N

Non-transformational grammar

The class of non-transformational generative grammars comprises frameworks that share many of the broad goals espoused in early transformational work (e.g., Chomsky 1957) but use different devices to pursue these goals. This class of grammars can be divided into three principal subclasses. The family of **feature-based** approaches, also known variously as 'unification-based', 'constraint-based' or 'description-based' grammars, makes essential use of complex-valued features in the analysis of local and non-local dependencies. **Generalised phrase-structure grammar**, **head-driven phrase structure grammar** and **lexical functional grammar** are among the most important members of this class. There are two basic varieties of **relational** approaches – **relational grammar** and **arc pair grammar** – which both accord primacy to grammatical relations and relation-changing rules. The class of **categorial** approaches uses flexible category analyses and highly schematic rules to combine expressions that often do not correspond to syntactic constituents in other approaches. Categorial approaches fall into three main groups: versions of the **Lambek calculus**, **combinatory categorial grammars** and offshoots of **Montague grammar**.

This article identifies the distinctive characteristics that broadly define the three primary subclasses and summarises some significant properties and insights of individual frameworks.

Feature-based grammars

It is customary to divide feature-based grammars into 'tools' and 'theories'. The class of tools includes versions of the **PATR** formalism (Shieber 1986), along with approaches, such as **functional unification grammar** (Kay 1979), which have mainly provided a basis for **grammar implementations**. While theories such as **generalised phrase-structure grammar** (**GPSG**), **head-driven phrase-structure grammar** (**HPSG**) and **lexical-functional grammar** (**LFG**) have also been successfully implemented, these formalisms provide a more general framework for theoretical analysis.

A distinguishing property of this class of formalisms is the use of complex feature values to regulate grammatical dependencies that are attributed to constituent structure displacements in transformational accounts. The analysis of **subject–verb agreement** provides a useful illustration. The subject agreement demands of an English verb such as *walks* may be expressed by assigning *walks* a complex-valued SUBJ(ECT) feature which contains the features that represent third person and singular number. In a simple feature system, these might be [PERS 3RD] and [NUM SG].

Agreement between the 3sg verb *walks* and the 3sg subject *he* in Figure 1 is then keyed to a non-directional requirement that the SUBJ features associated with the verb must be 'compatible' with the grammatical features of its syntactic subject. The execution details of this analysis vary slightly across approaches, though in all accounts the conditions that determine subject–verb agreement refer to the features introduced by the subject and verb, not to the elements *walks* and *he*. It is the ability to refer to such features, independently of the expressions on which they are introduced, that permits feature-based

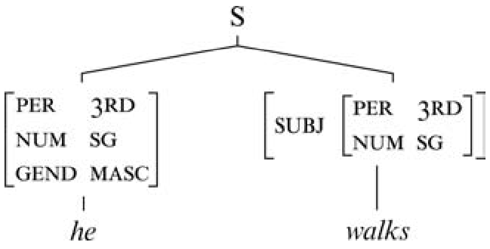

Figure 1 Subject–verb agreement.

approaches to dispense with the constituent-structure displacements that induce the 'flow' of feature information in transformational accounts.

Grammatical compatibility is usually determined 'destructively' in feature-based approaches. What this means in the present case is that the SUBJ features of the verb phrase are directly amalgamated or **unified** with the features of the syntactic subject. The result of combining two sets of compatible features is a single feature structure that contains the information from both. Unifying the features of *he* with the SUBJ features of *walks* yields a structure that just preserves the features of *he*, because these features already contain the SUBJ features of *walks*. If the input features are incompatible, unification is said to 'fail', in virtue of the fact that no consistent structure can contain conflicting values for a single feature 'path'. (The possibility of failure distinguishes unification from the formally similar set union operation.) The central role of unification in GPSG, LFG and HPSG underlies the now largely deprecated term 'unification-based grammars'.

Feature structure unification or, equivalently, structure sharing, retains a key role in most feature-based frameworks. It is nevertheless important to realise that the 'constructive' strategy of determining compatibility by actually combining the features of input structures does not in any way require a fully 'destructive' mechanism that overwrites the inputs in the process. To regulate agreement in Figure 1, we must combine the SUBJ features of *walks* and the features of its syntactic subject. It is, of course, more efficient to merge the original inputs than it is to copy their feature information and amalgamate it in another location, e.g., on the common S mother in Figure 1. Yet there is evidence that this efficiency incurs a significant descriptive cost in coordinate structures and other environments in which a single element is subject to multiple compatibility demands. The fact that such elements may satisfy incompatible demands suggests that, in at least some cases, valence and concord demands must be regulated by the non-destructive or semi-destructive mechanism suggested in recent accounts (Dalrymple and Kaplan 2000; Blevins forthcoming).

Another general issue concerns the symmetrical or non-directional character of operations such as unification. This 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 symmetrical operations can provide illuminating analyses of all of the cases that motivate the traditional distinction between agreement 'controllers' and 'targets' (Corbett 1991).

Generalised phrase-structure grammar

Although the descriptive potential of complex syntactic features is set out clearly by Harman (1963), this potential was not fully realised until the emergence of GPSG nearly twenty years later. A decisive step in the development of GPSG – and non-transformational approaches generally – was the demonstration in Gazdar (1981) that any non-local dependency that could be described in terms of transformational 'movement' rules could also be described by a local mechanism that 'passes' the features of a dislocated element successively from daughters to mothers in a phrase-structure tree. This demonstration effectively refuted long-standing claims that transformational devices were necessary for the description of non-local dependencies. The intervening decades have seen the development of a range of other non-transformational strategies (see, e.g., the discussion of **domain union**, **functional uncertainty** and **function composition** below), as well as a general recognition that derivational structure is not an intrinsic property of natural languages or of the language faculty, but rather a purely contingent property of transformational approaches.

Yet, with the benefit of hindsight, the success of the GPSG analysis of unbounded dependencies can be seen as something of a blessing and a curse. On the positive side, the discovery that

phrase-structure grammars could define structural descriptions isomorphic to those attributed to transformational devices threw open a number of issues that many linguists had taken to be settled. On the negative side, the successful use of features to mimic the effects of 'movement' rules encouraged two somewhat conservative tendencies in later GPSG work. The first was a tendency to push feature-based strategies into areas where they did not provide an illuminating analysis. The second was a tendency to use features to 'emulate' existing transformational analyses.

