Which relatives$_i$ should we send snapshots of —$_i$ to —$_i$?

Such structures satisfy the constraints on the occurrence of NONLOCAL features. There is a much wider range of acceptability judgements for such constructions than is usually acknowledged,[50] and after considering various alternatives, Pollard & Sag (1994) conclude that in more restrictive dialects, grammars are constrained to allow the first element on a lexical head's argument–structure list to contain a gap only if something else on the list does.[51]

[49] In strictly monotonic accounts, the NONLOCAL features are constrained by distinct constraints that live on different types (namely, *head-filler-ph, head-nexus-ph*, *head-valence-ph* (a subsort of *head-nexus-ph* that excludes *head-filler-ph*).

[50] For example, sentences like (i) and (ii) are routinely cited in scholarly discussions as unacceptable (in support of a claim that gaps in non-heads are dependent on coreferential gaps in a head). In fact, however, they are perfectly acceptable to many native speakers.

i. That is the rebel leader who$_i$ rivals of —$_i$ shot the British consul.

ii. Which rebel leader did you think [[my talking to —] would be dangerous]?

[51] An alternative, based on the claim in Postal (1994) that only NPs (and not say, PPs) are involved in so-called parasitic gaps, is to approach the analysis of these sentences in terms of coreference but not argument identity, i.e., as "null" resumptive pronouns (Sag 1997: 447-448). Levine, Hukari & Calcagno (1999) offer evidence against Postal's claim, and in support of the P&S94 analysis.

10.5 Empty categories

Neither extraction traces nor so-called null or zero pronouns need to be represented as abstract or invisible constituents. Extractions are represented as structure–sharing between a member of a lexical head's SLASH set, and the LOCAL value of an element that is on its ARG–ST list, but not on its COMPS or SUBJ lists. The lexical rules (Pollard & Sag 1994: 446-451) which define this relation have the form of a function (schematized in (50))[52] from lexical entries to lexical entries which are identical except that

a. they contain on their ARG-ST list an element whose LOCAL value is the same as its SLASH value;

b. an element with the same LOCAL value is absent from the COMPS list;

c. the SLASH set is augmented by that LOCAL value.

(50)
$$\left\langle \begin{bmatrix} \text{ARG-ST} \ \langle \ldots, \boxed{3} \begin{bmatrix} \text{LOCAL} & \boxed{1} \end{bmatrix}, \ldots \rangle \\ \text{COMPS} \ \boxed{5} \bigcirc \boxed{3} \\ \text{SLASH} \ \boxed{2} \end{bmatrix}, \begin{bmatrix} \text{ARG-ST} \ \left\langle \ldots, \boxed{4} \begin{bmatrix} \text{LOCAL} \ \boxed{1} \\ \text{SLASH} \ \left\{ \boxed{1} \right\} \end{bmatrix}, \ldots \right\rangle \\ \text{COMPS} \ \boxed{5} \\ \text{SLASH} \ \boxed{2} \cup \left\{ \boxed{1} \right\} \end{bmatrix} \right\rangle$$

Note that $\boxed{4}$ is the same as $\boxed{3}$, except that the SLASH specification is added, which entails a different tag; this formulation is consistent with the idea that specifications of the range of a lexical rule are the same as in the domain except where specified to be different.

The more recent analysis of Bouma et al. (1998) requires all words to satisfy the constraint that the list of synsems constituting the COMPS list is the list of synsems which is the value of ARG-ST minus the elements on the SUBJ list and any *gap-synsems* which are on the ARG-ST list.[53] In such analyses, the SYNSEM value of *ate* in *Bagels, Kim ate* would be represented as in (51).

(51)
$$\begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{SUBJ} \ \left\langle \boxed{4} \right\rangle \\ \text{COMPS} \ \langle \ \rangle \\ \text{ARG–ST} \ \left\langle \boxed{4}\text{NP}_{\boxed{1}}, \boxed{5}\text{NP} \begin{bmatrix} \text{LOCAL} \ \boxed{3}_{\boxed{2}} \\ \text{SLASH} \ \left\{ \boxed{3} \right\} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \textit{eat} \\ \text{EATER} \ \boxed{1} \\ \text{EATEN} \ \boxed{2} \end{bmatrix} \end{bmatrix} \\ \text{NONLOCAL} \mid \text{SLASH} \ \left\{ \boxed{3} \right\} & \end{bmatrix}$$

[52] The shuffle of lists $\boxed{m}$ and $\boxed{n}$, $\boxed{m} \bigcirc \boxed{n}$, is any of the lists consisting of the elements of $\boxed{n}$ interleaved with those of $\boxed{m}$ which retain the order defined for each list, like each shuffle of a deck of cards.

[53] See Appendix B for statement of these constraints.

