Article Number: LALI: 01984



a0005

# Features and Feature Structures

**J P Blevins**, University of Cambridge, Cambridge, UK



p0005 Classical grammars approach grammatical analysis essentially as a problem of classification, as Matthews (1991: Sect. 10) observes. An utterance is divided into parts, which are assigned to word classes and then subclassified in terms of their 'accidents' or properties. A fundamentally similar conception of grammatical analysis is incorporated in modern feature-based grammars. Modern approaches have developed more complex and somewhat more imposing constraint languages and formal representations. But underlying this feature technology is a core of relatively straightforward claims and hypotheses. Foremost among these is the idea that grammatical properties and dependencies—local and nonlocal—can be described in terms of complex-valued features. It is the use of complex feature values to express grammatical dependencies that principally distinguishes approaches such as GPSG, LFG, and HPSG from models that attribute the 'flow' of feature information to constituent-structure displacements or other mechanisms.

p0010 A second key idea is that feature matching within an endocentric phrase need not be confined to the word-class features that define major categories such as noun, verb, and adjective. All feature-based approaches provide strategies, of one sort or another, which allow a phrase to share the inflectional properties of its syntactic head. These strategies allow the number features of a noun to be shared with the noun phrase that it heads, and permit the tense and agreement features of a verb to be shared with the verb phrase or clause that it heads. The resulting syntactic analyses are, so to speak, 'bushy,' as the feature information at a given node tends to be relatively dense, but there are comparatively few nodes between the 'root' and 'leaves' of an analysis.

p0015 Feature-based approaches also tend to regulate grammatical dependencies constructively, by consolidating the properties of heads and arguments, agreement targets, and controllers, and even 'fillers' and 'gaps.' In most approaches, the elements in a grammatical dependency are identified as the same structure. The early feature-based accounts, described by Shieber, 1986, identify feature structures by applying a symmetrically destructive unification operation. In more recent accounts, the elements in a grammatical dependency are interpreted as 'codescriptions' of a common structure. This reflects a general shift to a description-based perspective, in which grammatical expressions, such as rules, constraints, or entries, are interpreted as descriptions, and objects such as trees and/or feature structures are interpreted as structures that satisfy those descriptions. This separation has led to the development of feature logics for grammatical description (Johnson, 1988; King, 1989; Carpenter, 1992), providing a point of contact between feature-based approaches and other deductive grammar formalisms, such as the Lambek calculus.

## The Structure and Distribution of Feature Information

s0005

p0020 To a great extent, the form of a syntactic analysis reflects assumptions about the organization and distribution of morphosyntactic information. In simple terms, feature-based models adopt comparatively simple constituent structures and more articulated feature structures, whereas transformational accounts assume relatively simple feature bundles, but highly complex constituent structures. This contrast can be traced to different ideas about the distribution of features.

p0025 The prerequisites of a feature-based approach are fully present in standard theory transformational grammars, which contain atomic features, complex symbols, and a lexical insertion rule whose nondistinctness condition (Chomsky, 1965: 121) anticipates operations like unification. Nevertheless, the descriptive potential of these devices is severely constrained in the standard theory by the fact that complex values are confined to the preterminal level on which they are introduced. Similar constraints on the distribution of morphosyntactic properties likewise determine the form of analyses in late transformational models, such as Chomsky, 1995.

p0030 Syntactic analyses in these accounts are much more 'spindly' than those in feature-based models, as the information at any given node tends to be sparse, but there are many more nodes over which information is distributed. This organization reflects the view that word class features are fundamentally different in character from other types of lexical or inflectional properties. This view in turn can be traced to the fact that the 'morpheme class marks' in Harris (1951: 263) only represent word class and contrasts between simple and complex expressions. This focus on word class and phrase type was retained when Chomsky (1970) recast Harris' distributional model as X-Bar Theory. The X-bar conventions thus provide a general mechanism for matching the word-class features of an endocentric phrase and its syntactic head. These conventions do not, however, regulate the compatibility

Article Number: LALI: 01984



of other types of features. Within the transformational paradigm, this has led to the wholesale reclassification of inflectional properties as word-class features, associated with various types of abstract formatives. Reclassifying traditional properties such as tense or aspect as features of functional heads allows these properties to be regulated by the restricted X-bar conventions. Agreement properties are similarly balkanized into a sequence of abstract formatives, each with a category that corresponds to a traditional agreement feature.