GPSG treatments of co-ordination display the first tendency, while analyses of passivisation illustrate the second. GPSG accounts of co-ordinate structures squarely address the problems posed by cases of unlike constituent co-ordination, such as *Max is a Guardian reader and passionate about penal reform*. In this example, the noun phrase *a Guardian reader* appears to be conjoined with the adjective phrase *passionate about penal reform*, violating the widely assumed constraint that conjuncts must be of the same category. The solution developed within GPSG assigns a coordinate mother the **generalisation** of the features of its conjunct daughters, so that *a Guardian reader and passionate about penal reform* is assigned the features from each conjunct that does not conflict with the other conjunct. GPSG accounts acknowledge that this account does not extend to cases of non-constituent co-ordination, and subsequent work suggests that a generalisation-based account also does not apply correctly to verbs with unlike valence demands. At an even more basic level, one might question the grounds for treating *a Guardian reader and passionate about penal reform* as a constituent in the first place. While the precise analysis of these constructions remains a matter of dispute, it is generally accepted that the solution is not likely to lie in an innovative strategy for combining the features associated with unlike conjuncts or non-constituent sequences.

By pushing a feature-based strategy to its limits, GPSG analyses of coordination can be seen to obtain a useful, if somewhat negative, result. GPSG treatments of passivisation in terms of **meta-rules** are perhaps best regarded in much the same way. These accounts demonstrate that the structure-to-structure mapping invoked in transformational analyses can be mimicked by a meta-rule that maps phrase-structure rules that introduce active VPs onto derived rules that introduce detransitivised passive VPs. Yet, by re-implementing the transformational analysis, GPSG accounts inherit the weaknesses of this analysis, while exposing limitations of a standard phrase-structure formalism. As LFG accounts in particular have shown, passivisation is a lexical – indeed, derivational – process, which is most insightfully expressed by analyses that relate entries, rather than structures or syntactic rules. This type of analysis is unavailable in a standard phrase-structure grammar, which represents the lexicon implicitly in the rules that rewrite preterminals. GPSG extends this conception by introducing entries that are cross-indexed with rules, though these entries still do not carry the information required for a lexicalist analysis of the passive.

GPSG accounts are arguably most successful in cases where they address a traditional issue or present an essentially new approach. For example, the GPSG **head feature convention** (or **principle**) illustrates how complex features yield an insightful treatment of traditional notions like 'endocentricity'. This principle requires that a syntactic head and the phrase that it heads must have the same values for the various 'head' features that represent part of speech and syntactically relevant inflectional properties. The inclusion of inflectional features contrasts with versions of **X-bar theory** in which parts of speech features are, without any explicit justification, singled out as the only head features. In GPSG, the features of a finite clause may be inherited from a finite verb on the assumption that clauses are endocentric verbal projections. In transformational accounts, the distribution of tense features must again involve recourse to a movement rule.

The definitive presentation of GPSG in Gazdar et al. (1985) displays some of the other insights developed in this framework, along with the attendant formal complications. A significant feature of later versions of GPSG is the decomposition of standard **phrase-structure rules** into separate **immediate dominance** (ID) and **linear precedence** (LP) constraints. This division of labour permits an elegant and often highly general description of various types of word order patterns and word order. To take just one example, the relatively free ordering of a verb and its complements in a language like Russian may be described by introducing no rule that imposes a relative order on these elements. However, the usefulness of the structure/order dissociation is severely constrained by the desire to keep the GPSG formalism within the class of **context-free grammars**. One consequence of this meta-theoretical constraint is that precedence rules must have the same domain as dominance rules and thus may not order non-siblings. This entails that the free ordering of a verb and a VP-external subject in Russian cannot be attributed simply to the lack of an applicable linear constraint. Although **liberation meta-rules** were proposed to telescope a set of rules and define essentially flat constituent structures, the use of these rules undercuts the motivation for the original structure/order division. The descriptive challenge posed by free constituent order languages was not met in a satisfactory way until the advent of **linearisation grammars** in HPSG (see below).

As is generally the case with feature-based approaches, GPSG accounts are explicitly – often painstakingly – formalised. The difficulties that this formalisation may present to contemporary readers reflect a genuine tension between the simple architecture and complex 'control structure' of GPSG. At one level, a GPSG can be viewed as a set of constraints, interpreted uniformly as 'tree licensing conditions'. Dominance rules license tree structure, precedence rules dictate the relative order of siblings, and feature constraints determine the distribution of features on non-terminal nodes. Yet this straightforward conception is complicated in GPSG by the numerous types of feature conditions and their often intricate interactions. A general source of complications is the default interpretation of conditions, such as **feature specification defaults** or, indeed, the **head feature convention**. This aspect of GPSG has not been taken up directly in other syntactic approaches, though defaults appear in a different guise in recent **optimality extensions of LFG**.

Head-driven phrase-structure grammar

HPSG is in certain respects a direct descendant of GPSG. However, it also includes features of Head Grammars (Pollard 1984), along with properties of categorial grammars and systems of feature logic. The two book-length expositions of HPSG, Pollard and Sag (1987, 1994), outline a general sign-based conception that integrates these diverse influences.

HPSG incorporates a number of evident improvements over GPSG. Foremost among these is a 'description-based' perspective that clarifies some of the representational issues that remained unresolved in GPSG. Like LFG, HPSG proceeds from a fundamental distinction between **(feature structure) descriptions**, which are sets of grammatical constraints, and the **feature structures** that actually model the expressions of a language. This distinction is clearly illustrated by the treatment of lexical entries, which are not viewed as structures, but rather as descriptions of lexical structures. Descriptions in HPSG are represented as standard **attribute-value matrices** (AVMs), similar to the bracketed 'feature bundles' familiar from phonological analyses. Structures are rarely exhibited in HPSG accounts, though they are conventionally depicted as directed acyclic graphs. The correspondence between descriptions and the structures that they describe is defined in terms of a standard satisfaction relation, as in **model-theoretic semantic** approaches. The structures that satisfy a description must, at the very least, preserve all of the information in the description, and also identify all of the values that are specified as token-identical in the description.

