8

Feature-Based Grammar

James P. Blevins

8.1   Introduction

This chapter considers some of the basic ideas about language and linguistic analysis that define the family of feature-based grammars. Underlying the complex constraint languages and representations that give this family its characteristic "look and feel" are a number of relatively straightforward claims and hypotheses. Foremost among these is the idea that many of the distinctive properties within a grammatical system can be described in terms of (morpho)syntactic features. A related claim is that many grammatical dependencies – both local and nonlocal – can be regulated by strategies that determine the compatibility of the feature information associated with grammatical dependents. A third claim is more formal than substantive and is formulated in somewhat different ways in different approaches. But the shared intuition is that the strategies that determine compability do not merely compare or "check" grammatical dependents to see if they have conflicting values for common features. Instead, the compatibility of two or more dependents is determined "constructively," by invoking principles that are satisfied only if there is an object that in some way combines the feature information associated with each of the dependents.

  These substantive claims interact with auxiliary assumptions and implementation choices in ways that define the different variants of feature-based frameworks. Specific inventories of features and values differ considerably across approaches, as do basic terminological and interpretive conventions. Traditional morphosyntactic properties, such as tense, aspect, agreement, and case, are usually represented by syntactic features, though there is less of a consensus regarding the treatment of phenomena such as voice alternations or word-order variation. The organization of features within a syntactic analysis also varies a great deal from one approach to the next, and tends to reflect general properties of a model, especially assumptions about the relation between features and constituent structure. In these respects, feature-based approaches can appear to comprise a family of approaches separated by a common metalanguage.

There is, however, more of a consensus regarding the formal strategies that determine feature compatibility within contemporary feature-based approaches. Both Lexical-Functional Grammar (LFG; Kaplan & Bresnan 1982; Dalrymple et al. 1995) and Head-Driven Phrase Structure Grammar (HPSG; Pollard & Sag 1987, 1992) adopt a model-theoretic or *description-based* perspective. This type of approach is distinguished from more traditional accounts by its rigid separation between linguistic *expressions* and the *objects* that those expressions describe. On the one side are grammatical rules, constraints, and lexical entries, which are treated as types of expressions. On the other side are feature structures and constituent structures, which are treated as types of linguistic objects. The formal properties of linguistic objects again vary across approaches, and these differences correlate with variation in the form of expressions and the nature of the relations between expressions and objects. But the central motivation for a description-based perspective is much the same in all approaches, and derives from the fact that the "satisfiability" of an expression or the mutual compatibility of a set of expressions can be determined by whether there is a well-formed object that is described by the expression or set of expressions. Treating notions like negation and disjunction – and possibly even "re-entrancy" – as properties of descriptions also greatly simplifies feature structures.[1]

From a broader linguistic perspective, feature-based grammars can be placed within the general post-Bloomfieldian tradition, representing a line of development parallel to the transformational tradition. The notions of constituent structure that make their way into feature-based and transformational approaches derive ultimately from the models of immediate constituent (IC) analysis developed by Bloomfield (1933) and his immediate successors. The models of constituent analysis outlined in Harris (1946), Wells (1947), and Hockett (1958: section 17), among others, were in many respects more sophisticated than the models that followed, particularly in their dissociation of hierarchical structure from linear arrangement and in their represention of suprasegmental properties such as intonation. But their treatment of other types of grammatical relations and dependencies are more rudimentary. Syntactic features have a highy restricted role and mainly serve to encode word class. Morphosyntactic properties tend to be encapsulated in abstract morphemes and morpheme classes, and there is no means of referring to features as independent components of a represention. The descriptive limitations of IC models are reflected in the treatment of concord and agreement as types of part–whole relations in which the dependents form a "discontinuous morpheme" (Harris 1951: section 82). Other local dependencies, such as case government and valence alternations, raise similar difficulties for models based almost exclusively on techniques of part–whole analysis.

The different lines of post-Bloomfieldian research differ largely in how they address these weaknesses. Early transformational accounts attempt to compensate for the limitations of individual IC analyses by relating complex structures to simpler structures. The transformational model developed by Harris (1957, 1965) is designed to reduce complex structures algebraically to "kernel" structures, which can be assigned an IC analysis. Chomsky (1957) pursues a similar intuition by deriving complex clauses from the structures that underlie kernel sentences, whereas Chomsky (1965) deals with the limitations of singule IC analysis by introducing derivations that consist of multiple IC analyses.

Feature-based models proceed from a different observation, namely that the descriptive limitations of IC analyses can be overcome by enriching the information associated with a single constituent analysis. This point is made initially by Harman (1963), who notes that the many of the limitations attributed to IC analyses are not intrinsic to part–whole analyses but artefacts of the way that these analyses are *formalized* in the model of phrase structure in Chomsky (1956). Two of the restrictions on phrase structure analysis are particularly relevant. First of all, the models proposed by Chomsky exclude discontinuous constituents, even though "most … conceptions of grammatical structure" developed to that point had "involve[d]

some notion of phrase structure with discontinuous elements" (Harman 1963: 96). Moreover, by allowing only simple non-terminal symbols such as "S," "NP," "VP," "N," "V," etc., "the amount of grammatical information made available by a grammar" is "restricted to information about the grammatical category of words and phrases" (Harman 1963: 94). The exclusion of discontinuous constituents deprives phrase structure analyses of the treatments of phrasal verbs, subject–auxiliary inversion, and other types of discontinuous dependencies that had been represented in IC analyses.

For the most part, the innovative aspects of feature-based models derive from the type of grammatical information that they associate with constituents in a syntactic analysis. By associating words and phrases with morphosyntactic properties, as well as information about valence and even filler–gap dependencies, feature-based models can regulate a wide range of local and nonlocal grammatical dependencies. The explicit way in which these models regulate grammatical dependencies also clarifies the scope and limits of feature-based strategies, while highlighting the general trade-off between the complexity of constituent structures and the complexity of feature information that is associated with the elements of those structures. By enriching the information in individual surface representations, feature-based approaches define classes of analyses that differ markedly from the more abstract structures characteristic of transformational accounts. Feature-based frameworks are also distinguished by their "lexicalist" orientation, in which grammatical properties are predominantly associated with "overt" lexical items or even – somewhat incongruously – with subword units. At the same time, their focus on details of *representations* places feature-based approaches squarely in the post-Bloomfieldian tradition, in contrast to traditional grammars, which organize syntactic systems more in terms of exemplary patterns and constructions.

The body of this chapter summarizes some of the strategies developed with the feature-based tradition, examines a number of choices that arise within these strategies, and considers the implications of particular choices. Sections 8.2 and 8.3 introduce features, feature structures, and feature-based mechanisms for regulating grammatical dependencies. Section 8.4 examines some empirical patterns that bear on the choice between strategies based on unification or structure-sharing and those based on a weaker subsumption relation. Section 8.5 concludes with a summary of a range of issues that serve to distinguish individual approaches. These include the treatment of locality, the formal interpretation of underspecification, and the relation between feature structures and constituency. Section 8.6 gives some concluding remarks.

8.2   Features and Values

It is useful at the outset to delimit the broad class of "feature-based" grammars. Some of the early approaches, such as Functional Unification Grammar (FUG: Kay 1979) and versions of the PATR formalism (Shieber 1986), have mainly been relevant for grammar implementations. The most theoretically oriented models include LFG, HPSG, and Generalized Phrase Structure Grammar (GPSG; Gazdar et al. 1985). Although these approaches have also provided a basis for practical implementations, they are formulated as general frameworks for broad-coverage linguistic description and theoretical analysis. A third set of approaches, which includes Ackerman and Webelhuth (1998) and Andrews and Manning (1999), attempt to combine properties of different feature-based models.

The foundation of all feature-based models is an inventory of feature *attributes* and feature *values* that describe the distinctive properties of a linguistic system.[2] The atomic values that represent individual properties are the simplest elements of these inventories and, in fact, the simplest types of *feature structures*. A property such as morphological case is typically represented by a CASE attribute with atomic values that might include NOM(INATIVE), ACC(USATIVE),

and GEN(ITIVE). Person properties are represented by a PER(SON) attribute with atomic values such as 1(st), 2(nd), 3(rd). Features with two possible values are often represented by the "boolean" values "+" and "−." However, nothing hinges on the choice between boolean and other types of atomic values unless a model incorporates a notion of "markedness" (Jakobson 1932, 1936) or otherwise distinguishes the interpretation of positive "+" and negative "−" values.[3] The fact that English nouns show a binary contrast between singular and plural is expressed by a boolean-valued PLU(RAL) feature in GPSG (Gazdar et al. 1985: 214). The same contrast is expressed by a NUM(BER) feature with the atomic values SG and PL in LFG (Kaplan & Bresnan, 1982: 177) and by the values SING(ULAR) and PLUR(AL) in HPSG (Pollard & Sag 1994: 397), even though LFG and HPSG both assign boolean values to other binary-valued features.

8.2.1   Complex-valued features

All syntactic models impose some structure on morphosyntactic properties, minimally organizing them into "bundles" of attributes and values. In most approaches, structured feature bundles are represented as attribute–value matrices (AVMs). AVMs represent a class of feature structure termed *categories* in GPSG and a slightly different class of structures termed *f(unctional)-structures* in LFG. HPSG also makes use of AVMs but, as in Blackburn (1994: 19), interprets AVMs as sets of *constraints*, not as feature structures.[4] The AVMs that represent the properties of the German pronouns *er* 'he' and *wir* 'we' in (1) illustrate simple feature structures with only atomic-valued attributes.

(1)   **Syntactic subjects and subject demands**:

$$
\text{er:} \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{MASC} \\ \text{CASE} & \text{NOM} \end{bmatrix} \cdot \quad \text{wir:} \begin{bmatrix} \text{PER} & 1 \\ \text{NUM} & \text{PL} \\ \text{CASE} & \text{NOM} \end{bmatrix} \quad \text{singt:} \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{TENSE PRES} \end{bmatrix}
$$

The most fundamental innovation within feature-based models is the use of attributes with "complex" or nonatomic structures as values. This extension is illustrated by the SUBJ(ECT) attribute associated with German *singt* 'sings' in (1). The value of the SUBJ attribute is not an atom, such as SG or NOM, but is itself a complex structure, consisting of features and atomic values. It is the fact that the value of the SUBJ attribute is a structure of the same kind as the structures associated with *er* and *wir* that permits a straightforward treatment of agreement. The subject demands of *singt* can be enforced by determining the compatibility between the SUBJ structure and the structures associated with syntactic subjects. The unacceptability of **Wir singt* '*We sings' correlates with the conflict between the plural NUM value of *wir* and the singular NUM value in the SUBJ value of *singt*. Conversely, the acceptability of *Er singt* 'He sings' is attributable to the lack of any conflict between the features of *er* and singt.