Null pronouns are also treated as synsems selected via a head's argument–structure list. They have no corresponding phrases in any daughter's specifications, and for this reason they do not appear on any valence feature list. Since they are not in a syntactic dependency relation with anything, they are not represented in SLASH sets either. Null pronouns are represented on the argument–structure list because they behave the same way as phonologically realized pronouns with respect to binding, and binding is a relation among elements on argument–structure lists (see Sec. 11). Because null pronouns are represented on the argument–structure list and in CONTENT, but not on SUBJ or COMPS lists, expressions containing them do not count as being unsaturated on that account. To illustrate, the representation of *ate* in an answer utterance like *Ate squid* (cf. Japanese *Ika-o tabeta*) is sketched in (52):

(52)
$$\left[\begin{array}{ll} \text{LOCAL} & \left[\begin{array}{ll} \text{CAT} & \left[\begin{array}{l} \text{SUBJ } \langle\ \rangle \\ \text{COMPS } \left\langle \boxed{5} \right\rangle \\ \text{ARG–ST } \left\langle \textit{pronoun}_{\boxed{1}},\ \boxed{5}\text{NP}_{\boxed{2}} \right\rangle \end{array}\right] \\ \text{CONT} & \left[\begin{array}{l} \textit{eat} \\ \text{EATER } \boxed{1} \\ \text{EATEN } \boxed{2} \end{array}\right] \end{array}\right] \\ \text{NONLOCAL} \mid \text{SLASH } \{\} & \end{array}\right]$$

Implied arguments (e.g., the implied argument of *ate* in *I already ate*) have been analyzed since Pollard & Sag (1987) as distinct from null pronouns. They have generic or nonspecific rather than definite referents,[54] and are represented in the existence of appropriate roles in the CONTENT representation. However, no indexes are assigned to those roles, and there are no corresponding synsems in ARG-ST and valence feature lists or SLASH sets. The nonspecific reference is attributable to the fact that no constraints are specified in the grammar on the index for the relevant role in the CONTENT value, as illustrated in (53).

(53)
$$\left[\begin{array}{ll} \text{LOCAL} & \left[\begin{array}{ll} \text{CATEGORY} & \left[\begin{array}{l} \text{SUBJ } \left\langle \boxed{4} \right\rangle \\ \text{COMPS } \langle\ \rangle \\ \text{ARG–ST } \left\langle \boxed{4}\ \text{NP}_{\boxed{1}} \right\rangle \end{array}\right] \\ \text{CONTENT} & \left[\begin{array}{l} \textit{eat} \\ \text{EATER } \boxed{1} \\ \text{EATEN } \textit{index} \end{array}\right] \end{array}\right] \\ \text{NONLOCAL} \mid \text{SLASH } \{\} & \end{array}\right]$$

For discussion of the many issues involved in the analysis of extraction of and from adjuncts including examination of issues involving so-called parasitic gaps, see Levine & Pollard (this volume).

[54]With a few verbs (e.g., *eat, drink*), they have indefinite referents referring to salient exemplars. See Cote 1996 for discussion.

10.6 Constraints on gaps

Following initial work by Chomsky (1964) and Ross (1967), research in syntax has sought to identify environments in which filler–gap linkages are precluded, and to formulate independent conditions which would predict them. As early as the 1970s, however, alternative explanations began to be fleshed out for the facts which the syntactic constraints were supposed to account for (e.g., Grosu 1972). More recently, it has become apparent that much of the data on which many of these constraints have been based are not representative of the syntactic classes originally assumed; acceptability often turns on the choice of lexical items or use of, e.g., a definite article rather than an indefinite article. In other cases it has become clear that phenomena used to support universal claims are in fact language-specific.

Nonetheless, certain broad classes of effects do emerge from the analysis of extraction constructions. For example, linking gaps to lexical selection[55] predicts most of the English (generalized) "left branch" phenomena discussed by Ross and by Gazdar (1981), given that non-head left branches are not sisters of a lexical head. Similarly, the Conjunct Condition of Ross' Coordinate Structure Constraint (precluding the extraction of only one of a set of conjuncts, as illustrated in (54a)) is also a consequence of the NONLOCAL feature inheritance principle and the fact that extracted items must be arguments of a lexical head, as illustrated in the impossibility of across-the-board extraction in (54b).

(54) a. *What table$_i$ did he buy $\_\__i$ and two chairs?

b. *What even positive integer$_i$ is four the sum of $\_\__i$ and $\_\__i$?

On the other hand, the Element Condition of Ross' Coordinate Structure Constraint, which permits only across–the–board extractions from within conjoined constituents, now seems not to be a syntactic constraint at all[56] in the face of such acceptable sentences as (55).

(55) a. Concerts that short you can leave work early, hear all of $\_\_$ and get back before anyone knows you're gone.

b. Don't tell me they drank all the whisky which I walked four miles to the store, and paid for $\_\_$ with my own money!

c. There was a new episode of *The Simpsons* on last Saturday, which I watched $\_\_$, and noted another bogus word.

If there were some reason to represent the Element Condition syntactically, it would just be the addition of the boldface clause in an independently needed Coordination Principle along the lines of (56):

(56) In a coordinate structure, the CATEGORY (**and NONLOCAL**) value of each conjunct daughter is an extension of that of the mother.