## The Role of Complex-Valued Features

Feature-based models take a more conservative view, in which the properties traditionally associated with a word class are directly assigned to the lexical items in that class. Verbal properties, such as tense, aspect, mood, and voice, are directly associated with verbs, and nominal properties such as person, number, gender, and case are associated with nouns. However, the key innovation in feature-based approaches involves the use of complex feature values to regulate grammatical dependencies. The analysis of subject-verb agreement provides a useful illustration. The subject agreement demands of an English verb such as *skis* may be expressed by assigning *skis* a complex-valued subj(ect) attribute that contains the features that represent third person and singular number. In a simple feature system, these might be [per 3] and [num *sg*]. Agreement between *skis* and *he* in (1) is then keyed to a requirement that the subj features associated with the verb must be compatible with the grammatical features of its syntactic subject. The details of this analysis vary across approaches, though the leading idea is that the properties of the agreement 'controller' *he* and the agreement 'target' *skis* must contain no features that conflict. In the early feature-based models described in Shieber (1986), this is regulated by first assigning the subject *he* and the verb *skis* the partial feature structures in (1). The compatibility of these structures is then determined by unifying them, which defines the structure associated with *he skis* in (1) (**Figure 1**).

(1) Agreement via complex-valued agreement features

In models such as LFG (Kaplan and Bresnan, 1982) or HPSG (Pollard and Sag, 1994), the agreement constraints in the entries for *he* in (2a) and *skis* (2b) are interpreted as partial descriptions of an agreement matrix. Compatibility between the agreement properties of *skis* and those of its syntactic subject is determined by interpreting the constraints in (2a) and (2b) as descriptions of the same feature structure, in accordance with the constraint stated informally in (2c).

(2a) *he*: {(per = 3), (num = *sg*), (gend = *masc*)}
(2b) *skis*: {(subj per = 3), (subj num = *sg*) . . . }
(2c) A verb's subj features describe the same structure as the features of its syntactic subject.

On either a unification- or description-based alternative, agreement is a symmetrical relation, unlike the asymmetrical 'copying' relations assumed in many transformational accounts. There is no requirement that agreement properties must be uniquely associated with the agreement target or controller, or that either element must in any sense be more informative than the other. Symmetry is widely viewed as a virtue, as order-independent formalisms fit particularly well with incremental models of comprehension or production. Nevertheless, it remains to be seen whether a symmetrical approach can provide illuminating analyses of all of the cases that motivate the traditional distinction between agreement 'controllers' and 'targets' (Corbett, 1991).

The introduction of complex-valued features overcomes many of the descriptive limitations of simple phrase structure systems, as originally noted in Harman (1963). For example, the subcategorization demands of a transitive verb can be represented straightforwardly by a complex-valued obj(ect) attribute. The verb's demands will be satisfied if its obj properties of the verb are compatible with the properties of its syntactic object, where compatibility can again be determined in unification- or description-based terms. Nonlocal dependencies can likewise be described by means of a chain of local dependencies involving complex-valued attributes.