Other types of local demands can be represented in a similar way. The fact that German *hilft* 'helps' governs a dative object is expressed in (2) by associating *hilft* with an OBJ(ECT) attribute whose value contains a CASE attribute with a DAT(IVE) value. The unacceptability of the verb phrase **hilft ihn* 'helped him' then correlates with the conflict between the dative case associated with *hilft* and the accusative value assigned to *ihn*. The acceptability of *hilft ihm* 'helped him' correlates with the lack of any conflict between the features associated with the OBJ of the governor *hilft* and the dative features of the object *ihm*.

More generally, as the structures assigned to *singt* and *hilft* show, complex-valued features represent the valence demands of a predicate independent of any syntactic context.

Constraints can thus refer to the subject or object *demands* or *requirements* imposed by a verb, which permits a lexical description of patterns that tend to be classified as syntactic in transformational accounts.

(2) **Syntactic objects and case government**:

$$
\text{ihm:}
\begin{bmatrix}
\text{PER} & 3 \\
\text{NUM} & \text{SG} \\
\text{GEND} & \text{MASC} \\
\text{CASE} & \text{DAT}
\end{bmatrix}
\quad
\text{ihn:}
\begin{bmatrix}
\text{PER} & 3 \\
\text{NUM} & \text{SG} \\
\text{GEND} & \text{MASC} \\
\text{CASE} & \text{ACC}
\end{bmatrix}
\quad
\text{hilft:}
\begin{bmatrix}
\text{SUBJ} & \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\
\text{OBJ} & \begin{bmatrix} \text{CASE} & \text{DAT} \end{bmatrix} \\
\text{TENSE} & \text{PRES}
\end{bmatrix}
$$

8.2.2   Local dependencies

Feature-based treatments of 'raising' constructions show exactly how complex-valued features extend the scope of a lexical analysis. Since at least Jespersen (1937), it is conventional to recognize a class of raising verbs that take a predicative complement and a syntactic subject that is ultimately selected by the complement. The English verbs *seem, appear*, and *tend* are all canonical raising verbs in this sense, as their syntactic subjects reflect the demands of their infinitival complements. The role that the complement plays in dictating syntactic properties of the "raised" subject is particularly evident with constructions that select "exceptional" subjects, such as the expletive elements *there* or *it* or parts of an idiom, such as *tabs*. The observation that the subjects of raising verbs obey the selectional demands of their complements is illustrated in (3).

(3) **Preservation of exceptional subject selection in raising:**
   a. There is a transit strike in France. ~
      There seems to be a transit strike in France.
   b. It rains more in coastal regions. ~
      It tends to rain more in coastal regions.
   c. Tabs were kept on the dissidents. ~
      Tabs appear to have been kept on the dissidents.

   The term "raising" derives from transformational analyses in which the subjects in (3) are taken to originate as the subject of the predicative complement and are then "raised" to become the syntactic argument of the raising verb. However, complex-valued features permit an analysis in which raising involves the *sharing* of information within the argument structure of a raising predicate. While this type of analysis has been applied to the English examples considered above, patterns involving the sharing of purely morphological properties offer an even clearer illustration of the role of complex-valued features. As discussed by Andrews (1982), among others, modern Icelandic contains verbs that may govern "quirky" non-nominative subjects. One such verb is *vanta* 'to want', which occurs with the accusative subject *hana* 'her' in (4a). These quirky case demands are preserved by raising verbs such as *virðast* 'to seem'. As example (4b) shows, *virðast* is, in effect, "transparent" to the accusative case demands of *vanta*, which are imposed on its own syntactic subject.

(4) **Quirky case in Icelandic raising constructions** (Andrews 1982):
   a. Hana   vantar   peninga.
      her.ACC lack.3SG money.ACC
      'She lacks money.'

b. Hana   virðist   vanta peninga.
   her.ACC seem.3SG lack   money.ACC
   'She seems to lack money.'

The transparency of *virðist* is represented in (5) by the pair of boxed integers. These "tags" indicate that the SUBJ attribute of *virðist* literally shares its value with the SUBJ value of its predicative complement. Identifying the values of the two SUBJ attributes ensures that any constraints that apply to the SUBJ of the complement of *virðist* will apply to its own syntactic SUBJ. The structure associated with *vanta* in (5) shows that *vanta* selects an accusative subject. Hence when *vanta* occurs as the complement in a phrase such as *virðist vanta peninga* 'seems to lack money', its accusative SUBJ demands will be identified with the SUBJ demands of *virðist*. An accusative subject, such as *hana* in (5), can then satisfy these demands, as in sentence (4b). But *hana* does not combine syntactically with the complement *vanta peninga* on this feature-based analysis. Instead, *virðist* inherits the demands of its complement, and imposes them in turn on its own syntactic subject.

(5) **Raising and quirky case government**:

$$
\text{hana:}
\begin{bmatrix}
\text{PER} & 3 \\
\text{NUM} & \text{SG} \\
\text{GEND} & \text{FEM} \\
\text{CASE} & \text{ACC}
\end{bmatrix}
\quad
\text{vanta:}
\begin{bmatrix}
\text{SUBJ} & [\text{CASE ACC}] \\
\text{OBJ} & [\text{CASE ACC}] \\
\text{TENSE} & -
\end{bmatrix}
\quad
\text{virðist:}
\begin{bmatrix}
\text{SUBJ} & \boxed{1} \\
\text{XCOMP} & \begin{bmatrix}\text{SUBJ} & \boxed{1} \\ \text{TENSE} & -\end{bmatrix} \\
\text{TENSE} & \text{PRES}
\end{bmatrix}
$$

As in the analyses of agreement and subcategorization, it is the use of complex-valued SUBJ attributes that permits feature-based models to identify the subject demands of a raising verb with those of its complement. As in previous analyses, the precise formal representation of shared values varies across individual approaches. The classification of predicative complements also tends to be highly theory-dependent. The structure in (5) follows LFG in treating infinitival complements as "unsaturated" XCOMP functions. However, nothing hinges on this choice, and the analysis in (5), is all relevant respects, identical to the HPSG analysis of the Icelandic constructions in Sag et al. (1992).

8.2.3   Nonlocal dependencies

As shown initially by Gazdar (1981), feature-based strategies for regulating local dependencies can be extended to accommodate potentially "unbounded" dependencies by breaking nonlocal dependencies into a sequence of local dependencies. By expressing information about a "gapped" element in an extraction construction as the value of a complex-valued "slash" feature, GPSG and HPSG accounts are able to match dislocated "fillers" with missing "gaps". This analysis can be illustrated with reference to the simple embedded question in (6), in which the initial question word *what* functions as the direct object of *saw*.

(6)   They wonder [what$_i$ Max saw $\_i$?]

In the analysis in (7), the filler *what* is linked to the "gap site" by a chain of SLASH attributes. At one end of the chain, a preterminal node dominating a "gap" is matched against the valence demands imposed by *saw*. At the other end, the value of the SLASH attribute is identified with the structure associated with the filler *what*. The intervening nodes have identical values for the SLASH attribute, ensuring that demands imposed at the gap site are applied to the filler.

(7) **Slash-category analysis of extraction** (Gazdar et al. 1985):

```
                     S
              ┌──────┴──────┐
             NP          S[SLASH [NP]]
              |       ┌──────┴──────┐
            what     NP         VP[SLASH [NP]]
                      |        ┌─────┴─────┐
                     Max       V       NP[SLASH [NP]]
                               |           |
                              saw          e
```
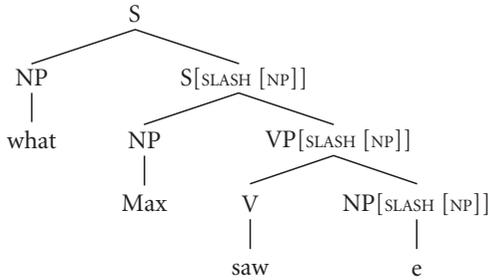
Early versions of HPSG and LFG similarly use chains of category-valued SLASH attributes or "bounded metavariables" (Kaplan & Bresnan 1982) to link fillers and gaps. Subsequent HPSG analyses (Pollard & Sag 1994; Sag & Fodor 1994) refine this analysis by introducing additional complex-valued attributes and by eliminating the null preterminal *e*. LFG accounts formulated in terms of "functional uncertainty" (Kaplan & Zaenen 1989) shift the locus of unbounded dependencies from c(onstituent)-structures similar to that in (7) to f(unctional)-structures of the sort illustrated in (8). The key representational innovation in this structure is the information-structure attribute FOCUS, which shares a value with the OBJ attribute. The FOCUS attribute in (8) is parasitic on the governed OBJ attribute elsewhere in the structure in much the way that the dislocated filler is dependent on the gap site in (7).[5]

(8) **f-structure representation of unbounded dependencies**:

$$
\begin{bmatrix}
\text{FOCUS} & \boxed{1}\begin{bmatrix} \text{PRON} & \text{WH} \\ \text{PRE} & \text{'PRO'} \end{bmatrix} \\
\text{TENSE} & \text{PAST} \\
\text{SUBJ} & \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{PRED} & \text{'MAX'} \end{bmatrix} \\
\text{OBJ} & \boxed{1} \\
\text{PRED} & \text{'SEE} \langle \text{SUBJ OBJ '} \rangle
\end{bmatrix}
$$
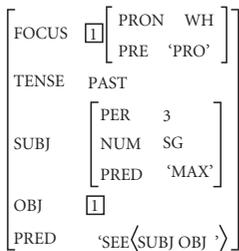
Associating fillers and gaps in f-structures rather than c-structures permits a simplification of the constituent analyses assigned to unbounded dependencies in LFG. Like IC analyses and the phrase structure trees of early transformational accounts, c-structures represent little more than word class, linear order, and constituent structure. Yet unlike in IC analyses, the part–whole relations represented by c-structures are not grammatically significant except insofar as they determine constituent order or properties of an an associated f-structure. The resulting division of labour is illustrated by the paired analyses in (9), in which the correspondence between c-structure nodes and f-structure elements is expressed by the indices f_1, f_2, and f_3. The index f_2 associates the filler *what* with the value of the FOCUS attribute, and f_3 associates the subject NP *Max* with the value of the SUBJ attribute. The index f_1 associates the verb, the verb phrase, and the clausal projections with the entire f-structure in (9).