The interpretation of the basic HPSG formalism is thus relatively straightforward, as is the interpretation of feature distribution constraints. A distinctive aspect of HPSG is the assumption that structures are **typed**, and that types may be organised into general **type hierarchies**, in which properties may be inherited from general types to their subtypes. For example, the general type *sign* contains the subtypes *word* and *phrase*. Features common to all signs, i.e. the fact that they are associated with a phonological form, are associated with the type *sign* and inherited down to its subtypes. The properties that distinguish words from phrases are in turn associated with the corresponding subtypes. A general strategy of type-based inheritance achieves considerable concision, while eliminating some of the vagaries of the heterogeneous feature distribution conditions in GPSG. To take a simple example, the open-class categories 'noun', 'verb' and 'adjective' are represented by the *head* subtypes *noun*, *verb* and *adjective*. Properties that are only distinctive for a particular part of speech may be associated with the appropriate subtype. The features that represent tense/aspect properties or distinguish infinitives from participles are associated with the *verb* type and thereby restricted to verbs and their projections. Declensional features like case may likewise be associated with nouns and/or adjectives. Current models of HPSG extend the use of type inheritance to classes of construction types (Sag 1997). The **feature declarations** that are directly associated with a given type or inherited from a more general type then represent the features for which that type may – and in current versions of HPSG must – be specified.

Moreover, it is possible to introduce a qualified notion of 'default' within this kind of type hierarchy. HPSG-type hierarchies make use of multiple inheritance, meaning that a given type may inherit properties from different general types. This permits a maximally general cross-classification and avoids the need to introduce the same properties at different points in a hierarchy. However, multiple inheritance also raises an issue of consistency, since different general types may introduce conflicting properties. Multiple inheritance systems usually address this issue by assigning a relative priority to general types, so that one type may 'outrank' or 'take precedence over' another type. In cases of conflict, the inheritance of properties from a higher-ranking type may then pre-empt the inheritance from a lower-ranking type. Controlling the inheritance of properties in this way provides an 'offline' default mechanism that expresses a limited notion of defeasibility, while retaining a standard non-default interpretation of the constraints themselves.

In addition to these largely technical improvements, the neo-Saussurean perspective adopted in HPSG permits a highly flexible treatment of the relation between form and features. The form associated with a *sign* is represented as the value of a PHON(OLOGY) attribute, rather than by a terminal or sequence of terminals, as in other approaches. This difference is illustrated by the descriptions of the noun *book* in Figure 2 (the SUBCAT(EGORISATION) feature is described below).

While these alternatives may look rather like notational variants, the description in Figure 2(a) implicitly supports the feature–form mapping characteristic of word and paradigm (WP) models of morphology (Anderson 1992; Stump 2001). At the lexical level, a sign-based system provides the formal prerequisites for morphological analyses in which a given form is said to 'spell out' or 'realise' a particular feature combination. Further, as Ackerman and Webelhuth (1998) argue at some length, this **exponence-based** conception extends straightforwardly to a range of periphrastic constructions in which multiple words may realise a notion like 'perfect' or 'passive'.

At the level of phrasal analysis, the introduction of a *marker* type reconstructs the distinction that Hockett (1958) draws between the **immediate constituents** (ICs) of a construction, and formatives that merely serve to identify or 'mark' the construction. There is a direct parallel between the WP treatment of *-s* in *books* as a marker of plurality, rather than a morphological constituent proper, and the HPSG treatment of complementisers and coordinating conjunctions as markers of subordination and coordination, respectively, rather than defective 'functional' heads. The HPSG formalism likewise permits, in principle, a description of non-biunique patterns of exponence. To turn to a construction discussed by Hockett 1958, iterative coordinate structures, in which a coordinating conjunction is repeated before or after each conjunct, may be treated as a case of 'extended exponence' (Matthews 1974/1991) where the distinct occurrences

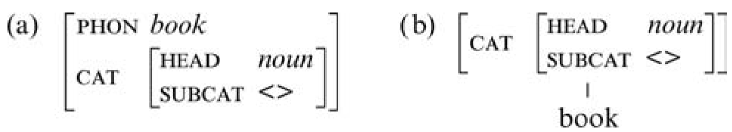

Figure 2 Lexical signs in HPSG.

of the conjunction collectively 'spell out' or 'realise' the features that represent the notion 'coordinate category'.

In sum, the simple representational shift illustrated in Figure 2 avoids a commitment to the rigid 'item and arrangement' perspective that many generative approaches have uncritically inherited from their structuralist predecessors. The basic design of HPSG also frees analyses from other, similarly anachronistic, assumptions.

Linearisation-based accounts of word order variation provide perhaps the most striking illustration. The form associated with a node in a phrase-structure tree is standardly defined as the concatenation of the terminals dominated by that node. Thus the tree in Figure 3(a) represents the sentence *He should walk*. On the conventional assumption that sister nodes are strictly ordered, it is not possible to interleave constituents that occur at different levels. In particular, there is no way to assign the subject–predicate structure in Figure 3(a) to the corresponding question *Should he walk?*

This is precisely the sort of word order alternation that American structuralists took to justify discontinuous IC analyses and which motivated non-concatenative 'wrap' operations in Head Grammars and Montague Grammar. Linearisation-based models of HPSG (Reape 1993; Kathol 2000) develop a general approach to this phenomenon in terms of independent **word order domains**. In the default case, the DOM(AIN) of a phrase is just a list containing its daughters, so that the form or 'yield' of the phrase is defined in much the same way as for a phrase-structure tree. However, by allowing daughters to pass up their DOM values to their mother, linearisation grammars also make it possible to interleave or 'shuffle' non-siblings. The intuition underlying these approaches can be illustrated with reference to Figure 3(b). To simplify this illustration, DOM values are assumed to be lists of signs, as in Reape (1993). The boxed integer 'tags' in Figure 3(b) represent token identity and indicate that the DOM value of the VP contains its actual V and NP daughters. Precedence constraints apply to DOM elements, determining a sequence whose order defines the relative order of PHON elements.

The yield of the S in Figure 3(b) thus depends on how its DOM list is defined. If this list contains the daughters of S, [1] and [4], it will only be possible to concatenate *he*, the yield of the subject daughter, to the yield of the predicate, i.e. the entire string *should walk*. However, if the VP in Figure 3(b) instead passes up its own DOM value, the DOM value of the S will contain the elements [1], [2] and [3]. This expanded domain '**unions**' the subject into the domain of the predicate. Precedence constraints that place the head initially in this domain will determine the list ([2], [1], [3]). Concatenating the yields of these elements produces the 'inverted' order *should he walk*.