As for Ross' Complex NP Constraint (which was supposed to preclude gaps in noun complements and relative clauses as in (57), it has been known for decades that the noun-complement cases are often completely acceptable, as shown in (58).

[55] See Pollard & Sag 1994: 175, and Sag 1997: 446-447 for two approaches to achieving this effect.

[56] Cf. Goldsmith 1985 and Lakoff 1986. Postal 1998 has a syntactic account of "element extraction" which provides for the acceptability of examples like (55); Levine (To appear) argues that it makes incorrect predictions so that the class of prohibited element extractions, if non-empty, is even smaller than what Postal's account predicts.

(57) a. *Nelson, they quashed the report that the player choked —.

b. *Nelson, they quashed the report which Kim gave to —.

(58) That coach, they heard a claim that someone choked —.

Consequently, any constraints on them are pragmatic, not syntactic, in nature. In fact, the HPSG treatment of gaps predicts that in the general case "extractions" from the clausal complements of nouns will be syntactically well-formed, since the finite clause is just a complement of the noun (*fact, proposal, idea...*), and nothing prevents extraction of arguments from complements. Relative clauses, on the other hand, do seem to be strictly and syntactically constrained not to contain gaps, and HPSG analyses in effect stipulate that in a relative clause, the SLASH value set is a singleton whose index matches that of the relative pronoun. In many analyses, subject relative clauses are analyzed as having in situ subjects, and therefore allow complement extractions, as illustrated in (59).

(59) Okra$_i$, I don't know ANYONE who likes $—_i$.

Pollard & Sag (1994) and Sag (1997) correctly predict (59) to be well–formed, since they take *who* here to be an in situ subject. However, Ginzburg & Sag (To appear, Chapter 8) and Bouma et al. (1998), following Hukari & Levine 1995, 1996, argue that even highest subject WH-phrases should be treated as extracted, attributing the special properties of (59) to the indefinite head of the apparent relative clause; with a definite head, this construction is much less acceptable:

(60) ??Okra$_i$, I don't know THE CLIENT who likes $—_i$.

In addition, some sort of Sentential Subject Condition seems to be required to exclude gaps in clausal subjects. In fact, it is not just clausal subjects which prohibit gaps in subject position: all subject phrases headed by verbs or complementizers (*verbals* in Sag (1997)) display the same property, and the same property holds for gerundive NPs:

(61) a. *Lou, to argue with — makes me sick.

b. *Lou, that Terry argued with — irritated everyone.

c. *Who do you think arguing with — would infuriate Terry?

d. ?Which of the political candidates do you think that [my arguing with —] could be productive]?

How to define a natural class comprising these structures remains an open question.

11 Binding

The HPSG account of binding phenomena, treated in more detail in Levine & Pollard (this volume) starts from the premise that theoretical inconsistencies and documented counterexamples to familiar binding theories require that sentence–internal dependencies between referential noun phrases and coreferential reflexives, reciprocals, and personal pronouns be stated in a way which does not make reference to syntactic configurations. The HPSG account was developed to account for the following facts:

- Anaphoric personal pronouns ("pronominals") and reflexives ("anaphors") are in complementary distribution, to the extent that reflexives must have a clausemate antecedent, and pronouns may not have one.

- However:
  - Reflexives in picture–NPs can have antecedents in higher clauses. (*John$_i$ thought that pictures of himself$_i$ would make a good gift.*)
  - Reflexives in picture–NPs can have antecedents outside the sentence. (*John$_i$ thought about the situation. Pictures of himself$_i$ would make a good gift.*)
  - Reflexives with plural reference can be bound to noun phrases which jointly do not form a syntactic constituent. (*Kim$_i$ told Sandy$_j$ that there were pictures of themselves$_{i,j}$ on display.*)

The HPSG account is framed in terms of constraints on relations between coindexed subcategorized arguments of a head (i.e., between SYNSEM objects on an argument–structure list which have the same INDEX value). On Pollard & Sag's account, extraction gaps will have the same index–sort as any filler they are bound to, since what is structure–shared in the unbounded extraction dependency is a LOCAL value, and the value for INDEX is a part of a LOCAL value. The HPSG account of binding is stated in terms of obliqueness–binding (o–binding), which is dependent on the notion of obliqueness-command (o–command), defined in terms of the obliqueness relation which orders the SYNSEMS on an argument–structure list.

> For SYNSEM objects Y and Z, Y is **less oblique than** Z iff Y precedes Z on the ARG–ST value of a lexical head.
>
> For SYNSEM objects Y and Z with distinct LOCAL values, and Y referential, Y **locally o–commands** Z iff
>
> - Y is less oblique than Z,
> - or Y locally o–commands some X that subcategorizes for Z.
>
> For SYNSEM objects Y and Z, with distinct LOCAL values, Y referential, Y **o–commands** Z iff
>
> - Y is less oblique than Z,
> - or Y o–commands some X that subcategorizes for Z,
> - or Y o–commands some X whose HEAD value is token-identical to that of Z.
>
> Y **(locally) o–binds** Z iff
>
> - Y and Z have the same index
> - and Y (locally) o–commands Z.
>
> Z is **(locally) o–free** if Z is not (locally) o–bound.