(9) **Associated c- and f-structure analyses of unbounded dependency:**

$$
\begin{array}{c}
\text{S}_{f1} \\
\end{array}
\qquad
\begin{bmatrix}
\text{FOCUS} & f_2 \begin{bmatrix} \text{PRON} & \text{WH} \\ \text{PRED} & \text{'PRO'} \end{bmatrix} \\
\text{TENSE PAST} \\
\text{SUBJ} & f_3 \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{PRED} & \text{'MAX'} \end{bmatrix} \\
\text{OBJ} & f_2 \\
\text{PRED} & \text{'SEE} \left\langle \text{SUBJ OBJ} \right\rangle \text{'}
\end{bmatrix}
$$

Tree structure:
- S_{f1}
  - NP_{f2} — what
  - S_{f1}
    - NP_{f3} — Max
    - VP_{f1}
      - V_{f1} — saw
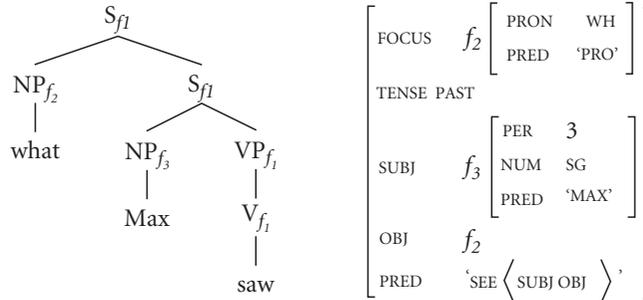
   Despite evident differences in execution, the analyses in (7) and (9) represent variations on a common strategy that uses complex-valued features to link fillers and gaps. The contrasts between the analyses principally reflect different views of the relation between constituent structure and feature structure. The GPSG analysis in (7) introduces feature information in the labels that annotate the nodes of a phrase structure tree. The LFG analysis in (9) instead consolidates feature information into a separate structure, whose parts correspond to the nodes of a c-structure tree. HPSG accounts develop a third strategy, which, in effect, inverts the GPSG organization. Rather than treating tree structures as grammatical "skeletons" that are annotated with feature information, HPSG treats feature structures as basic and expresses constituency relations within feature structures by means of DAUGHTERS attributes that take SIGNS – that is, structures representing subconstituents – as values.

8.2.4   Features, categories, and constituency

The analysis of unbounded dependencies also brings out the way that feature-based analyses tend to enrich the feature information associated with syntactic representations, while retaining the simple model of constituent structure from early phrase structure grammars. In the case of phenomena such as government or agreement, complex-valued features appear to offer an advantage over constituency-based analyses that admit discontinuous morphemes (Harris 1951) or invoke operations like "affix hopping" (Chomsky 1957). Yet, in other cases, notably those involving discontinuous dependencies, there is no principled reason why pairs of dependents should be linked by complex-valued features rather than by constituency relations. The preference for feature-based analyses comes down ultimately to ease of formalization or implementation. Feature-based models have formal techniques for linking the feature information associated with non-adjacent constituents, but lack comparably formalized strategies for extending constituency relations over larger domains.[6]

   In this repect, feature-based approaches are something of a mirror image of earlier Bloomfieldian models. Lacking a means of representing feature information directly, Bloomfieldian models tended to "overload" constituency relations. Nevertheless, the flexible model of constituency developed within this tradition permitted the assignment of IC analyses to be guided by empirical considerations, rather than dictated by constraints on a grammatical formalism. The benefits of this flexibility are particularly clear in connection with IC analyses of phrasal verbs and other types of complex predicate. As Wells (1947) argues, a phrasal verb such as *let out* is a grammatical unit, whether its parts occur contiguously, as in *let out the cat*, or are separated by another element, as in *let the cat out*. Hockett (1958) represents the general view of his contemporaries when he suggests that polar questions have the same constituent analysis as the corresponding declaratives, but are distinguished by their linear arrangement.

> On the other hand, two sentences may involve exactly the same constituents at all hierarchical levels, and yet differ in meaning because of different patterns … The difference [between *John is here* and *Is John here*] lies not in constituents, but in their arrangement: *John* respectively before or within *is here*. (Hockett 1958: 158)

The model of IC analysis suggested in Gleason (1955: 142) would likewise treat the "filler" *what* in (7) and (9) as the surface object of the verb *saw*. Most feature-based models are unable to treat non-adjacent elements as surface constituents because, like transformational accounts, they adopt a model of constituent analysis that derives from phrase structure grammars (rather than from the models of IC analyses that phrase structure grammars were meant to formalize). There is no evidence that the constraints on constituent analyses assumed by feature-based models have any psychological relevance. In particular, there is no reason to believe that speakers have any more difficulty recognizing *is … here* or *put … out* as syntactic units in *Is John here?* or *let the cat out* than they do in treating *un … likely* as a morphological unit in *un-bloody-likely*.

The treatment of unbounded dependencies illustrates more general points about feature-based approaches. On the one hand, these analyses show that complex-valued features can be used to relate grammatical dependencies over a potentially unbounded domain, so that the existence of nonlocal dependencies does not establish the need for transformations or any type of derivational mechanism. On the other hand, these analyses highlight the influence that transformational accounts have exerted on feature-based approaches. This influence is particularly clear in the way that early GPSG and LFG analysis adopted the "operator–variable" analysis from the "Extended Standard Theory" (Chomsky 1977), and merely employed different devices to link operators/fillers with variables/gaps.[7] From constraints on the class of constituent structures through to analyses of individual constructions, assumptions native to transformational approaches have molded the development of feature-based formalisms.

8.3   Feature Compatibility

The analyses in section 8.2 have shown how complex-valued features can act as repositories of grammatical information. This section considers the strategies for regulating dependencies between repositories, which constitute the second key component of feature-based models. The basic idea expressed by these strategies is that grammatical dependents must be *compatible*, and that compatibility mainly comes down to the lack of conflicting atomic-valued features. This notion of compatibility can be determined in a number of different ways. At one extreme are strategies that determine the compatibility of multiple structures by *unifying* them (or, equivalently, by treating multiple descriptions as descriptions of the *same* structure). These unification-based (or description-based) strategies can be said to be *destructive*, because the compatibility of multiple structures is established by the existence of a unified structure that does not record the individual "contribution" of the structures whose compatibility was being determined. By consolidating information from different sources, destructive strategies induce what can be informally described as a "flow" of information within a representation. This "information flow" allows the principles that govern grammatical dependencies to be stated over a local domain, without the mediation of constituent structure displacements. At the other extreme are "checking" strategies that inspect structures to verify whether or not they contain conflicting atomic values. These strategies are often described as *nondestructive* because the compatibilty check that they perform does not combine the input structures into a structure that amalgamates their features, nor does it alter the inputs in any way. Because "checking" strategies do not modify the properties of checked structures, they are proposed in the analysis of constructions in which a single element appears to be satisfying multiple,

incompatible demands. Yet because checking does not induce information flow, it cannot be used to regulate dependencies over a nonlocal domain.

Between these positions lies a third possibility, which combines the complementary virtues of destructive and nondestructive strategies. The information flow induced by destructive strategies comes from combining the information of compatible "inputs" in an amalgamated "output." The usefulness of checking strategies arises in contexts, such as coordinate structures, where an underspecified element is simultaneously subject to multiple incompatible demands. However, it is possible to induce information flow without sacrificing or "resolving" the neutrality of input structures. The compatibility of input structures $s_1 \ldots s_n$ can established by the existence of a SEPARATE structure S that "pools" the features in the inputs without "overwriting" them. More precisely, the compatibility of a set of structures can be determined by a "semi-destructive" strategy that merely requires compatible structures to subsume a common structure. This common structure will often correspond to the mother of the "inputs." A subsumption constraint will determine the same "information flow" as unification, but without the problematic "side effect" of folding the inputs into the consolidated output.

8.3.1   Unification

The importance of unification to feature-based models such as FUG (Kay 1979) and versions of the PATR formalism (Karttunen 1984; Shieber 1986) is reflected in the description "unification-based," which is now somewhat deprecated among proponents of feature-based accounts. Shieber (1986) provides a particularly straightforward definition of feature-structure unification in terms of a *subsumption* or "relative informativeness" relation. Shieber begins by specifying subsumption relations for two types of simple feature structures: *variables* (or 'empty' structures), represented '[]', and *atomic structures* like 3, PL or ACC. "An atomic feature structure neither subsumes nor is subsumed by a different atomic feature structure. Variables subsume all other feature structures, atomic or complex, because, as the trivial case, they contain no information at all" (Shieber 1986: 15).

These simple structures provide the base for a general subsumption relation "⊑," which imposes a partial informativeness order on arbitrary feature structures S and T. In Shieber's formulation, feature structures are treated as partial functions from features to values, so that the expression "$S(f)$" denotes the value that a structure S assigns to a feature f. Similarly, $dom(S)$ denotes the *domain* of features to which a structure S assigns a value. The expression "$S(p)$" denotes the value assigned a sequence or *path* of attributes. Applying a feature structure S to a path (fg) provides a convenient reference to the value obtained by applying S successively to f and g. Thus applying the structure S_3 in (8) to the sequence (SUBJ NUM) denotes the value obtained by applying S_3 to SUBJ and applying the resulting function S_3 (SUBJ) to the attribute num.

The general definition of subsumption in (10) below imposes two very different conditions. Clause (i) specifies the core "relative informativeness" relation. This clause stipulates that a structure S subsumes another structure T only if the value of every attribute f in S subsumes its value in T. If f is an atom-valued feature, then the values assigned to f in S and T can be compared. If the value of $S(f)$ is a variable, then it subsumes any value that T assigns to f.[8] If the value of $S(f)$ is an atomic structure, such as PL, then $S(f)$ subsumes $T(f)$ only if T assigns the same value to f. If, on the other hand, f is a complex-valued feature, then clause (i) applies recursively to each of the features in the complex value and keeps recursing down until it reaches atom-valued features, which are then compared. Clause (ii) preserves structural *re-entrancies* of the sort introduced in the analysis of *virðist* in (5). This clause requires that if a structure contains a path of attributes that lead to a shared value, then it subsumes only structures in which the same paths lead to the same shared value.

(10) **Feature structure subsumption** (cf. Shieber 1986: 15):
A structure S SUBSUMES a complex structure T if and only if (i) $S(f) \sqsubseteq T(f)$, for all $f \in dom(S)$, and (ii) $T(p) = T(q)$, for all paths p and q such that $S(p) = S(q)$.

The subsumption relation in (10) provides a definition of what it means for one structure to contain the information of another, in terms of both the content and organization of attributes. The *unification* of feature structures can then be defined, as in (11), as the least informative structure that they subsume.

(11) **Feature structure unification** (cf. Shieber 1986: 17 f.):
The unification of two feature structures S and T [is] the most general feature structure U, such that $S \sqsubseteq U$ and $T \sqsubseteq U$.

Unification provides a general mechanism for determining the compatibility of information from different sources. Applied to a pair of compatible structures, unification returns the least informative structure that contains the information in both. Applied to incompatible structures, unification is said to "fail" or to return the inconsistent object "⊤." Unification can be described as "destructive," in the sense that it amalgamates actual inputs, rather than "copies" of those structures. The empirical value of unification can be illustrated by using it to distinguish compatible from incompatible feature structures. The first two structures in (12) repeat the structure associated with *er* 'he' and the value of the SUBJ attribute of *singt* 'sings' from (1). The third structure in (12) represents their unification, which combines the information from the first two structures.