The dissociation of structure and order illustrated in Figure 3(b) likewise accommodates the free ordering of a subject and VP-internal object in Russian, which was identified above as a problem that defied analysis in GPSG. While these cases are both extremely local, linearisation approaches provide a general mechanism for describing constituency-neutral ordering variation. Reape (1993) and Kathol (2000) present analyses of the ordering freedom characteristic of the *Mittelfeld* in German, while recent extensions extend a linearisation approach to cases of scrambling (Donohue and Sag 1999) and extraction (Penn 1999). Linearisation accounts thus permit a simple and uniform

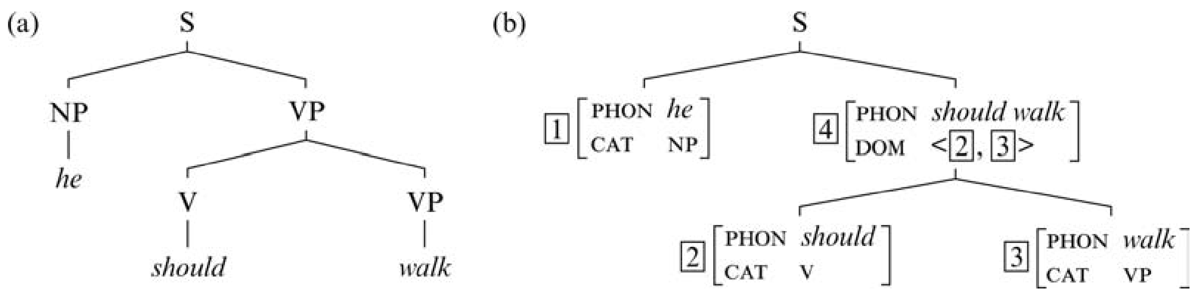

Figure 3 Linearisation of order domains.

treatment of hierarchical structure within HPSG, avoiding the spurious structural variation characteristic of transformational and some categorial approaches. Yet the introduction of word order domains also potentially undermines the feature-based technology for handling word order variation, including the feature-based account of unbounded dependencies.

An aspect of HPSG that reflects the influence of categorial approaches is the treatment of valence. The initial version of HPSG in Pollard and Sag (1987) introduced a single SUBCAT feature that consolidated all of the subcategorised arguments of a head. Pollard and Sag (1994) subsequently distinguished separate SUBJ(ECT) and COMP(LEMENT)S lists, while retaining an argument structure list, ARG-S, as a lexical counterpart of the SUBCAT list. Some current versions of HPSG add a further DEP(ENDENT) S list to integrate grammatical dependants that are neither subjects nor complements.

A significant difference between argument structure and valence features is that the elements of SUBJ and COMPS lists are removed or 'cancelled' as syntactic arguments are encountered. Thus the transitive verb *hit* begins with the singleton SUBJ list and singleton COMPS lists in Figure 4, which each contain an element from the ARG-S list. The VP *hit Max* retains a singleton SUBJ list, but has an empty COMPS list, signifying that it does not select any further complements. The S *Felix hit Max* has both an empty SUBJ and COMPS list, signalling that it is fully 'saturated'.

The tags on the syntactic subject and object in Figure 4 indicate that the features of these arguments are shared or, in effect, unified with the corresponding valence elements.

The flow of feature information represented in Figure 4 highlights the strongly 'head-driven' nature of some versions of HPSG. The head in Figure 4 functions as the ultimate repository of the grammatical information in this sentence, since the features of the verb and its arguments are consolidated in the ARG-S value. In contrast, the projections of the verb become progressively less informative as elements are popped off their valence lists. The 'head-directed' flow in Figure 4 thus represents the transitivity of a head, while tightly restricting access to information about 'cancelled' arguments.

In addition to the properties discussed above, HPSG signs also represent constituent structure in terms of DAUGHTERS attributes. It is nevertheless common for HPSG analyses to be expressed informally as annotated tree structures, as in Figure 3(b) and Figure 4. Semantic and pragmatic information is also expressed via CONTENT and CONTEXT attributes. Yet the empirical consequences of bundling this disparate information together in a single data structure are not always obvious. The non-syntactic properties in signs rarely show significant interactions with grammatical processes, such as sub-categorisation. Agreement features, which HPSG accounts introduce as part of the CONTENT, are an exception, though these features are more traditionally regarded as syntactic.

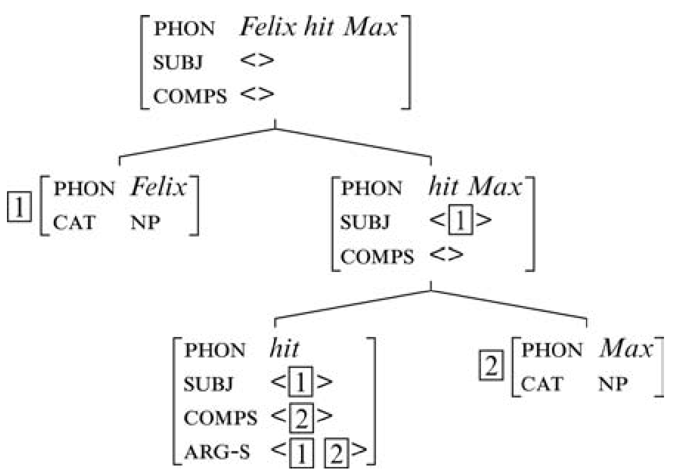

Figure 4 Valence and argument structure in HPSG.

Lexical-functional grammar

In some regards, LFG straddles the classes of feature-based and relational approaches. On the one hand, the lexicalist and description-based framework outlined in Kaplan and Bresnan (1982) is close to the perspective subsequently adopted in HPSG, though there is also a number of significant respects in which these approaches diverge. At the same time, the analyses developed in Bresnan (1982a) and subsequent work show an affinity with relational accounts, both in the importance they attach to grammatical functions and in their comparatively broad typological coverage.

LFG exhibits a clean formal architecture, with well-defined interfaces between levels of representation. A unique aspect of LFG is the separation between **c(onstituent)structures**, which represent category and ordering information, and **f(unctional) structures**, which represent the features that represent valence properties and feed semantic interpretation. The c-structure in Figure 5(a) and the f-structure in Figure 5 (b) express the analysis assigned to *Felix hit Max*.