Despite its non-configurational basis, the HPSG binding theory has a familiar look:

A. A locally o–commanded anaphor must be locally o–bound.

B. A personal pronoun must be locally o–free.

C. A non-pronoun must be o–free.

Nonetheless, the obliqueness account differs crucially from typical configurational accounts in that it has an inherently narrower scope. Principle A does not constrain ALL anaphors to be locally o–bound (that is, coindexed to something before them on an argument–structure list), but only those which are locally o–commanded (i.e., those which are non-initial on the list). This makes strong, vulnerable, and apparently correct claims. First, pronouns which are initial elements on argument–structure lists are unconstrained—free to be anaphors, coindexed to anything, and vacuously satisfying Principle A, or to be pronouns, substantively satisfying Principle B. Thus, the theory predicts that pronominal objects in these "exempt" conditions which are coindexed to anything anywhere in a higher clause, or outside the sentence altogether, can be either anaphors or pronouns. The following kinds of phrases are thus exempt:

- pre-nominal possessives (These are determiners, with CONTENT values equivalent to NPs, but they are the first or even the unique items on the argument structure lists of nominal heads, so they are not subject to Principle A, since they are not locally o–commanded.)

  (62) Bush and Dukakis charged that Noriega had contributed to each other's campaigns.

- objects of (prepositions in) Picture NPs (These are also a unique item on an argument structure list, and so not locally-o–commanded.)

  (63) a. The children$_i$ thought that pictures of themselves$_i$ were on sale.
  b. I suggested that portraits of themselves$_i$ would amuse the twins$_i$.
  c. John$_i$ knew there was a picture of himself$_i$ in the post office.

- objects, when the subject is expletive (The would-be o–commander is not referential, but o–command is not defined for non-referential sorts, therefore the next item on the list is not locally o–commanded.)

  (64) a. They$_i$ made sure that it was clear to each other$_i$ why Kim had to go.
  b. John$_i$ knew that there was only himself$_i$ left.

- accusative subjects (As subjects, they are not locally-o–commanded. Therefore they are exempt, and can be anaphors.)

  (65) a. John$_i$ wanted more than anything for himself$_i$ to get the job.
  b. What John$_i$ would prefer is for himself$_i$ to get the job.

This is correct; these reflexives that contradict the naive versions of Principle A are generally replaceable with pronouns with the same reference.

Second, because non-predicative ("case-marking") prepositions have a CONTENT value which is structure-shared with that of their object (since the preposition makes no contribution to the meaning), the prepositional phrase has a CONTENT value of the same sort as its object, and constrained by the binding theory just as if it were an NP. Thus, in contrast to a configurational binding theory, they pose no problem; when its nominative and accusative NPs are coindexed with each other, *depends on* requires an anaphoric accusative and disallows a prounoun, just as *trust* does.

(66) a. John$_i$ depends on himself$_i$ to get things done.

b. *John$_i$ depends on him$_i$ to get things done

(67) a. John$_i$ trusts himself$_i$ to get things done.

b. *John$_i$ trusts him$_i$ to get things done.

Third, a pronoun or anaphor cannot have a nonpronominal "antecedent" in a lower clause because the coindexing would put the nonpronominal in violation of the HPSG Principle C.

(68) *They told him$_i$ that John$_i$ would get things done.

Fourth, the analysis of extraction gaps predicts that the missing element is of the same sort as the filler, and therefore predicts that (69a) is a Principle C violation, while (69b) is not.

(69) a. *John$_i$, he$_i$ said you like $\_\_{}_i$.

b. Him$_i$, he$_i$ said you like $\_\_{}_i$.

Finally, the HPSG account of binding phenomena predicts that with multiple complements of the same verb, more oblique arguments cannot bind less oblique ones, regardless of their relative phrase order, so that (70a) and (70b) are correctly predicted to be unacceptable since the anaphor goal phrase is less oblique than the non-pronominal *about*-phrase.[57]

(70) a. *Marie talked about John$_i$ to himself$_i$.

b. *Marie talked to himself$_i$ about John$_i$.

12 Further directions

HPSG has proven attractive to researchers (both scientists and engineers) seeking to harness the systematic knowledge of natural languages in applications such as automated translation assistants and natural language interfaces to a wide variety of electronically stored databases. This chapter has tried to provide an introductory survey of the domain that HPSG aims to give an account of, and the major strategies used in that endeavor.