(12) **Unification of features of *er* 'he' and SUBJ demands of *singt* 'sings':**

$$
\begin{bmatrix}
\text{PER} & 3 \\
\text{NUM} & \text{SG} \\
\text{GEND} & \text{MASC} \\
\text{CASE} & \text{NOM}
\end{bmatrix}
\sqcup
\begin{bmatrix}
\text{PER} & 3 \\
\text{NUM} & \text{SG} \\
\text{CASE} & \text{NOM}
\end{bmatrix}
=
\begin{bmatrix}
\text{PER} & 3 \\
\text{NUM} & \text{SG} \\
\text{GEND} & \text{MASC} \\
\text{CASE} & \text{NOM}
\end{bmatrix}
$$

The structures in (13) exhibit the conflict between the features of 1pl *wir* 'we' and those associated with the SUBJ value of 3sg *singt* 'sings'. This conflict leads to a "failure" of unification, represented by ⊤.

(13) **Unification "failure" due to feature conflict:**

$$
\begin{bmatrix}
\text{PER} & 3 \\
\text{NUM} & \text{SG} \\
\text{GEND} & \text{MASC} \\
\text{CASE} & \text{NOM}
\end{bmatrix}
\sqcup
\begin{bmatrix}
\text{PER} & 1 \\
\text{NUM} & \text{SG} \\
\text{CASE} & \text{NOM}
\end{bmatrix}
= \top
$$

8.3.2 Constraint satisfaction

The use of unification to regulate grammatical dependencies in models such as PATR is sometimes taken to reflect an "operational" perspective, in which grammatical dependents are associated with feature structures and the compatibility of structures is determined by unifying these structures. However, the effect of unification can be recast in description-based terms, by treating the regulation of grammatical dependencies as a case of constraint

satisfaction. As noted in the introduction, this type of approach begins by separating expressions, such as rules, principles, and entries, from the structures that they describe, usually trees and/or feature structures. Rather than associating *er* and *singt* directly with feature structures, a description-based account would assign these items lexical entries that, like those in (14), consist of sets of constraints.

(14)   **Partial lexical entries for *er* and *singt***:

$$
\text{er: NP,} \left\{ \begin{array}{l} (\uparrow \text{PER} = 3) \\ (\uparrow \text{NUM} = \text{SG}) \\ (\uparrow \text{GEND} = \text{MASC}) \\ (\uparrow \text{CASE} = \text{NOM}) \end{array} \right\} \quad \text{singt: V,} \left\{ \begin{array}{l} \uparrow \text{ SUBJ PER} = 3) \\ \uparrow \text{ SUBJ NUM} = \text{SG}) \\ \uparrow \text{ SUBJ CASE} = \text{NOM}) \\ \uparrow \text{ TENSE} = \text{PRES}) \end{array} \right\}
$$

The constraints in lexical entries are interpreted as descriptions of feature structures. The LFG notation "↑" in the constraints in (14) indicate that the constraints associated with *er* apply to the feature structure associated with its preterminal NP mother, whereas the constraints associated with *singt* apply to the feature structure associated with its preterminal V mother. The precise relation between constraints and the structures that they describe varies across approaches, reflecting different assumptions about the form of constraint languages and the nature of satisfying structures. Explicit formalizations of these relations can be found in Kaplan and Bresnan (1982), who present a procedure for solving sets of functional equations in LFG, or in King (1989) and Carpenter (1992), who provide model theories and definitions of constraint satisfaction that apply to the types of descriptions proposed within HPSG.

But to clarify the basic relation between unification and constraint satisfaction, it is useful to retain the intuitive conception of a feature structure as a function from features to values. A structure S will satisfy a constraint $(f = \alpha)$ whenever S assigns the value α to the attribute (or path of attributes) f. If f is atom-valued, then S satisfies $(f = \alpha)$ whenever $S(f) = \alpha$. If f is a finite sequence of attributes, S is applied successively to the attributes in this sequence; the constraint is satisfied if this process eventually yields the value α.

This simple model of constraint satisfaction can be illustrated with reference to the structures initially associated with *er* and *singt* in (1) and repeated in (15). The feature structure S_1 directly satisfies the constraints in the entry for *er* in (13): $S_1(\text{PER}) = 3$, $S_1(\text{NUM}) = \text{SG}$, $S_1(\text{GEND}) = \text{MASC}$, and $S_1(\text{CASE}) = \text{NOM}$. The feature structure S_2 satisfies the tense constraint in the entry for SINGT, given that $S_2(\text{TENSE}) = \text{PRES}$. A valence constraint such as "$(\uparrow \text{SUBJ CASE}) = \text{NOM}$" in the entry for *singt* is evaluated in two steps. The value that the structure S_2 assigns to the attribute SUBJ is determined first. This value is the structure S_3 in (15). Next, the value that S_3 assigns to the attribute CASE is determined. This value is the atom NOM. Hence S_2 satisfies the constraint "$(\uparrow \text{SUBJ CASE}) = \text{NOM}$" because applying S_2 successively to the attributes SUBJ and CASE yields the value NOM. Although the satisfaction of constraints containing paths of attributes is developed in greater formal detail in the works cited above, it should at least be intuitively clear at this point how a constraint containing a finite number of attributes can be evaluated by successively determining the value that a structure assigns to a single attribute.[9]

(15)   **Structures satisfying the entries of *er* and *singt***:

$$
S_1 : \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{MASC} \\ \text{CASE} & \text{NOM} \end{bmatrix} \quad S_2 : \begin{bmatrix} \text{SUBJ} & S_3 \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \end{bmatrix}
$$

The constraints in the entry of *er* in (14) describe structure S_1 in (15), while the constraints in the entry of *singt* describe structure S_2. The observation that *er* satisfies the subject agreement demands of *singt* is reflected in the fact that the constraints associated with *er* and the constraints associated with the SUBJ attribute of *singt* can be interpreted as descriptions of the same structure. As it happens, structure S_1 in (15) satisfies both sets of constraints. Moreover, the fact that S_1 is the same structure as the unified output in (12) shows that two sets of constraints can be interpreted as descriptions of the same structure if they independently describe a pair of structures that are unifiable.

8.3.3 Destructive regulation of grammatical dependencies

The notion of constraint satisfaction outlined above can used to regulate agreement relations and other types of grammatical dependencies, given constraints that (a) sanction a constituent analysis in which *er* occurs as the syntactic subject of *singt*, and (b) identify the SUBJ value of *singt* with its syntactic subject. The annotated phrase structure rules in (16) offer a particularly simple and transparent notation for expressing both types of constraint.
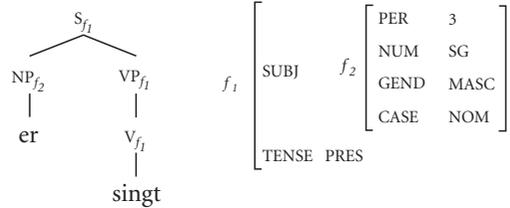
(16) **Annotated phrase structure rules** (Kaplan & Bresnan 1982):

    a. $S \rightarrow$ NP   VP
                          $(\uparrow \text{SUBJ}) = \downarrow$ $\uparrow = \downarrow$

    b. $VP \rightarrow$ V   NP
                         $\uparrow = \downarrow$ $(\uparrow \text{OBJ}) = \downarrow$

The phrase structure "backbone" of these rules can be interpreted as node admissibility conditions, as suggested by McCawley (1968) and Gazdar et al. (1985). The rule in (16a) admits a subtree consisting of an S node immediately dominating NP and VP nodes, while (16b) introduces a VP node with V and NP daughters. The annotations on these rules then project corresponding feature structure from the constituent structure admitted by the phrase structure component. In the sentence rule (16a), the constraint "($\uparrow$ SUBJ) = $\downarrow$" on the NP identifies the feature structure associated to the NP subject (designated "$\downarrow$") with the SUBJ value of its S mother (designated "($\uparrow$ SUBJ)"). In the verb phrase rule (16b), the constraint "($\uparrow$ OBJ) = $\downarrow$" on the NP similarly unifies the features of the NP object with the OBJ value of its VP mother. The constraint "$\uparrow = \downarrow$" on the VP in (16a) identifies the features of the VP with those of its S mother, while the same constraint on the V in (16b) identifies the features of the V with those of its VP mother. These constraints ensure that the SUBJ and OBJ features of a lexical verb are preserved by the structures corresponding to VP and S nodes, where they can be identified with the features of syntactic objects and subjects.

Example (17) shows how the annotated rules in (16) regulate subject agreement requirements. The agreement demands and tense properties of *singt* are satisfied by the structure f_1. The structure f_1 also corresponds to the VP and S nodes, due to the constraints $\uparrow = \downarrow$ in (16), which express the traditional view that *singt* is the head of the finite clause *er* singt. The structure f_2 satisfies the SUBJ constraints in the entry of *singt* and the features of the syntactic subject *er*. The "constructive" nature of constraint satisfaction (or unification) is reflected in the fact that the compatibility of *er* and *singt* is established by constructing a structure f_2 that preserves the properties of both dependents. The "destructive" character of constraint satisfaction is reflected in the fact that the properties of the dependents are not represented independently of the "consolidated" structure f_2.
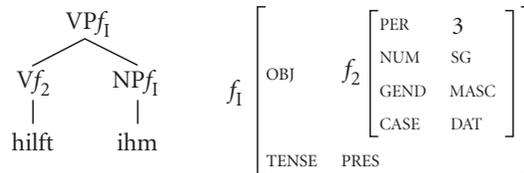
(17) **Associated c- and f-structure analyses of subject agreement**:

$$
\begin{array}{c}
S_{f_1} \\
\diagup \quad \diagdown \\
NP_{f_2} \qquad VP_{f_1} \\
| \qquad\qquad | \\
er \qquad\qquad V_{f_1} \\
\qquad\qquad\qquad | \\
\qquad\qquad\qquad singt
\end{array}
\qquad
f_1 \begin{bmatrix} \text{SUBJ} & f_2 \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{MASC} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \end{bmatrix}
$$

A parallel analysis applies to case government. In (18), the case governement demands of *hilft* and the properties of the syntactic object *ihm* again describe a common structure f_2. The existence of this structure establishes the compatibility of *hilft* and *ihm*; had the demands of the verb conflicted with the properties of its object, the value of the attribute OBJ would have been ⊤.

More generally, the analyses in (17) and (18) show how valence demands can be regulated by treating the complex values of SUBJ and OBJ attributes as the same structure as the structure described by the features of the corresponding syntactic subject or object. Constraint satisfaction or unification is characteristic of the mechanisms that combine features in feature-based approaches in that they are symmetrical, and "consolidate" information from different sources without keeping track of the provenance of any information or assuming that any one source will be more informative than another.[10]

(18) **Associated c- and f-structure analyses of case government**:

$$
\begin{array}{c}
VP f_1 \\
\diagup \quad \diagdown \\
V f_2 \qquad NP f_1 \\
| \qquad\qquad | \\
hilft \qquad ihm
\end{array}
\qquad
f_1 \begin{bmatrix} \text{OBJ} & f_2 \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{MASC} \\ \text{CASE} & \text{DAT} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \end{bmatrix}
$$

The interaction of destructive constraint satisfaction and complex-valued features also provides an analysis of nonlocal phenomena via the iteration of local identifications. For example, the transparency of a subject raising verb, such as *virðist* in (4), can be captured by identifying its SUBJ value with the SUBJ value of its complement. The raising verb and infinitival complement (XCOMP) are introduced by the annotated rule in (19a). Including the constraint in (19b) in the entry of *virðist* identifies its SUBJ value with its complement's SUBJ value.