In the subject raising example in (3a), the expletive *it* occurs as the subject of *seems* but ultimately satisfies the valence demands of the embedded weather verb *raining*. The term 'raising' reflects the fact that transformational descriptions of this pattern tend to introduce *it* as a syntactic subject of *raining*, which is then advanced to the matrix subject position by means of constituent-structure displacements. Complex-valued features permit an account of this pattern in terms of a sequence of identities. The idea is just that the subj features of a subject-raising verb are interpreted as describing the same feature structure as the subj features of its complement, in accordance with (3c). If, as often assumed, *seems, have*, and *been* are all classified as subject-raising verbs, then the subj features of *seems* will describe the same structure as the subj features of *have*. The subj features of *have* will describe the same structure as the subj values of *been*, and the subj features of *been* will in

Article Number: LALI: 01984



turn describe the same structure as the subj features of *raining*. Through this chain of identities, the subj demands of each of the verbs in (3a) will come to describe the same structure. In order to satisfy the demands of *raining*, this structure must be the structure associated with the expletive subject *it*. Hence *seems* must select *it* in (3a).

(3a) It seems to have been raining.
(3b) The subject-raising (SR) verbs in English include *seem*, *have*, and *be*.
(3c) A SR verb's subj features describe the same structure as its complement's subj features.

p0060 Complex-valued features also permit the description of grammatical dependencies that span potentially unbounded domains, as initially demonstrated by the treatment of extraction constructions in GPSG (Gazdar *et al*., 1985). The principle is the same as in the treatment of raising constructions, though different grammatical attributes may mediate between a 'filler' and 'gap.'

## s0015 Compatibility and Identity

p0065 Individual feature-based approaches differ in a variety of technical respects, such as the relative informativeness of the structures that are taken to satisfy a grammatical description. Whereas structures in LFG are 'minimal' solutions to a set of constraints, a structure in HPSG is assumed to be specified for any properties that are appropriate for structures of its type. The organization of structures into hierarchies of 'types' is another distinctive characteristic of HPSG accounts.

p0070 A number of more general issues also arise within the broad class of feature-based formalisms. One basic issue concerns the use of unification or identity to model grammatical 'compatibility.' It has often been observed that a single element appears to be able to satisfy multiple, conflicting demands if it is neutral for the conflicting features. In German, for example, a verb form such as *kaufen* allows either the 1pl subject *wir* or a 3pl subject such as *sie*. This correlates with the fact that *kaufen* is compatible with both *wir* and *sie* in the coordinate structure in (4).

(4) . . . *weil wir das Haus und sie den Garten kaufen*.
. . . because we.1pl the house and they.3pl the garden bought
'. . . because we bought the house and they the garden.'

p0075 As Ingria (1990) notes, 'destructive' operations like unification would be expected to eliminate the feature neutrality that allows *kaufen* to combine with *wir* and *sie* in (4), since unifying the agreement features of *kaufen* with the features of either subject will yield an agreement matrix with person features that conflict with those of the other subject. A similar issue arises on a description-based alternative that interprets the subj features of *kaufen* as describing the same structure as the features of *wir* and *sie*, since no consistent structure can satisfy the properties of *kaufen*, *wir*, and *sie*. Ingria proposes that compatibility is more appropriately regulated by a 'nondistinctness' check. However, a simple check would sacrifice feature-based accounts of nonlocal dependencies. It is the identification of subj values in (3a) that associates *seems* with the subject demands of *raining*. If the subj features of a raising verb need only be nondistinct from those of its complement, nothing would ensure a higher verb 'inherits' subj features of a lower verb.

There is a variety of options between distinctness p0080
and identity, and there is as yet no clear consensus regarding the best general solution to the challenge that Ingria articulates. One possibility is that the properties of a head and dependent are, in effect, 'consolidated upward' and associated with the constituent formed by combining the head and dependent. On this alternative, the third structure in (1) would be associated with the sentence *he skis*. The subject-verb agreement constraint would require that subj properties of *skis* in (2b) and the properties of *he* in (2a) must both describe this structure. If, as in (4), a verb is affiliated with multiple clauses, its subj properties would be required to describe the structure associated with each clause. The properties of a subject would also be required to describe the structure associated with the clause that contains it. But nothing would enforce the compatibility of subjects in different clauses. The subj features of each verb in (3a) would likewise be required to describe the structures associated with the verb phrases under which it is embedded. However, nothing would require an embedded structure to satisfy the properties of a higher verb. Thus the subj structure associated with the matrix verb phrase in (3a) would satisfy the subj features of *raining*, but the subj structure associated with the verb phrase headed by *raining* would not satisfy the agreement demands of *seems*.