Sag & Wasow (1999) provides a tutorial on the spirit and the mechanics of HPSG, and is accessible to anyone with a minimal background in linguistics. Sag (1997) provides a comprehensive treatment of the syntax and semantics of English relative clauses; Ginzburg and Sag (To appear) gives an even more in-depth treatment of English interrogative constructions. The chapter by Kathol in this volume addresses problems of binding, extraction, and linear order in more detail than it has been possible to provide here. Levine & Green (1999) collects analyses of a variety of phenomena in English, German, and Japanese. Nerbonne, Netter & Pollard (1994) is exclusively HPSG analyses of a variety of phenomena in German, while Balari & Dini (1998) collects analyses of several phenomena in Romance languages, and Webelhuth, Koenig, & Kathol (1999) offers detailed analyses of phenomena in both Western and non-Western languages.

[57] For the same reason, this account of binding correctly predicts (i) to be acceptable, but fails to predict that (ii) is unacceptable.

i. I talked to John$_i$ about himself$_i$.

ii. *I talked about himself$_i$ to John$_i$.

Appendix A: Some basic sorts[58]

SORT	CONSTRAINTS	IS-A
sign	$\begin{bmatrix} \text{PHON} & \textit{list(phonological string)} \\ \text{SYNSEM} & \textit{canonical-synsem} \end{bmatrix}$	*ftr-str*
canonical-synsem		*synsem*
gap-synsem	$\begin{bmatrix} \text{LOCAL} & \boxed{1} \\ \text{SLASH} & \left\{\boxed{1}\right\} \end{bmatrix}$	*synsem*
PRO-synsem	$\begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} \mid \text{HEAD} & \begin{bmatrix} \textit{n} \\ \text{CASE} & \textit{acc} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \textit{reflexive} \\ \text{INDEX} & \textit{referential} \end{bmatrix} \end{bmatrix} \end{bmatrix}$	*synsem*
word	$\begin{bmatrix} \text{SYNSEM} \mid \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{ARG-ST} & \boxed{1} \oplus \boxed{2} \oplus \boxed{3} \\ \text{SUBJ} & \boxed{1} \\ \text{SPR} & \boxed{2} \\ \text{COMPS} & \boxed{3} \end{bmatrix} \\ \text{BIND} & \textit{set(local)} \end{bmatrix} \end{bmatrix}$	*sign*
phrase	$\begin{bmatrix} \text{HEAD-DTR} & \textit{sign} \\ \text{NON-HD-DTRS} & \textit{list(sign)} \end{bmatrix}$	*sign*
synsem	$\begin{bmatrix} \text{LOCAL} & \textit{local} \\ \text{NONLOCAL} & \textit{nonlocal} \end{bmatrix}$	*ftr-str*
local	$\begin{bmatrix} \text{CATEGORY} & \textit{category} \\ \text{CONT} & \textit{content} \\ \text{CONTEXT} & \textit{context} \end{bmatrix}$	*ftr-str*

[58] The sequence union of lists $\boxed{m}$ and $\boxed{n}$, $\boxed{m} \oplus \boxed{n}$, is the list consisting of $\boxed{n}$ appended to $\boxed{m}$.

nonlocal	$\begin{bmatrix}\textsc{slash} & \textit{set(local)}\\ \textsc{rel} & \textit{set(index)}\\ \textsc{que} & \textit{set(content)}\end{bmatrix}$	*ftr-str*
category	$\begin{bmatrix}\textsc{head} & \textit{part-of-speech}\\ \textsc{subj} & \textit{list(synsem)}\\ \textsc{spr} & \textit{list(synsem)}\\ \textsc{comps} & \textit{list(synsem)}\end{bmatrix}$	*ftr-str*
part-of-speech	$\begin{bmatrix}\textsc{mod} & \textit{none} \vee \textit{synsem}\end{bmatrix}$	*ftr-str*
v	$\begin{bmatrix}\textsc{vform} & \textit{vform}\\ \textsc{aux} & \textit{boolean}\\ \textsc{inv} & \textit{boolean}\end{bmatrix}$	*part-of-speech*
vform		*ftr-str*
n	$\begin{bmatrix}\textsc{case} & \textit{case}\end{bmatrix}$	*part-of-speech*
fin, bse, prp, psp, psv		*vform*
nom, acc		*case*
content	$\begin{bmatrix}\textsc{mode} & \textit{mode}\\ \textsc{index} & \textit{index}\\ \textsc{restriction} & \textit{set(predication)}\end{bmatrix}$	*ftr-str*
index	$\begin{bmatrix}\textsc{person} & \textit{person}\\ \textsc{number} & \textit{number}\\ \textsc{gender} & \textit{gender}\end{bmatrix}$	*ftr-str*
1st, 2nd, 3rd		*person*
sg, pl		*number*
fem, masc, neut		*gender*
prop, dir, int, reference		*mode*
referential, expletive		*index*
indiv-ind, sit-ind		*referential*
it, there		*expletive*

Appendix B: Some phrasal types[59]