(19) **Subject raising rule and constraint** (Kaplan & Bresnan, 1982):

a. $\text{VP} \rightarrow \begin{array}{cc} V & VP \\ \uparrow = \downarrow & (\uparrow \text{XCOMP}) = \downarrow \end{array}$

b. $(\uparrow \text{SUBJ}) = (\uparrow \text{XCOMP SUBJ})$

By virtue of the identification of SUBJ values in (20), *virðist* "inherits" any subject selection requirements associated with its complement. Since the entry of *vanta* contains a constraint specifying an accusative subject, this constraint will also be interpreted as describing the syntactic subject of *virðist*.

(20) **Associated c- and f-structure analyses of subject raising**:
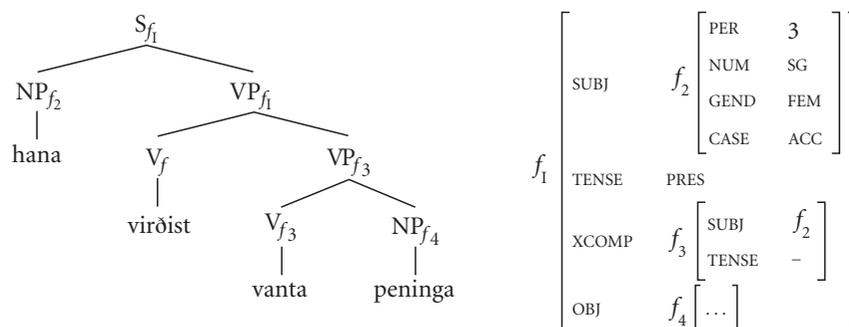
8.4 A Subsumption-Based Alternative

The analysis of raising constructions in (20) shows how destructive mechanisms induce a "flow" of feature information within a representation. The syntactic subject *hana* is not at any point a syntactic argument of *vanta*. However, because the constraints on the subj value of *vanta* and the constraints on the subj value of *virðist* are interpreted as describing the same structure, the syntactic subject of *virðist* must satisfy case constraints that are associated with the lexical entry of *vanta*. At the same time, the analysis of "mediated" dependencies like raising highlights distinctive properties of destructive strategies that are less obvious in the analysis of local dependencies. In local case or agreement dependencies, it is not entirely obvious whether there are two feature structures, corresponding to "controller" and "target" dependents, or whether there is just one structure, which is co-described by different entries. In the case of raising constructions, there are two predicates, each of which governs a subj value, and an independent constraint that identifies these values. In a feature-based analysis, the subj value of *virðist* must obey the case constraint in the entry of *vanta*. But it is unclear that there is any reason for the subj value of the infinitival complement *vanta* to share the agreement features of the finite form *virðist*. That is, the grammatical dependency in a raising construction is, in effect, asymmetrical; the raising verb must "inherit" the subj features of its complement, but the complement does not depend on the features of the raising verb.

A comparison with transformational accounts provides an instructive perspective. Transformational accounts incorporate two independent assumptions: first, that information is propagated "upwards" in a syntactic representation, and second, that this propagation is achieved through constituent-structure displacement. Accounts that substitute structure-sharing for "NP movement" revise both of these assumptions. However, a feature-based model can also express an asymmetrical dependency between raising verbs and their complements by replacing the constraint in (19b) with the subsumption-based counterpart in (21).

(21) **Subsumptive subject raising constraint**:
$$(\uparrow\text{xcomp subj}) \sqsubseteq (\uparrow \text{subj})$$

Associating (21) with *virðist* will ensure that its subj value satisfies any constraints imposed on the subj value of its complement. The analysis in (22) illustrates the effect of this revision. The top-level subj value f_2 is no longer identified with the subj value of the xcomp, as in (20). Instead, the subj value of the xcomp is an independent structure, which subsumes f_2.

(22) **Associated c- and f-structure analyses of raising**:
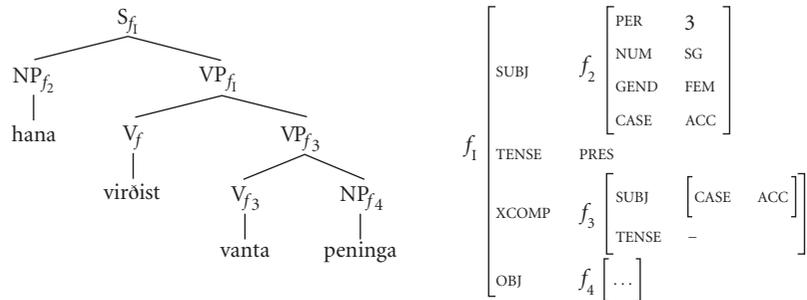
A similarly "constructive" analysis of other dependencies can be obtained by replacing the constraints in (16) with the subsumption-based counterparts in (23).

(23) **Subsumption-based rules**:

a. $\quad$ S $\rightarrow$ $\underset{\downarrow\sqsubseteq(\uparrow \text{SUBJ})}{\text{NP}} \quad \underset{\downarrow\sqsubseteq\uparrow}{\text{VP}}$

b. $\quad$ VP $\rightarrow$ $\underset{\downarrow\sqsubseteq\uparrow}{\text{V}} \quad \underset{\downarrow\sqsubseteq(\uparrow \text{OBJ})}{\text{NP}}$

In simple constructions, identification-based and subsumption-based analyses have no evident empirical differences. However, constructions in which a single element is simultaneously subject to conflicting grammatical demands provide relevant test cases, since it is in these contexts that the two approaches diverge. The treatment of feature neutrality provides the basis for this test. A subsumption-based approach preserves the neutrality of shared elements, and thus permits them to participate in multiple grammatical dependencies. In contrast, an identification-based account will tend to resolve neutrality, which should prevent an item from satisfying incompatible demands in different constructions. The following subsections review a range of constructions that appear to preserve feature neutrality and thus lend a measure of support to the use of subsumption rather than destructive mechanisms in feature-based approaches.

8.4.1   Neutrality and the limits of identification

Groos and van Riemsdijk (1981) identify free relative clauses in German as one construction in which conflicting case demands may be satisfied by neutral elements. They suggest that the case of relative pronouns in German free relatives must match not only the case governed by the lower verb, but also the case governed by the verb with which the entire free relative is associated. This requirement determines the ill-formedness of (24a), in which nominative *wer* 'who' violates the dative demands of *geholfen* 'helped' and dative *wem* 'whom' violates the nominative demands of *sein* 'be'. Groos and van Riemsdijk assert, however, that incompatible demands are satisfied in (24b) by non-oblique *was* 'what', which neutralizes the case conflict between *gegeben* 'given', which governs an accusative object, and *ist* 'is', which governs a nominative subject.

(24) a. $\quad$ *[Wer/Wem $\quad\quad$ nicht geholfen $\quad$ wird]muß klug $\quad$ sein.
$\quad\quad\quad$ who.NOM/whom.DAT not $\quad$ helped.DAT is $\quad\quad$ must clever be.NOM
$\quad\quad\quad$ 'Who(ever) is not helped must be clever.'

b. [Was   du mir gegeben hast] ist   prächtig.
    what.NOM/ACC you me given.ACC have is.NOM wonderful
    'What you have given me is wonderful.'

Coordinate structures provide another environment in which neutral elements appear to satisfy conflicting demands. Pullum and Zwicky (1986) note that complements of a coordinate verb must satisfy the case demands of each verb in German. The unacceptability of (25a) reflects the fact that neither conjunct can be construed with both verbs: dative plural *Männern* 'men' violates the case demands of accusative-governing *finden* 'find', while the accusative plural form *Männer* conflicts with dative-governing *helfen* 'help'. These conflicting demands are resolved in (25b) by the case-neutral plural *Frauen* 'women', which can be construed with both *finden* and *helfen*.

(25) a. *Sie findet   und hilft   Männer/Männern.
     she finds.ACC and helps.DAT men.ACC/men.DAT
     'She finds and helps men.'
   b. Er findet   und hilft   Frauen.
     he finds.ACC and helps.DAT women
     'He finds and helps women.'

Eisenberg (1973) describes a converse pattern in German, in which a neutral verb form agrees with distinct 1pl and 3pl subjects. The unacceptability of (26a) is attributed to the fact that 3sg *kauft* 'buys' conflicts with 1sg *ich*, while 1sg *kaufe* conflicts with 3sg *Franz*. No such conflict arises in (26b), as the plural form *kaufen* can be construed both with 1pl *wir* and with 3pl *die Müllers*.

(26) a. *weil Franz   das Haus und ich den Garten kauft/kaufe
     because Franz.3SG the house and 1SG the garden buy.3SG/buy.1SG
     'because Franz buys the house and I the garden'
   b. weil wir   das Haus und sie   den Garten kaufen
     because we.1PL the house and they.3PL the garden buy.PL
     'because we buy the house and they the garden'

The cases in (24)–(26) all involve the satisfaction of local valence demands. An analogous pattern arises in constituent questions in Polish, as Dyła (1984) notes. Whereas inanimate *co* violates the demands of *nienawidzi* in (27a), animate *kogo* satisfies both verbs in (27b).

(27) a. *Co   Janek lubi   i   Jerzy nienawidzi.
     what.NOM/ACC Janek likes.ACC and Jerzy hates.GEN
     'What does Janek like and Jerzy hate.'
   b. Kogo   Janek lubi   i   Jerzy nienawidzi.
     who.GEN/ACC Janek likes.ACC and Jerzy hates.GEN
     'Who does Janek like and Jerzy hate.'

8.4.2   Nondistinctness and information "flow"

The implications of these patterns for feature-based approaches are stated most cogently by Ingria (1990). Ingria notes that a destructive operation like unification immediately resolves the underspecification that permits neutral elements to satisfy conflicting demands. Consider, for example, the satisfaction of accusative and dative demands by case-neutral *Frauen* in (25).

Enforcing the [CASE ACC] requirements of *finden* by unifying these specifications with the case-neutral entry for *Frauen* yields a category with a [CASE ACC] value that conflicts with the [CASE DAT] requirements of *helfen*, and blocks further unification. Conversely, unifying the entry for *Frauen* with the [CASE DAT] requirements of *helfen* yields a category with a [CASE DAT] value that blocks unification with the accusative specifications of *finden*. These alternatives are schematized in (28).