The functional annotations in Figure 5(a) define the correspondence between c-structure nodes and their f-structure counterparts in Figure 5(b). The equation '↑ = ↓' expresses the LFG counterpart of the **head feature principle** by associating the V, VP and S nodes with the same f-structure in Figure 5(b). This shared f-structure is the complete or 'outermost' f-structure in Figure 5(b). The equations 'SUBJ' and 'OBJ' unify the properties of the syntactic subject and object in Figure 5(a) into the values of the SUBJ and OBJ attributes in Figure 5(b).

The structures in Figure 5 are defined by annotated phrase-structure rules in conjunction with the lexical entries for the items *Felix*, *hit* and *Max*. The rules in Figure 6(a) determine the tree in Figure 6(a). The entry in Figure 6(b) likewise represents the properties of the verb *hit*.

The category symbol 'V' specifies the preterminal mother of *hit* in Figure 5(a). The functional equations in Figure 6(b) are both satisfied by the f-structure in Figure 5(b). The TENSE feature specified in Figure 6(b) is obviously present in Figure 5(b), as is the PRED value. The LFG **completeness** and **coherence** conditions, which are keyed to PRED features, are also satisfied in Figure 5(b). Informally, an f-structure is **complete** if it contains all of the grammatical functions governed by its predicate and **coherent** if all of its governable grammatical functions are governed by its predicate. Governable functions are essentially those that can be selected by a predicate. The functions governed by the predicate 'hit < (SUBJ)(OBJ) >' are just SUBJ and OBJ. Since exactly these functions are present in Figure 5(b), the f-structure is complete and coherent.

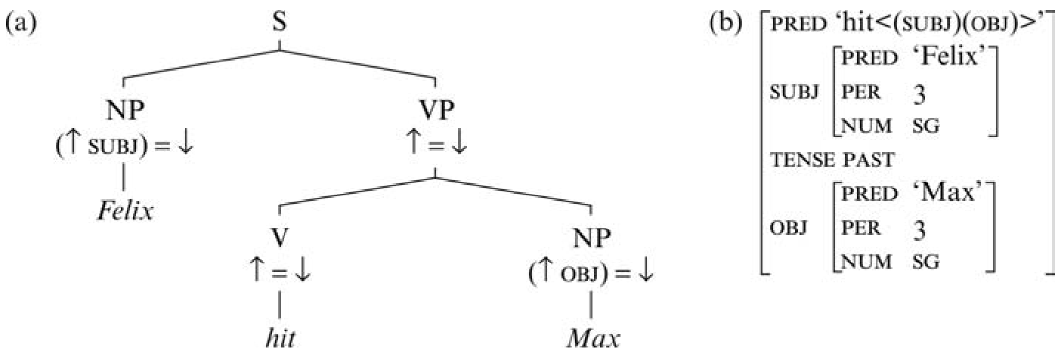

Figure 5 LFG c-structure and f-structure analysis.

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

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

(b) *hit*: V, $(\uparrow \text{TENSE}) = \text{PAST}$
$(\uparrow \text{PRED}) = \text{'hit<(SUBJ)(OBJ)>'}$

Figure 6 Annotated phrase-structure rules and lexical entry.

The analyses in Figure 5 highlight some important contrasts with GPSG and HPSG. One unfortunate notational difference concerns the interpretation of AVMs. HPSG accounts use AVMs as a convenient graphical representation of descriptions, i.e. as sets of constraints. LFG interprets AVMs like Figure 5(b) as structures that provide the solution to a set of constraints.

The role of annotated phrase-structure rules in LFG reflects a more substantive difference. The separation of order and structure in GPSG and HPSG reflects an interest in unbundling the different types of information expressed by phrase-structure rules. The addition of functional annotations moves in precisely the opposite direction, by incorporating a further sort of information into phrase-structure rules. The use of an augmented phrase-structure formalism has a number of formal advantages, though it also severely constrains the role of constituency relations. Thus in interleaved constructions, such as Germanic cross-serial dependencies, a verb and its complements cannot form a syntactic constituent. Instead, these elements are introduced on parallel c-structure 'spines' and only associated in the corresponding f-structure. The c-structures proposed for cross-serial dependencies in Bresnan et al. (1982) exhibit other remarkable properties, including verb phrases that consist entirely of noun and prepositional phrases. The patently expedient nature of these c-structures clearly signals the diminished importance of constituent structure in LFG.

Indeed, LFG c-structures are in many respects closer to the derivational structures of a categorial grammar than to the part–whole structures represented by **IC analyses**. Much as derivational structures are essentially by-products, produced in the course of deriving semantic representations, c-structures are the by-product of deriving f-structures in LFG. In versions of LFG that introduce a notion of **functional precedence** (Bresnan 1995; Kaplan and Zaenen 1995), c-structures do not even retain their original role as the unique locus of ordering relations and constraints.

The centrality of grammatical functions is another distinctive property of LFG, one which has contributed to highly influential analyses of relation-changing rules. Beginning with the analysis of the passive in Bresnan (1982b), LFG accounts have succeeded not only in establishing the viability of lexicalist analyses, but often in showing the essential correctness of such analyses. The influence of these analyses is perhaps most obvious in the treatment of passivisation and other lexical rules in HPSG (Pollard and Sag 1987). The structure-neutral analyses proposed in relational approaches likewise strongly suggest a lexical reinterpretation. Moreover, the ultimately lexical basis of relation-changing rules is also tacitly conceded in transformational accounts that invoke a morphological operation to detransitivise a verb by 'absorbing' its case or thematic properties.

While the locus of relation-changing rules has remained constant in LFG, the form of these rules has undergone significant changes. This evolution is reflected in the contrast between the treatments of passive represented in Figures 7–9. Figure 7 summarises the analysis in Bresnan (1982b), while Figure 9 outlines the lexical mapping approach of Bresnan and Kanerva (1989).

The form for *bite* in Figure 7(b) identifies the mapping between argument structure, thematic structure and grammatical functions characteristic of a transitive verb. The rule in Figure 7(a) applies to this lexical form, and defines the derived form in Figure 7(c). The first operation in Figure 7(a) suppresses arg1, which is lexically associated with the agent role, by reassigning arg1 the null grammatical function 'Ø'. This determines a 'short' passive in which the agent is not realised. The second operation in Figure 7(b) 'promotes' arg2 by reassigning it the SUBJ function.