SORT	CONSTRAINTS	IS-A
headed-phrase	$\begin{bmatrix} \text{SYNSEM}\|\text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{4} \\ \text{SUBJ} & /\ \boxed{1} \\ \text{SPR} & /\ \boxed{2} \\ \text{COMPS} & /\ \boxed{3} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{MODE} & \boxed{5} \\ \text{INDEX} & \boxed{6} \\ \text{RESTR} & \boxed{7} \cup \boxed{8} \ldots \boxed{n} \end{bmatrix} \end{bmatrix} \\ \text{HEAD-DTR}\|\text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{4} \\ \text{SUBJ} & /\ \boxed{1} \\ \text{SPR} & /\ \boxed{2} \\ \text{COMPS} & /\ \boxed{3} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{MODE} & \boxed{5} \\ \text{INDEX} & \boxed{6} \\ \text{RESTR} & \boxed{7} \end{bmatrix} \end{bmatrix} \\ \text{NON-HD–DTRS} & \left\langle \begin{bmatrix} \ldots\text{RESTR} & \boxed{8} \end{bmatrix}, \ldots \begin{bmatrix} \ldots\text{RESTR} & \boxed{n} \end{bmatrix} \right\rangle \end{bmatrix}$	*phrase*
head-nexus-ph	$\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOCAL}\|\text{CONTENT} & \boxed{1} \\ \text{NONLOCAL} & /\ \boxed{2} \end{bmatrix} \\ \text{HEAD-DTR}\|\text{SYNSEM} & \begin{bmatrix} \text{LOCAL}\|\text{CONTENT} & \boxed{1} \\ \text{NONLOCAL} & /\ \boxed{2} \end{bmatrix} \end{bmatrix}$	*headed-phrase*
head-comps-ph	$\begin{bmatrix} \text{HEAD-DTR} & \textit{word} \end{bmatrix}$	*head-nexus-ph*
head-su-ph	$\begin{bmatrix} \text{SUBJ} & \langle\rangle \\ \text{HEAD-DTR} & \begin{bmatrix} \text{SUBJ} & \left\langle \begin{bmatrix} \ \end{bmatrix} \right\rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \end{bmatrix}$	*head-nexus-ph*

[59]The right-leaning slash in a default specification has the interpretation 'unless otherwise specified, has the following value.'
The symbol $\uplus$ represents disjoint set union, which is just like the familiar set union, except that it is only defined for sets with an empty intersection.

head-filler-ph	$\left[\begin{array}{l} \text{SYNSEM}\mid\text{NONLOCAL}\mid\text{SLASH } \boxed{2} \uplus \boxed{3} \\ \text{HEAD-DTR} \left[\begin{array}{l} \text{LOCAL} \left[\begin{array}{l} \text{HEAD } v \\ \text{SUBJ } \langle\ \rangle \\ \text{COMPS } \langle\ \rangle \end{array}\right] \\ \text{NONLOCAL } \left[\text{SLASH } \boxed{2} \uplus \left\{\boxed{1}\right\}\right] \end{array}\right] \\ \text{NON-HD–DTRS} \left\langle \left[\text{SYNSEM} \left[\begin{array}{l} \text{LOCAL } \boxed{1} \\ \text{NONLOCAL}\mid\text{SLASH } \boxed{3} \end{array}\right]\right] \right\rangle \end{array}\right]$	*head-nexus-ph*
head-adjunct-ph	$\left[\begin{array}{l} \text{HEAD-DTR } \left[\text{SYNSEM } \boxed{1}\right] \\ \text{NON-HD-DTRS} \left\langle \left[\text{HEAD } \left[\text{MOD } \boxed{1}\right]\right] \right\rangle \end{array}\right]$	*headed-ph*

REFERENCES

Abeillé, Anne, Danièle Godard, and Ivan Sag. 1999. Two kinds of composition in French complex predicates. *Complex predicates in nonderivational syntax*, ed. by Erhard Hinrichs, Tsuneko Nakazawa, and Andreas Kathol. New York: Academic Press

Andrews, Avery. 1982. The representation of case in modern Icelandic. *The mental representation of grammatical relations*, ed. by Joan Bresnan. Cambridge, MA: MIT Press.

Balari, Sergio, and Luca Dini. 1998. *Romance in HPSG.* CSLI Publications: Stanford, CA.

Borsley, Robert. 1986. A note on HPSG. *Bangor Research Papers in Linguistics* 1: 77-85.

Borsley, Robert. 1989. An HPSG approach to Welsh. *Journal of Linguistics* 25: 333-354.

Bouma, Gosse, Rob Malouf, and Ivan A. Sag. (ms.) Satisfying constraints on extraction and adjunction.

Bresnan, Joan. 1977. Variables in the theory of transformations. *Formal Syntax*, ed. by Peter Culicover, Thomas Wasow, and Adrian Akmajian, 157-196. NY: Academic Press.

Clements, George N., James McCloskey, Joan Maling, and Annie Zaenen. 1983. String-vacuous rule application. *Linguistic Inquiry* 14: 1-17.

Chomsky, Noam. 1964. The logical basis of linguistic theory. *Proceedings of the Ninth International Congress of Linguists*, ed. by H. Lunt. The Hague: Mouton.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin and use.* NY: Praeger.

Cote, Sharon. 1996. Grammatical and discourse properties of null arguments in English. Ph.D. diss., University of Pennsylvania.