(28) **Resolution of case neutrality:**

$$
\begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{PL} \\ \text{GEND} & \text{FEM} \end{bmatrix} \sqcup \begin{bmatrix} \text{CASE} & \text{ACC} \end{bmatrix} = \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{PL} \\ \text{GEND} & \text{FEM} \\ \text{CASE} & \text{ACC} \end{bmatrix}
$$

$$
\begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{PL} \\ \text{GEND} & \text{FEM} \end{bmatrix} \sqcup \begin{bmatrix} \text{CASE} & \text{DAT} \end{bmatrix} = \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{PL} \\ \text{GEND} & \text{FEM} \\ \text{CASE} & \text{DAT} \end{bmatrix}
$$

The structures in the first column of (28) represent the features of *Frauen*. The first structure in the second column represents the case government demands of *findet*, while the second structure represents the demands of *hilft*. The first result category in the third column bears an accusative value that conflicts with the dative demands of *hilft*. The second result category likewise bears a dative value that conflicts with the accusative demands of *findet*. The other examples in (24)–(27) instantiate the same basic pattern, suggesting that valence demands cannot be regulated in these environments by unification, or by any symmetrically destructive mechanism.[11] The conflict illustrated in (28) also arises on a constraint-based formulation, as no consistent structure can simultaneously satisfy the constraints associated with *findet* and *hilft*.

To avoid the undesired side effects of unification, some accounts propose nondestructive mechanisms for enforcing valence requirements. Ingria (1990: 200) formulates a nondistinctness check that has the effect of "determining that two expressions are unifiable, without actually unifying them." This alternative is illustrated in (29), in which "≃" represents a nondistinctness relation that holds of two structures whenever they do not assign distinct atomic values to a common path of attributes.. Since the case-neutral properties of *Frauen* are simultaneously nondistinct from the accusative demands of *findet* and the dative demands of *hilft*, a treatment of subcategorization that merely checks for nondistinctness will correctly classify (25b) as well-formed.

(29) **Preservation of case neutrality:**

$$
\begin{bmatrix} \text{CASE} & \text{ACC} \end{bmatrix} \simeq \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{PL} \\ \text{GEND} & \text{FEM} \end{bmatrix} \simeq \begin{bmatrix} \text{CASE} & \text{DAT} \end{bmatrix}
$$

Given that conflicting values for atom-valued features cause destructive mechanisms to fail, a strategy that checks for conflicts will often make the same split between compatible and incompatible structures. A nondistinctness check will also tend to impose the same

compatibility demand as destructive mechanisms in the regulation of case government and other local dependencies. Nevertheless, a fully nondestructive checking mechanism is incompatible with feature-based accounts of mediated compatibility demands. Because nondistinctness is an intransitive relation, a nondistinctness check does not induce the information flow that is integral to feature-based accounts. This point is illustrated by the fact that nondistingush between the acceptable raising sentence in (30a), repeated from (4b), from the unacceptable (30b).

(30) **Quirky case in Icelandic raising constructions:**
     a.   Hana   virðist   vanta peninga.
         her.ACC seem.3SG lack   money.ACC
     b.   *Hún   virðist   vanta peninga.
         her.NOM seem.3SG lack   money.ACC
         'She seems to lack money.'

In (31), corresponding to (30a), the first structure satisfies the constraints associated with *hana* 'her', the second satisfies the constraints on the SUBJ value of *virðist*, and the third satisfies the constraints on the SUBJ value of *vanta*. In this case, the acceptability of (30a) correlates with the fact that the SUBJ value of *virðist* is simultaneously nondistinct from the other structures. However, a nondisinctness check cannot reliably identify unacceptable raising constructions. The sentence in (30b) is unacceptable, due to a conflict between the nominative case associated with the syntactic subject *hún*. But the SUBJ value of *virðist* is again nondistinct from the other structures in (32).

(31) **Compatible case features in acceptable raising construction (30a):**

$$
\begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{FEM} \\ \text{CASE} & \text{ACC} \end{bmatrix} \simeq \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \end{bmatrix} \simeq \begin{bmatrix} \text{CASE} & \text{ACC} \end{bmatrix}
$$

(32) **Compatible case features in unacceptable raising construction (30b):**

$$
\begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEND} & \text{FEM} \\ \text{CASE} & \text{NOM} \end{bmatrix} \simeq \begin{bmatrix} \text{PER} & 3 \\ \text{NUM} & \text{SG} \end{bmatrix} \simeq \begin{bmatrix} \text{CASE} & \text{ACC} \end{bmatrix}
$$

Because a raising verb does not govern any particular case, it can be simultaneously nondistinct from multiple structures with conflicting case features. The lack of information flow that is instrumental in allowing a single dependent to satisfy conflicting case demands in (29) is what prevents a nondistinctness check from detecting the case conflict in (32). Whereas unifying the elements in a grammatical dependency eliminates the neutrality that allows a single element to satisfy incompatible requirements, a nondestructive compatibility check avoids the side effects of identification, but sacrifices the benefits of combining information. A subsumption-based account preserves the advantages of combining information from different sources, but avoids the conflicts that arise when the consolidation of features modifies inputs.

8.4.3　Simultaneous dependencies in coordinate structures

To illustrate how the use of subsumption can avoid the case conflict in (25b), it is useful to adopt provisionally the set-based analysis of constituent coordination proposed in LFG. The schematic rule in (33) assigns a coordinate mother the set of feature structures assigned to its conjunct daughters.

(33)　**Set-based coordination** (cf. Bresnan et al. 1985; Kaplan & Maxwell 1995b):

$$X \rightarrow \begin{array}{ccc} X & & X \\ \downarrow \varepsilon \uparrow & \text{CONJ} & \downarrow \varepsilon \uparrow \end{array}$$

If *findet und hilft Frauen* is interpreted as a case of V coordination, the rule in (33) will associate the coordinate V in (35) with a set that contains the feature structures assigned to *findet* and *hilft*. The head constraint "$\downarrow \sqsubseteq \uparrow$" in (23b) requires that the features of the VP preserve those of its V daughter. For sets of feature structures, this requirement is satisfied if, as in (35), the set assigned to the VP preserves the cardinality of the set assigned to the V, and each feature structure in the set assigned to V subsumes a unique counterpart in the set assigned to the VP. This intuition is formalized in (34).

(34)　For any sets S, S', $S \sqsubseteq S'$ iff there is a one-to-one correspondence f from S to S', such that $\forall s \infty S$, $s \sqsubseteq f(s)$.

The complement constraint "$\downarrow \sqsubseteq (\uparrow \text{OBJ})$" in (23b) is likewise satisfied in (35) if the properties of *Frauen* subsume each OBJ value in the set of feature structures assigned to its VP mother. The distribution of the constraint "$\downarrow \sqsubseteq (\uparrow \text{COMP})$" falls under the standard LFG principle.

(35)　**Property distribution** (Bresnan et al. 1985; Dalrymple & Kaplan 2001): For any property P and set s : $P(s)$ iff $\forall f \in s, P(f)$
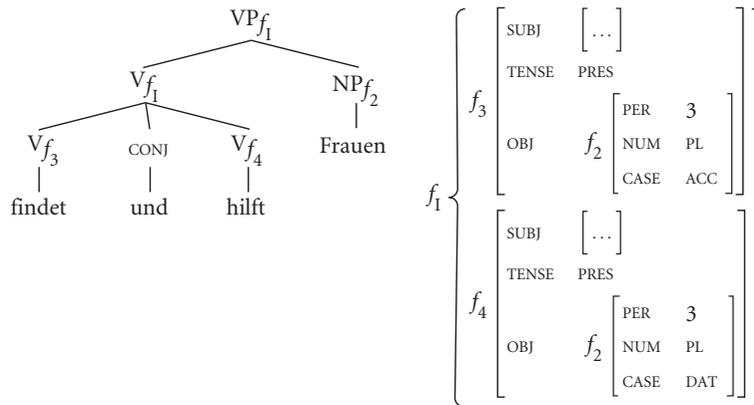
The constraints "$\downarrow \sqsubseteq \uparrow$" and "$\downarrow \sqsubseteq (\uparrow \text{OBJ})$" assign the VP in (35) the set f_1. The OBJ values of both structures in f_1 are subsumed by the features of *Frauen*. In addition, the first structure is subsumed by the OBJ features of *findet*, while the second structure is subsumed by the OBJ features of *hilft*. Thus the features of the NP complement are combined with those of the conjoined verbs, in contrast to a nondistinctness-based account. However, the resulting structures are assigned to distinct OBJ values on the VP mother, leaving the features of *Frauen* and

those of the coordinate V unmodified. Hence subsumption-based valence constraints determine the compatibility of a head and its syntactic complement by combining their features, but avoid the inconsistency that arises if the features of the head and complements are identified.

Although the analysis in (35) treats *findet und hilft* as a coordinate V, a subsumption-based account is equally compatible with an analysis that classifies *findet und hilft Frauen* as a case of VP coordination in which the conjuncts are *findet Frauen* and *hilft Frauen*. On this analysis, *findet und hilft Frauen* is a case of Right Node Raising in which the peripheral dependent *Frauen* is literally shared across multiple conjuncts (McCawley 1982; Levine 1985). In this case, the constraints "$\downarrow \sqsubseteq \uparrow$" and "$\downarrow \sqsubseteq (\uparrow\text{OBJ})$" will apply independently to the V and NP within each VP conjunct, determining the structures f_3 and f_4 in (35). The coordinate VP will again be associated with the set f_1, containing f_3 and f_4.

More generally, the use of subsumption provides a general strategy for avoiding conflicts in the other constructions in section 8.4.1, irrespective of the particular assumptions that are adopted regarding constituent structure and feature neutrality. In each of these constructions, subsumption-based constraints will have the effect of "pooling" feature information from syntactic arguments and verbal heads, determining an upward flow of feature information that terminates in the attribute values of a feature structure that corresponds to a sentence. Feature "pooling" imposes precisely the same compatibility demands as unification or structure-sharing, since the pooling of incompatible valence demands with syntactic arguments will violate subsumption constraints. However, the fact that the original elements remain unmodified by this process avoids conflicts when a single element is subjected to multiple compatibility demands.

8.4.4   Subsumption and exponence

Complex features are indispensable to all of the feature-based analyses outlined above. Equally important are mechanisms that consolidate information from different sources, rather than merely check for atomic conflicts. However, the difference between the destructive analyses in section 8.3 and the semi-destructive alternative in this section represents an essentially free choice within the class of feature-based approaches. Section 8.4.3 provides some empirical motivation for the use of subsumption, and Sells (2006) presents further support for subsumption-based treatments of raising and control constructions.

Subsumption-based constraints of the sort illustrated in (23) are also flexible in another, more fundamental, respect. These constraints merely require that the properties of syntactic heads and their dependents must be preserved on the corresponding feature structure. The constraints do not require that all of the properties of the mother must originate on one of its daughters. Hence a subsumption-based strategy is compatible with an *exponence-based* perspective in which the properties of a constituent *C* may be "realized" or "spelled out" by parts whose properties do not collectively exhaust the properties of *C*. Exponence-based approaches have been applied most systematically to the analysis of morphological systems (Matthews 1972; Anderson 1992; Aronoff 1994; Stump 2001), though recent accounts (summarized in Ackerman et al., this volume) extend this approach to periphrastic constructions and other syntactic formations.