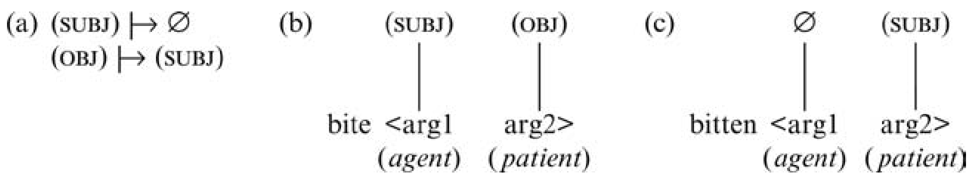

Figure 7 Passivisation by lexical rule.

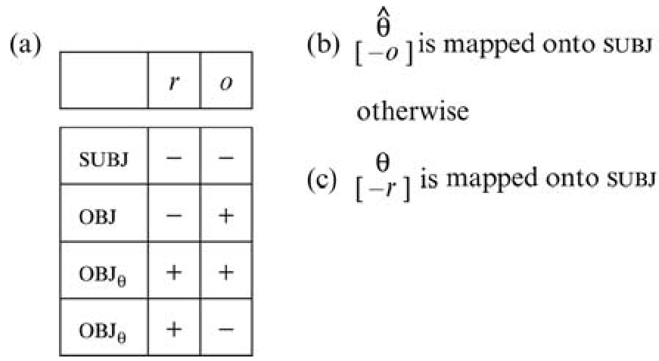

Figure 8 Argument classification and subject mapping principles.

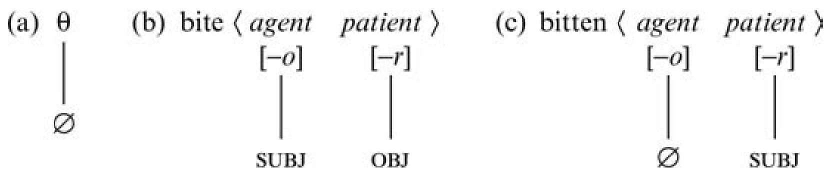

Figure 9 Passive via thematic suppression.

Given the completeness and coherence conditions, the form in Figure 7(c) determines an f-structure whose only governed function is a SUBJ which is associated with the patient role. The alternation between the forms in Figure 7(b) and (c) thus expresses the relation between active sentences such as *Cecilia bit Ross*, and corresponding passives such as *Ross was bitten*. More recent work in LFG has refined this analysis in the context of what is known as **lexical mapping theory** (LMT). The main prerequisites of LMT are set out in Figure 8.

The features [*r(estricted)*] and [*o(bjective)*] cross-classify the governable grammatical functions in Figure 8(a). These features then guide the mapping principles in Figure 8(b) and (c), which link up the subject with a semantic role. The role $\hat{\theta}$ designates the highest thematic role of a predicate, which is usually taken to be defined with reference to a universal thematic hierarchy. The principle in Figure 8(b) associates the highest role with the SUBJ function. If the highest role is not available, the principle in Figure 8(c) maps an unrestricted role on to the SUBJ.

The configuration in Figure 9(a) represents the LMT counterpart of the lexical rule in Figure 9(b). This mapping associates $\hat{\theta}$ to the null function Ø, thereby pre-empting the principles in Figure 8 and determining the contrast between the argument structures in Figure 9. The active structure in Figure 9(b) conforms to the principle in Figure 8(b), as the SUBJ is mapped onto the highest role, the agent role. In the passive structure in Figure 9, the agent is 'suppressed' or unavailable by virtue of its association to Ø. Hence the SUBJ is linked to the unrestricted patient role, in conformance with the mapping principle in Figure 8(c).

The LMT account in Figure 9 differs from the lexical rule analysis in two main respects. First, the LMT analysis uses monotonic (albeit conditionalised) mapping principles, in place of nonmonotonic attribute changes. More strikingly, suppression does not refer to subjects, but instead targets the highest thematic role. This shift implicitly rejects the traditional view, developed in greatest detail in relational approaches, that passivisation is restricted to verbs that select subjects.

These assumptions must of course be understood in the context of the larger LMT programme, and its ambitious goal of mapping out the correspondences between grammatical functions and lexical semantics. Nevertheless, one can question whether either of the revisions incorporated in the LMT analysis contributes to an improved treatment of passives. It is, for example, not altogether clear why monotonicity should be regarded as a desirable property of derivational rules, given that derivational processes are to a great degree defined by their non-monotonic, feature-changing, character. The benefits of a thematic role-based notion of suppression are similarly open to question. The rationale for this change rests on a number of prima facie cases of passives of **unaccusative** predicates (discussed in more detail in connection with relational approaches, below). Since unaccusative predicates, by definition, have no subject to target, a subject-sensitive passive rule cannot apply correctly to these cases. Yet the existing literature hardly considers the alternative, advocated by Postal (1986), that these cases involve impersonal rather than passive constructions, and thus are not directly relevant. Moreover, even a cursory examination of some of the 'passive' constructions in question suggests that they are equally problematic for role-based accounts. For example, the celebrated Lithuanian passive freely applies to 'weather' verbs (Ambrazas 1997), which are not standardly associated either with subjects or with thematic roles.

Contemporary work takes LFG in a number of different directions. One line of research involves incremental, carefully formalised, extensions to the original LFG formalism. Typical of this work is the f-structure treatment of extraction in terms of **functional uncertainty** (Kaplan and Maxwell 1995). In effect, this device identifies a dislocated TOPIC function with an *in situ* grammatical function GF by means of a regular expression of the form '(↑TOPIC) = (↑COMP* GF)'. A separate line of research explores more radical extensions that integrate ideas from **optimality theory** [*see* OPTIMALITY THEORY]. Bresnan (2000) provides a good point of entry into this literature.

Relational grammar

Relational grammar (**RG**) was initially developed in the mid-1970s by David Perlmutter and Paul Postal as a relation-based alternative to the highly configurational transformational accounts of that period. The three volumes of *Studies in Relational Grammar* (Perlmutter 1983; Perlmutter and Rosen 1984; Postal and Joseph 1990) provide a good survey of work in RG until the late 1980s, and display the descriptive detail and typological breadth that is typical of much of the work in this tradition. The insights developed in this framework have been highly influential and have often been directly integrated into other frameworks. The range of phenomena analysed within RG likewise provides a useful empirical 'test suite' for the validation of other approaches.