Copestake, Ann, Dan Flickinger, Carl Pollard, and Ivan A. Sag. To appear. (Ms.) Minimal recursion semantics: an introduction.

Culicover, Peter W. 1993. Evidence against ECP accounts of the *that*-t effect. *Linguistic Inquiry* 24: 557-561.

Davis, Anthony. 1996. *Lexical semantics and linking in the hierarchical lexicon.* Ph.D. Dissertation, Stanford University.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language*, 67: 547-619.

Dowty, David. 1996. Towards a minimalist theory of syntactic structure. it Discontinuous constituency, ed by Harry Bunt and Arthur van Horck, 11-62. Berlin: Mouton de Gruyter.

Gazdar, Gerald. 1979. *Pragmatics, implicature, presupposition, and logical form.* NY: Academic Press.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12: 155-184.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag. 1985. *Generalized phrase-structure grammar.* Cambridge, MA: Harvard University Press.

Gazdar, Gerald, and Geoffrey K. Pullum. 1981. Subcategorization, constituent order, and the notion 'head.' *The scope of lexical rules*, ed. by Michael Moortgat, H. v.d. Hulst, and T. Hoekstra, 107-123. Dordrecht: Foris.

Ginzburg, Jonathan, and Ivan A. Sag. To appear. *English interrogative constructions.* CSLI Publications.

Goldsmith, John. 1985. A principled exception to the Coordinate Structure Constraint. *Papers from the 21st regional meeting of the Chicago Linguistic Society*, ed. by William Eilfort, Paul Kroeber, and Karen Peterson. Chicago: Chicago Linguistic Society.

Green, Georgia M. 1974. *Semantics and syntactic regularity.* Bloomington, IN: Indiana University Press.

Green, Georgia M. 1995. The structure of CONTEXT. The representation of pragmatic restrictions in HPSG. *Proceedings of the Fifth Annual Conference, Formal Linguistics Society of Mid-America*, edited by James H. Yoon. *Studies in the Linguistic Sciences* 24: 215-232. Urbana, Illinois: Department of Linguistics, University of Illinois.

Green, Georgia M. 1997. Modelling grammar growth. *Language acquisition: Knowledge representation and processing. Proceedings of GALA '97.* University of Edinburgh.

Green, Georgia M. In prep. expanded version of Modelling grammar growth. To be submitted to Cognitive Science.

Green, Georgia M., and Robert D. Levine. 1999. Introduction. *Studies in contemporary phrase structure grammar.* Cambridge University Press.

Grosu, Alexander. 1972. The strategic content of island constraints. Ohio State University Working Papers in Linguistics 13: 1-225.

Haegeman, Lilliane. 1984. parasitic gaps and adverbial clauses. *Journal of Linguistics* 20: 229–232.

Hukari, Thomas, and Robert Levine. 1987. Rethinking connectivity in unbounded dependency constructions. *Proceedings of the Sixth West Coast Conference on Formal Linguistics*, ed. Megan Crowhurst. Stanford: Stanford Linguistics Association.

Hukari, Thomas, and Robert Levine. 1991. On the disunity of unbounded dependency constructions. *Natural Language and Linguistic Theory* 9: 97-144.

Hukari, Thomas, and Robert Levine. 1995. Adjunct extraction. *Journal of Linguistics* 31: 195–226.

Hukari, Thomas, and Robert Levine. 1996. Phrase-structure grammar: the next generation. *Journal of Linguistics* 32: 465–496.

Hukari, Thomas, and Robert Levine. Ms. Subject extraction. University of Victoria/Ohio State University; originally presented at the 1996 International Conference on HPSG, Marseilles.

Kathol, Andreas. 1995. Linearization-based German syntax. Ph.D. dissertation, Ohio State University.

Kasper, Robert. 1995. Semantics of recursive modification. Paper presented at workshop on HPSG, University of Tuebingen.

Keenan, Edward, and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63-99.

King, Paul. 1989. A logical formalism for head-driven phrase structure grammars. Doctoral dissertation, University of Manchester, Manchester.

Kornai, Andras, and Geoffrey K. Pullum. 1990. The X-bar theory of phrase structure. *Language* 66: 24-50.

Lakoff, George. 1986. Frame semantic control of the Coordinate Structure Constraint. *Papers from the 22nd regional meeting of the Chicago Linguistic Society (Papers from the Parasession)*, ed. by Anne M. Farley, Peter T. Farley and Karl-Erik McCullough, 152-167. Chicago: Chicago Linguistic Society.

Lascarides, Alex, and Ann Copestake. 1999. Default representations in constraint-based frameworks. *Computational Linguistics* 25: 55-106.

Levin, Beth. 1993. English verb classes and alternations; a preliminary investigation. Chicago: University of Chicago Press.

Levine, Robert. 1989. On focus inversion: syntactic valence and the role of a SUBCAT list. *Linguistics* 27: 1013-1055.