8.5   Foundations and Implications

Where possible, the analyses above attempt to suppress inessential differences between approaches, in order to highlight issues that arise for the general class of feature-based formalisms. This section considers some points on which individual approaches differ significantly. Section 8.5.1 examines the treatment of structural re-entrancies, and contrasts the treatment of identity constraints with standard treatments of negation and disjunction.

Section 8.5.2 considers how the treatment of valence and locality determines the distinctive "head-driven" flow of information in HPSG. Section 8.5.3 outlines the relation between the model structures of LFG and HPSG and their treatment of underspecification. Section 8.6 concludes with some general remarks about the role of formal considerations.

8.5.1   The status of re-entrancy

The analyses in section 8.4 show how the use of subsumption to regulate grammatical dependencies can largely eliminate the need for re-entrant feature structures. In particular, a standard model of LFG can be modified to run in "subsumption mode" by replacing all identity relations between attributes by appropriate subsumption relations. The same substitution is possible in GPSG, but less straightforward in HPSG, for reasons discussed in section 8.5.2.

There are various reasons why one might wish to eliminate re-entrant feature structures, independent of the treatment of feature neutrality in the constructions in section 8.4.1. In the first place, eliminating re-entrancies simplifies the class of feature structures, and correspondingly simplifies the definition of relations and principles that apply to structures, since there is no need to include dedicated clauses – such as (10ii) – to apply to re-entrant configurations. This revision also dispenses with a grammatically inert distinction between "type-identical" and "token-identical" structures. The point is clearest in connection with atomic values, but applies equally to complex structures. It would appear to be a purely implementational choice whether one regards all plural expressions in a given syntactic structure as sharing a single PLU value, or as being assigned different occurrences of PLU. Much the same is ultimately true of complex structures. There is simply no reason to think that any grammatical process is — or, indeed, could be — sensitive to a difference between type and token identity.

Moreover, interpreting identity constraints as a relating distinct structures, rather than as defining a re-entrant structure, is more compatible with a description-based perspective. A comparison with treatments of negation and disjunction brings this point out clearly. Early feature-based accounts, such as Kay (1979) and Karttunen (1984), model negation and disjunction directly in terms of negative and disjunctive feature structures. Within this sort of approach, it is natural to model identity relations by means of distinctive, re-entrant structures, in which features literally share a value. However, largely following the lead of LFG, current approaches tend to treat negation and disjunction as abbreviatory devices that are confined to descriptions. Within this type of approach, it is more natural to treat identity too as constraint-language relation between non-re-entrant structures. More specifically, a description-based approach interprets a disjunctive constraint of the form $(\phi \vee \psi)$ as describing a set of structures, each of which satisfies either ϕ or ψ. A negative constraint of the form $(\neg \phi)$ is likewise interpreted as describing the set of structures that either fail to satisfy ϕ or conflict with ϕ. Hence the operators "$\vee$" and "$\neg$" are modeled not by distinctive types of structures, but by distinctive relations between basic structures. The same logic applies to identity statements. Rather than interpreting $(\phi = \psi)$ as describing a special token-identical structure, a description-based approach can model "$=$" in terms of a mutual subsumption relation between structures ϕ and ψ.

Description-based analyses of negation and disjunction illustrate one way in which feature-based approaches exploit the trade-off between the complexity of descriptions and the complexity of structures. Confining logical operators to constraint-language expressions eliminates the logically complex structures proposed in earlier accounts. Extending this analysis to identity statements further simplifies the class of structures by eliminating re-entrant feature values. This leads to a model in which the values of complex-valued

features are, in fact, relatively simple. A feature structure is comprised of a collection of attributes and values, which can be represented formally by means of a function or a graph. Feature values may be atomic objects, such as PL or ACC, or they may be substructures, which consist again of attributes and values. The simplicity of these structures highlights the fact that a limited inventory of complex-valued attributes is what avoids the need for anything like the layers of "functional" constituent structures employed in transformational accounts (Chomsky 1995).

8.5.2 Valence and Locality

As suggested earlier, the treatment of valence outlined in section 8.4 amounts, in effect, to running an LFG model in "subsumption mode." The use of semi-destructive constraints to regulate valence demands is, however, more fundamentally incompatible with HPSG. One source of conflict derives from the fact that the valence terms in HPSG are "canceled" as syntactic arguments are encountered. A transitive verb will initially be associated with a feature structure containing singleton SUBJ and COMP(LEMENT)s lists. After combining with its syntactic complement, the verb will head a VP structure with a singleton SUBJ list and an empty COMPS list, indicating that complement demands have been met by a syntactic complement. The feature structure associated with the corresponding sentence will contain empty SUBJ and COMPS lists, indicating the satisfaction of all subcategorization demands.

Because valence terms are canceled, SUBJ and COMPS cannot represent structure-shared values in the feature structure associated with a sentence. To provide a persistent "repository" for the feature information of predicates and syntactic arguments, recent HPSG accounts follow Manning and Sag (1999) in consolidating the valence terms of a lexical head in an ARG(UMENT)-S(TRUCTURE) list. Yet, by classifying ARG-S as a lexical (word-level) feature, HPSG accounts ensure that information about the arguments of a head are not "visible" on its phrasal projections. The division of labor between cancelable valence terms and persistent argument structure terms enforces a fairly strict notion of locality, albeit a highly "configurational" conception, which implicitly keys the accessibility of feature information to its presence or absence in a structure.[12] This configurational view of locality is in no sense intrinsic to a feature-based approach, and, in fact, conflicts with other aspects of HPSG accounts. For example, HPSG accounts integrate syntactic and semantic properties into a single data structure, termed a "sign." Signs typically contain detailed semantic features that are, in principle, visible to subcategorizing heads, but tend not to be subcategorized for, or otherwise "accessed" from without. Hence, in HPSG, the observation that no process selects feature information at a certain place in a structure is not a reliable indicator that the information is absent from the structure.

The cancelation of valence demands also precludes the use of SUBJ and COMPS values as repositories for "pooled" features in a subsumption-based variant of HPSG. It is, nevertheless, possible to combine subsumption and cancelation with less "configurational" strategies for access control. The most straightforward account would follow Manning (1996) in associating ARG-S features with phrases, so that the ARG-S features of a mother provide a repository for pooling the properties of its daughters. The beneficial effects of a standard locality condition will be largely preserved if only head daughters can directly access the ARG-S features of their mothers. All other access to phrasal ARG-S values would be mediated through cancellable SUBJ or COMP features. This alternative grants a head daughter privileged access to the features of its mother and clarifies the fact that SUBJ and COMPS features function essentially as "accessor" functions in HPSG, controlling syntactic access to the argument structure of a predicate.

8.5.3   Underspecification and minimality

A more basic contrast between feature-based approaches concerns the general constraints that they impose on feature structures. These constraints are stated most clearly in description-based approaches, where they form part of the model theory of a formalism. The difference between the "minimal" models of LFG and the "maximal" models of HPSG bears directly on the treatment of feature neutrality. LFG imposes a *minimality* condition on f-structures that excludes features that are not specified by some constraint in an f-description. This condition captures the intuition underlying the use of underspecification, namely that only distinctive features are represented in linguistic analyses.

> In general, if an f-description has one solution, it has an infinite number of "larger" solutions. Of course, there is something counterintuitive about these larger solutions. The extra features they contain cannot possibly conflict with those specifically required by the f-description. In that sense they are grammatically irrelevant and should not really count as f-structures that the grammar assigns to sentences. This intuition, that we only countenance f-structures with relevant features and values, can be formalized in a technical refinement to our previous definitions that makes "*the* f-structure of a sentence" a well-defined notion. (Kaplan & Bresnan 1982: 202)

A minimality condition is compatible with underspecified analysis, such as the case-neutral analysis of *Frauen* in section 8.4. Admitting minimal models also ensures the well-formedness of the structures that result from substituting subsumption constraints for identity constraints in the rules in (16).

In contrast, underspecified analyses are not admissible structures in HPSG. Underspecification in general violates the requirement that structures must be *totally well-typed* (Carpenter 1992) in the sense that they must be assigned a value for each appropriate feature. This requirement bars a case-neutral interpretation for *Frauen* in (35). A separate requirement that structures must be *sort-resolved* (Pollard & Sag 1994: 18) permits only "fully specific" feature values and thus bars disjunctive case values from occurring in a well-formed structure. This condition prevents an analysis from trivially satisfying total well-typing by assigning *Frauen* a CASE value such as {NOM, ACC, DAT, GEN}, which just exhaustively lists the available case values. Yet sort resolution also bars the use of nontrivial values like {NOM, ACC} to express the fact that an item such as *was* in (24b) neutralizes the contrast between nominative and accusative case.

These requirements reflect the same idealizations that underlie the problems faced by the unification-based approach in section 8.4.2. Although the HPSG formalism can be revised to exploit the subsumption-based strategies in section 8.4.3, these revisions are incompatible with the HPSG model theory. As the contrast with LFG shows, this is a property of HPSG, rather than of feature-based or even description-based approaches in general. The assumption that structures must be totally well-typed and sort-resolved avoids the need for counterparts of the completeness and coherence conditions in LFG, and facilitates type-based inferencing within HPSG. However, like other foundational assumptions, these do not rest directly on any empirical considerations, nor, correspondingly, are they susceptible to direct empirical refutation. As such, these assumptions do not express empirical hypotheses so much as constitute part of the definition of an HPSG grammar. It thus remains an open question whether the phenomena described in section 8.4.1 can be reconciled with the foundational assumptions of HPSG, at least as they are presently understood.

8.6 Conclusions

The tension between the foundational assumptions of HPSG and the apparent persistence of feature neutrality in the constructions in section 8.4.1 illustrates the type of choice between empirical and formal considerations that is often faced by feature-based approaches. Just as the model theory of HPSG dictates the treatment of underspecification and neutrality, initial decisions about the role of constituent order determine the role of SLASH and NONLOCAL features in GPSG and HPSG and motivate the division of labor between c-structure and f-structure in LFG. These sorts of "foundational" assumptions are largely shielded from scrutiny within a particular approach, since any revision effectively takes one outside that approach. Foundational issues thus provide an instructive contrast with the subsumption-based strategies outlined in previous sections, as these strategies preserve the form and character of feature-based approaches.

More remains to be said about these issues, and about other design choices made within individual feature-based frameworks. But the primary goal of the present discussion has been to isolate some basic issues, clarify their interaction with other assumptions, and, where possible, assess their empirical impact.

Notes

1 The model-theoretic foundations of feature-based models have also led to the development of full feature logics (Johnson 1988; King 1989; Carpenter 1992), providing a point of contact between feature-based approaches and other deductive grammar formalisms.

2 Matthews (1991) adopts the term "morphosyntactic category" for what are here termed "attributes" and "morphosyntactic property" or "morphosyntactic feature" for "values."

3 See Andrews (1982, 1990), Neidle (1982), and Blevins (2000) for some neo-Jakobsonian extensions of feature-based models.