RG incorporates two distinctive claims. The first is that grammatical relations are primitive constructs that cannot be defined in terms of phrase-structure configurations, morphological cases, thematic roles, or any other properties. RG recognises two classes of grammatical relations. The core relations are referred to as **terms** and designated by integers. Subjects are designated as '1s', direct objects as '2s', and indirect objects as '3s'. Term relations correspond to the elements of an ARG-S list in HPSG or unrestricted functions in LFG. There is also a distinguished non-term relation, the **chômeur** relation, which is assigned to an element that becomes 'unemployed' by the advancement of another. This relation has no direct counterpart in non-relational approaches.

The second basic claim is that grammatical systems are intrinsically multistratal, consisting of multiple syntactic levels at which expressions may be assigned distinct grammatical relations. Strata are subject to a variety of well-formedness conditions, usually stated in the form of 'laws'. Among the important laws are the **Stratal Uniqueness Law**, which allows at most one subject, object and indirect object; the **Final 1 Law**, which requires a subject in the final stratum; and the **Motivated Chômage Law**, which prevents elements from 'spontaneously' becoming chômeurs.

Grammatical descriptions in RG take the form of **relational networks** that represent the relations associated with an expression at different strata. The network associated with *Cecilia bit Ross* in Figure 10(a) illustrates the limiting case in which there is no change in relations. The arc labelled 'P' identifies the verb *bit* as the predicate of the clause. The '1 arc' likewise identifies *Cecilia* as the subject (i.e. the 1) while the '2 arc' similarly identifies *Ross* as the direct object.

Changes within a relational network provide a general format for expressing relation-changing processes such as passivisation or causativisation. These changes fall into two basic classes: **advancements**, which assign an element a higher-ranking relation, and **demotions**, which assign a lower-ranking relation. For example, passive is analysed as a case of a 2 → 1 advancement, in which an initial object becomes a final subject, thereby forcing the initial subject into chômage. This view of the passive is represented in the analysis of *Ross was bitten by Cecilia* in Figure 10(b).

In the initial stratum at the top of Figure 10(b), *Cecilia* and *Ross* bear the same grammatical relations as in the active clause in Figure 10(a). In the second and final stratum, *Ross* is advanced to subject, represented by the fact that it 'heads' the '1 arc'. Given the Stratal Uniqueness Law, *Cecilia* cannot also remain a 1 and thus must become a chômeur, heading the 'Cho arc'.

The multistratal perspective illustrated in this treatment of the passive also underlies the **unaccusative hypothesis** (UH), which represents one of the lasting contributions of RG. In effect, the UH sub-classifies predicates according to the initial grammatical relation associated with their subjects. Predicates whose final subjects are also initial subjects are termed **unergative**. The transitive verb *bit* in Figure 10(a) is unergative, as are intransitive verbs like *telephone* or *ski*. In contrast, predicates whose final subjects are initial non-subjects are termed **unaccusative**. This class is canonically taken to include intransitives like *exist*, *vanish*, *disappear*, *melt*, *faint*, etc. RG accounts also extend this class to include semi-transitive predicates such as *last* and *weigh*.

The networks in Figure 11 illustrate the advancement of non-subjects in initially unaccusative clauses. In the intransitive structure in Figure 11(a), representing *The manuscript vanished*, *the manuscript* is the direct object in the initial stratum and is advanced to subject in the final stratum. In Figure 11(b), representing *The concert lasted an hour*, *the concert* is analysed as an initial oblique, which heads the oblique GRx arc in the initial stratum. This oblique is advanced to subject in the final stratum, while *an hour* is an object in both strata (Perlmutter and Postal 1984).

A striking property of unaccusative predicates is their resistance to passivisation. Neither *last* nor *weigh* may be passivised in English, and the counterparts of *vanish* or *exist* tend to resist passivisation in languages that may otherwise form passives of intransitive verbs. Perlmutter and Postal took the robustness of this pattern as evidence of a universal constraint on advancement. Their 1-Advancement Exclusiveness Law (1AEX) had the effect of barring multiple advancements to subject in a single clause. Passives of unaccusative would violate the 1AEX, by virtue of the fact that they would involve both unaccusative and passive advancement in a single clause.

As mentioned in connection with the LMT treatment of passive in LFG, the factual basis of the 1AEX has subsequently come under scrutiny. Even if we were to assume that the putative counterexamples are not misanalysed, the observation that unaccusatives resist passivisation describes a highly pervasive pattern. This pattern would seem to call for some principled explanation.

One particular alternative, raised but subsequently rejected in the RG literature, is worth reviewing for the insight it lends to this framework. The naive reader might at first wonder

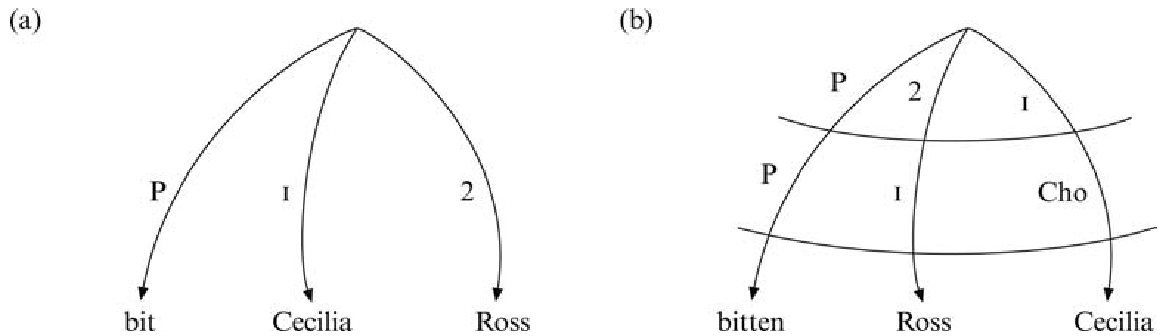

Figure 10 Active and passive relational networks.

why the 1AEX is needed at all in RG. If passivisation demotes initial subjects to chômeurs, and only unergative predicates have initial subjects, surely it follows directly that there can be no passives of unaccusatives? Further, as noted by Comrie (1977), an analysis along these lines applies to personal and impersonal passives, yielding a simple and uniform treatment of passive constructions.