In short, a transitive relation is required to 'propa- p0085
gate' compatibility demands over a nonlocal domain, and identity is a canonically transitive relation. However, relations weaker than identity (or operations less symmetrically 'destructive' than unification) might ultimately be more appropriate for regulating grammatical dependencies within a feature-based approach. One might then ask whether token-identity is ever really needed. The constraints in (2c) and (3c) both interpret two descriptions as referring to the same structure, and identity statements are generally interpreted as requiring token identity of structures. Yet it

is by no means clear that the same compatibility demands will not be enforced if identity constraints are satisfied by type-identical structures. In effect, this amounts to moving identity out of feature structures and confining it to the constraint language of a feature-based approach. The treatment of other logical operators provides a model. Karttunen (1984: 30) introduces negative constraints on pair nodes that 'prevent the nodes from ever becoming alike.' In description-based models, a negative (or disjunctive) constraint is interpreted as describing a family of satisfying structures. There are no negative or disjunctive structures, just negative or disjunctive descriptions of structures. This type of reinterpretation can also be applied to identity, so that there are no re-entrant (token-identical) structures, only identity constraints, which are satisfied by type-identical structures.

p0090 In summary, the classification of features and the relative complexity of feature structures and constituent structures are issues that clearly distinguish feature-based approaches from transformational alternatives. Within the class of feature-based models, one can identify strategies for describing grammatical dependencies, model-theoretic perspectives on grammatical analysis, and a number of formal and empirical questions that these strategies and perspectives raise.

*See also:* Functional categories (01966); HPSG (02040); LFG (02043); Unification: Classical & Default (01994); X-bar Theory (01996).

## Bibliography

Carpenter B (1992). *The logic of typed feature structures*. Cambridge: Cambridge University Press.

Chomsky N (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky N (1970). 'Remarks on nominalization.' In Jacobs R A & Rosenbaum P S (eds.) *Readings in English transformational grammar*. Waltham: Ginn and Company. 232–286. Reprinted in *Studies on semantics in generative grammar*, 1–61. The Hague: Mouton.

Chomsky N (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Corbett G (1991). *Gender*. Cambridge: Cambridge University Press.

Gazdar G, Klein E, Pullum G K & Sag I (1985). *Generalized phrase structure grammar*. Cambridge, MA: Harvard University Press.

Harris Z S (1951). *Methods in structural linguistics*. Chicago: University of Chicago Press.

Ingria R J (1990). 'The limits of unification.' In *Proceedings of the 28th annual meeting of the Association for Computational Linguistics*. 194–204.

Johnson M (1988). *Attribute-value logic and the theory of grammar*. Stanford: CSLI Publications.

Kaplan R M & Bresnan J (1982). 'Lexical-functional grammar: a formal system for grammatical representation.' In Bresnan J (ed.) *The mental representation of grammatical relations*. Cambridge, MA: MIT Press. 173–281.

Karttunen L (1984). 'Features and values.' In *Proceedings of COLING 84*. Association for Computational Linguistics. 28–33.

King P J (1989). *A logical formalism for head-driven phrase structure grammar*. University of Manchester 9.

Matthews P H (1991). *Morphology*. Cambridge: Cambridge University Press.

Pollard C & Sag I A (1994). *Head-driven phrase structure grammar*. Stanford: University of Chicago Press and CSLI.

Shieber S M (1986). *An introduction to unification-based approaches to grammar*. Stanford: CSLI.