4 The feature structures that satisfy AVMs are modeled by (directed acyclic) graphs in HPSG.

5 The structure in (8) also contains the LFG PRED attribute, which represents the semantic form associated with an item and functions implicitly as a unique index for each unique item.

6 Though constituency-preserving analyses of auxiliary inversion, phrasal verbs, and other constructions are developed in models of Montague Grammar (Bach 1979, 1980), as well as in the feature-based tradition represented by Head Grammars (Pollard 1984) and linearization-based models of HPSG (Reape 1994; Kathol 2000).

7 The introduction of "COMP" positions defined a "derived constituent structure" for extraction constructions, answering a critique raised by Robert Stockwell some twenty years earlier (Hill 1962: 158) and solving a problem that does not arise on nonderivational approaches.

8 The value of $S(f)$ will be a variable under two circumstances: when S assigns no definite value to the attribute f, and when f is not in the domain of S. However, most models do not distinguish a structure that contains an attribute with an undefined value from a structure that lacks that attribute altogether.

9 Within LFG, the device of "functional uncertainty" (Kaplan & Maxwell 1995a) provides solutions to constraints containing paths of arbitrary and even unbounded length.

10 The asymmetry of these mechanisms clearly distinguish them from the "copying" strategies assumed in transformational accounts, and implicitly deny the importance of asymmetrical relations between "targets" and "controllers" assumed in many traditional accounts.

11 There is no reason to suppose that these are isolated or unrepresentative examples. For instance, there are reports of similar patterns involving case government in French

(Kayne 1975) and Icelandic (Zaenen & Karttunen 1984), definiteness concord in Hungarian (Szamosi 1976), and noun class agreement in Xhosa (Voeltz 1971).

12   A similar conception underlies the "two-level" analysis of inherent case in Andrews (1982).

References

Ackerman, Farrell & Gect. Webelhuth. 1998. *A Theory of Predicates*. Stanford: CSLI.

Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Andrews, Avery. 1982. The representation of case in modern Icelandic. In Bresnan 1982: 427–503.

Andrews, Avery. 1990. Unification and morphological blocking. *Natural Language and Linguistic Theory* 8: 507–58.

Andrews, Avery & Christopher Manning. 1999. *Complex Predicates and Information Spreading in LFG*. Stanford: CSLI.

Aronoff, Mark. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.

Bach, Emmon. 1979. Control in Montague Grammar. *Linguistic Inquiry* 10: 515–31.

Bach, Emmon. 1980. In defense of passive. *Linguistics and Philosophy* 3: 297–341.

Blackburn, Patrick. 1994. Structures, languages and translations: the structural approach to feature logic. In C. J., Rupp, M. A. Rosner, & R. L. Johnson (eds.), *Constraints, Language and Computation*. London: Academic Press, 1–27.

Blevins, James P. 2000. Markedness and agreement. *Transactions of the Philological Society* 98(2): 233–62.

Bloomfield, Leonard. (1933). *Language*. Chicago: University of Chicago Press.

Bresnan, J. (ed.). 1982. *The Mental Representation of Grammatical Relations*, Cambridge, MA: MIT Press.

Bresnan, Joan, Ron Kaplan, Stanley Peters, & Annie Zaenen. 1982. Cross-serial dependencies in Dutch. *Linguistic Inquiry* 13.4: 613–35.

Bresnan, Joan, Ron Kaplan, & Peter Peterson. 1985. Coordination in LFG. MS, University of Newcastle.

Carpenter, Bob. 1992. *The Logic of Typed Feature Structures*. New York: Cambridge University Press.

Chomsky, Noum. 1956. Three models for the description of language. *Institute of Radio Engineers Transactions on Information Theory* II-2: 113–24.

Chomsky, Noum. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noum. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noum. 1977. *Essays on Form and Interpretation*. Amsterdam: North-Holland.

Chomsky, Noum. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Dalrymple, Many & Ron Kaplan. 2001. A set-based approach to feature resolution. *Language* 76.4: 759–98.

Dalrymple, Many, Ron Kaplan, John T. Maxwell III, & Annie Zaenen. (eds.). 1995. *Formal Issues in Lexical-Functional Grammar*. Stanford: CSLI.

Dyła, Stefan. 1984. Across-the-board dependencies and case in Polish. *Linguistic Inquiry* 15.4: 701–5.

Eisenberg, Peter. 1973. A note on 'identity of constituents'. *Linguistic Inquiry* 4: 417–20.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12: 155–84.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Gleason, Henry. 1955. *Workbook in Descriptive Linguistics*. New York: Holt, Rinehart & Winston.

Groos, Anneke & Henk van Riemsdijk. 1981. Matching effects in free relatives: a parameter of core grammar. In A. Belletti, L. Brandi, & L. Rizzi (eds.), *Theory of Markedness in Generative Grammar*. Pisa: Foris, 171–216.

Harman, Gilbert. 1963. Generative grammars without transformational rules. *Language* 39: 567–616.

Harris, Zellig. 1946. From morpheme to utterance. *Language* 22: 161–83.

Harris, Zellig. 1951. *Methods in Structural Linguistics*. Chicago: University of Chicago Press.

Harris, Zellig. 1957. Co-occurrence and transformation in linguistic structure. *Language* 33: 283–340.

Harris, Zellig. 1965. Transformational theory. *Language* 41: 363–401.

Hill, Archibald. (ed.). 1962. *Third Texas Conference on Problems of Linguistic Analysis in English*. Austin: University of Texas.

Hockett, Charles. 1958. *A Course in Modern Linguistics*. New York: Macmillan.

Ingria, Robert. 1990. The limits of unification. In *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*, 194–204.

Jakobson, Roman. 1932. Zur Structur des russischen Verbums. In *Charisteria Guilelmo Mathesio Quinquagenario*, 74–83. Reprinted in E. P. Hamp, F. W. Householder, & R. Austerlitz (eds.). 1966. *Readings in Linguistics II*. Chicago: University of Chicago Press, 22–30.

Jakobson, Roman. 1936. Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus. *Travaux du Cercle Linguistique de Prague VI*, 240–99. Reprinted as 'Contribution to the general theory of case: general meanings of the Russian cases' in Linda R. Waugh & Morris Halle (eds.). 1984. *Roman Jakobson: Russian and Slavic Grammar*. Berlin: Mouton, 59–103.

Jespersen, Otto. 1937. *Analytic Syntax*. London: Allen & Unwin. Reprinted 1984. Chicago: University of Chicago Press.

Johnson, Mark. 1988. *Attribute-Value Logic and the Theory of Grammar*. Stanford: CSLI.

Kaplan, Ron & Joan Bresnan. 1982. Lexical-functional grammar: a formal system for grammatical representation. In Bresnan 1982: 173–281.

Kaplan, Ron & John T. Maxwell III. 1995a. An algorithm for functional uncertainty. In M. Dalrymple, R. M. Kaplan, J. T. Maxwell III, & A. Zaenen (eds.), *Formal Issues in Lexical – Functional Grammar*, Stanford: CSLI, 177–197.

Kaplan, Ron & John T. Maxwell III. 1995b. Constituent coordination in Lexical Functional Grammar. In M. Dalrymple, R. M. Kaplan, J. T. Maxwell III, & A. Zaenen (eds.), *Formal Issues in Lexical – Functional Grammar*, Stanford: CSLI, 198–210.

Kaplan, Ron & Annie Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In M. R. Baltin & A. S. Kroch (eds.), *Alternative Conceptions of Phrase Structure*, Chicago: University of Chicago Press, 17–42.

Karttunen, Lauri. 1984. Features and values. In *Proceedings of COLING 84*. 28–33.

Kathol, Andreas. 2000. *Linear Syntax*. Oxford: Oxford University Press.

Kay, Martin. 1979. Functional grammar. In *Proceedings of the Fifth Annual Meeting of the Berkeley Linguistics Society*, 142–58.

Kayne, Richard. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

King, Paul. 1989. A logical formalism for Head-Driven Phrase Structure Grammar. PhD thesis, University of Manchester.

Levine, Robert D. 1985. Right node (non)-raising. *Linguistic Inquiry* 16: 492–7.

Manning, Christopher D. 1996. *Ergativity: Argument Structure and Grammatical Relations*. Stanford: CSLI.

Manning, Christopher D. & Ivan A. Sag. 1999. Dissociations between argument structure and grammatical relations. In G. Webelhuth, J.-P. Koening, & A. Kathol, (eds.), *Lexical and Constructional Aspects of Linguistic Explanation*. Stanford: CSLI, 63–78.

Matthews, Peter. 1972. *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press.

Matthews, Peter. 1991. *Morphology*. 2nd edn. Cambridge: Cambridge University Press.

McCawley, James D. 1968. Concerning the base component of a tranformational grammar. *Foundations of Language* 4: 243–67.

McCawley, James D. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13: 99–107.

Neidle, Carol. 1982. Case agreement in Russian. In Bresnan 1982: 391–426.

Pollard, Carol. 1984. Generalized phrase structure grammars, head grammars and natural language. PhD thesis, Stanford.

Pollard, Carol & Ivan A. Sag. 1987. *Information-Based Syntax and Semantics*. Stanford: CSLI.

Pollard, Carol & Ivan A. Sag. 1992. Anaphors in English and the scope of Binding Theory. *Linguistic Inquiry* 23: 261–303.

Pollard, Carol & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Stanford: University of Chicago Press and CSLI.

Pullum, Geoffrey K. & Arnold Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language* 62: 751–73.

Reape, Mike. 1994. Domain union and word order variation in German. In J. Nerbonne, K. Netter, & C. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*. Stanford: CSLI, 151–98.

Sag, Ivan A. & Janet Dean Fodor. 1994. Extraction without traces. In R. Aranovich, W. Byrne, S. Preuss, & M. Senturia (eds.), *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*. 365–84.

Sag, Ivan A., Lauri Karttunen, & Jeffrey Goldberg. 1992. A lexical analysis of Icelandic case. In I. A. Sag & A. Szabolcsi (eds.), *Lexical Matters*. Stanford: CSLI, 301–18.

Sells, Peter. 2006. Using subsumption rather than equality in functional control. In M. Butt & T. Holloway King (eds.), *Proceedings of LFG-06*, 455–73.

Shieber, Stuart M. 1986. *An Introduction to Unification-Based Approaches to Grammar*. Stanford: CSLI.

Stump, Gregory T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

Szamosi, Michael. 1976. On a surface structure constraint in Hungarian. In J. D. McCawley (ed.), *Syntax and Semantics 7: Notes from the Linguistic Underground*. New York: Academic Press, 409–25.

Voeltz, Erhard. 1971. Surface constraints and agreement resolution: some evidence from Xhosa. *Studies in African Linguistics* 21: 37–60.

Wells, Rulon. 1947. Immediate constituents. *Language* 23: 81–117.

Zaenen, Annie & Lauri Karttunen. 1984. Morphological non-distinctness and coördination. In *Eastern States Conference on Linguistics*, 309–20.