Levine, Robert D. (To appear) Review of Postal 1998. *Journal of Linguistics.*

Levine, Robert D., and Georgia M. Green. 1999. *Studies in contemporary phrase-structure grammar.* Cambridge University Press.

Levine, Robert D., Thomas E. Hukari, and Michael Calcagno. 1999. Parasitic gaps in English: Non-nominal, non-pronominal and case-consistent. Paper presented at HPSG'99, Edinburgh, Scotland.

Levine, Robert D., and Carl Pollard. (this volume) Unbounded dependencies and the theory of binding in HPSG.

Meurers, W. Detmar. 1995 Towards a semantics for lexical rules as used in HPSG. Paper presented at the Conference on Formal Grammar. Barcelona.

Meurers, W. Detmar, and Guido Minnen. 1997. A computational treatment of lexical rules in HPSG as covariation in lexical entries. *Computational Linguistics* 23: 543-568.

Miller, Philip, and Ivan Sag. 1997. French clitic movement without clitics or movement. *Natural language and linguistic theory*, 15: 573-639.

Nakazawa, Tsuneko, and Erhard Hinrichs. 1999. Partial VP fronting and split NP topicalization in German: an HPSG analysis. *Studies in contemporary phrase structure grammar*, Robert D. Levine and G. M. Green, eds. Cambridge University Press.

Nerbonne, John, Klaus Netter, and Carl Pollard, eds. 1994. *German in Head-driven Phrase Structure Grammar.* Lecture Notes Series, CSLI: Stanford, CA.

Pollard, Carl. 1999. Strong generative capacity in HPSG. Webelhuth, Gert, Jean-Pierre Koenig, and Andreas Kathol, eds., *Lexical and constructional aspects of linguistic explanation.* CSLI: Stanford, CA.

Pollard, Carl, and Drew Moshier. 1990. Unifying Partial descriptions of sets. *Information, Language and Cognition. Vancouver Studies in Cognitive Science*, Vol. 1: 285-322. Vancouver: University of British Columbia Press.

Pollard, Carl, and Ivan Sag. 1987. *Information-based syntax and semantics, Vol. 1.* Stanford University, Center for the Study of Language and Information.

Pollard, Carl, and Ivan Sag. 1994. *Head-driven phrase structure grammar.* Chicago: University of Chicago Press.

Postal, Paul M. 1994. Parasitic and pseudoparasitic gaps. *Linguistic Inquiry* 25: 63-117.

Postal, Paul M. 1998. *Three investigations of extraction.* Cambridge, MA: MIT Press.

Postal, Paul M., and Geoffrey K. Pullum. 1988. Expletive noun phrases in subcategorized positions. *Linguistic Inquiry* 19: 635-670.

Pullum, Geoffrey K. 1982a. Free word order and phrase structure rules. *Proceedings of the Twelfth Annual Meeting of the North Eastern Linguistic Society*, ed. by James Pustejovsky and Peter Sells, 209–220. Graduate Linguistics Student Association, University of Massachusetts at Amherst.

Pullum, Geoffrey K. 1982b. Syncategorematicity and English Infinitival *to. Glossa* 16: 181-215.

Pullum, Geoffrey K. 1985. Assuming some version of x-bar theory. *Papers from the 21st regional meeting of the Chicago Linguistic Society, General Session*, ed. by William H. Eilfort, Paul D. Kroeber, and Karen L. Peterson, 323-353. Chicago: Department of Linguistics, University of Chicago.

Reape, Michael. 1994. Domain union and word order variation in German. *German in Head-driven phrase structure grammar*, ed. by John Nerbonne, Klaus Netter, and Carl Pollard. No. 46 Lecture Notes. Stanford: CSLI Publications.

Reape, Michael. 1996. Getting things in order. *Discontinuous constituency*, ed. by Harry Bunt and Arthur Van Horck. Berlin: Mouton de Gruyter.

Ross, John R. 1967. Constraints on variables in syntax. Ph.D. dissertation, Massachusetts Institute of Technology. (Published 1983 as *Infinite Syntax.* Norwood, NJ: Ablex).

Sag, Ivan. 1997. English relative clause constructions. *Journal of Linguistics* 33: 431-485.

Sag, Ivan A., Lauri Karttunen, and Jeffrey Goldberg. 1992. A lexical analysis of Icelandic case. *Lexical matters*, ed by Ivan A. Sag and Anna Szabolcsi. CSLI Lecture Notes No. 24. Stanford: Center for the Study of Language and Information (distributed by the University of Chicago Press).

Sag, Ivan, and Thomas Wasow. 1999. *Syntactic theory: a formal introduction.* Stanford, CA: CSLI Publications.

Shieber, Stuart. 1986. *An introduction to unification-based approaches to grammar.* CSLI Lecture Notes Series. Stanford: Center for the Study of Language and Information. [Distributed by University of Chicago Press]

Webelhuth, Gert, Jean-Pierre Koenig, and Andreas Kathol, eds. 1999. *Lexical and constructional aspects of linguistic explanation.* CSLI: Stanford, CA.