Alas, however, this account runs foul of the Motivated Chômage Law (MCL), since the initial subject of an intransitive must go into chômage 'spontaneously', not as the result of an antecedent advancement to subject. One might have expected this conflict to lead to a reassessment of the MCL, along with other laws, such as the Final 1 Law, which disallows genuinely impersonal (i.e. subjectless) constructions. Instead, Permutter and Postal mounted a spirited and ultimately successful defence of the MCL. The arguments advanced in support of the MCL featured a number of ingenious and innovative strategies, including the advancement of invisible 'dummy' objects to force subjects of unergative intransitives into chômage. However, the defence of the MCL was something of a pyrrhic victory.

The MCL and Final 1 Law were upheld, and with them an intrinsically promotional treatment of the passive. Yet this orthodoxy was maintained at great cost. A general and largely theory-neutral treatment of passives was discarded, at a time when most competing approaches were only beginning to register the existence of impersonal passives. The analyses adopted in RG to preserve the MCL also contributed to the alienation of linguists, such as Comrie, who were sympathetic to the goals of RG, but were more interested in broad-based description and analysis than in the interactions of increasingly theory-internal relational laws.

The treatment of passivisation and unaccusativity in RG illustrates a tendency within this framework to express fundamental, theory-neutral, insights in terms of a highly idiosyncratic and often inscrutable system of formal laws and principles. This tendency reaches its apogee in the closely related Arc Pair Grammar (APG) framework (Johnson and Postal 1980). APG shows more attention to formal detail than RG, facilitating comparisons with other non-transformational approaches. For example, the notion of 'overlapping arcs' proposed in Johnson and Postal (1980) corresponds quite closely to structure sharing in HPSG, and to the identity implicated in functional control in LFG. APG analyses likewise provide a distinctive perspective on issues of broad relevance, as in the case of the impersonal re-analysis of the passive constructions in Postal (1986). Unfortunately, these analyses tend to be formulated in an extremely uncompromising fashion, confronting the reader with an often impenetrable thicket of definitions and examples, illustrated or, at any rate, accompanied by, whimsically labelled and exotically annotated diagrams.

Nevertheless, the range of analyses developed in RG and APG provide a sustained argument for an intrinsically relational and multistratal perspective. This perspective also casts interesting light on the goals and methods of more structurally oriented approaches. For example, the transformational claim that constructions are mere 'epiphenomena' or 'taxonomic artifacts' (Chomsky 1995) makes perfect sense from the standpoint of RG. If the essential properties of constructions are indeed relational, it is only to

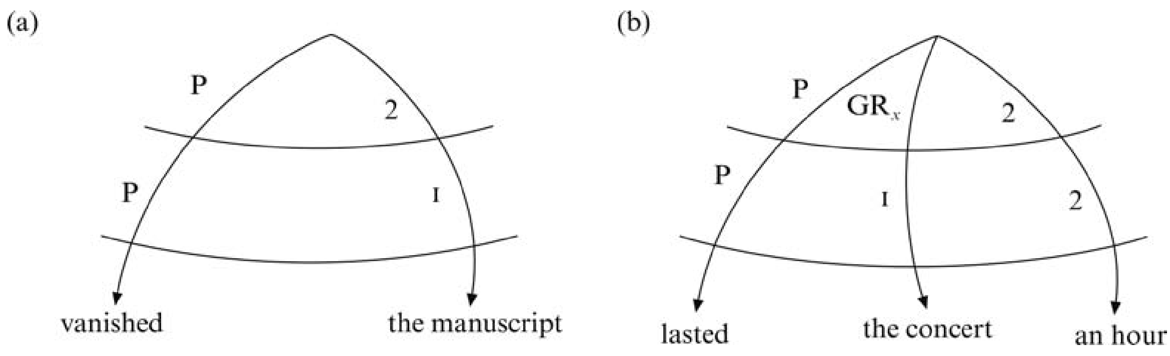

Figure 11 Unaccusative advancement.

be expected that analyses that make almost exclusive reference to features of form and arrangement will never yield a unified account of passive constructions.

Conversely, the lexicalist orientation of other non-transformational approaches suggests a basis for the strikingly non-structural character of RG analyses. Although these analyses are quite explicitly presented as syntactic, they conspicuously suppress all but the most superficial features of form and arrangement. In effect, the properties just suffice to associate the elements in a relational network with expressions in the clause it represents. One might of course regard RG as merely underspecified or incomplete in these regards. However, a more principled explanation can be obtained by reinterpreting RG as a covert theory of lexical alternations, in which grammatical relations are associated with the argument positions specified by a predicate, rather than with the syntactic arguments that ultimately fill those positions. The lack of configurational properties then follows from the fact that such properties are simply not defined in the lexical entries of predicates.

The strata in RG can likewise be associated with the lexical levels or strata assumed by nearly all approaches to morphology. A standard distinction between derivational stems and inflectional words provides morphological counterparts of initial and final strata. Where there is evidence for intermediate strata, these can be imported from approaches that recognise further lexical levels. Multistratalism thus does not require the notion of a syntactic derivation, and the derivational interpretation of RG is perhaps best regarded as a legacy of its transformational origins.

Categorial grammar

Categorial grammars are in some respects the most venerable systems of formal analysis, deriving originally from the proposals of Ajdukiewicz (1935), particularly as these were developed in Bar-Hillel (1953) and Lambek (1961). A central feature of categorial systems is the assignment of expressions to **functor** and **argument** categories, and the use of a general rule of function application to combine functors with their arguments. Ajdukiewicz postulated two basic categories – 'sentence' and 'name'. All functor categories are non-basic, defined ultimately in terms of basic categories. Intransitive verbs or verb phrases are assigned the functor category s/n, denoting a function that applies to a name and yields a sentence. A transitive verb is likewise assigned the category, denoting a function that applies to a name to yield an intransitive verb phrase.

The combination of functors and arguments is sanctioned by highly general rules. The formulation of these rules depends on the interpretation of the slash '/' used to represent functor categories, a notational point on which there is no general consensus across different approaches. To facilitate the comparison of alternatives, this entry adopts the convention '*result/argument*', in which arguments occur uniformly to the right of the slash and results to the left. This convention is followed by the category s/n, in which the name n is the argument and s is the result. The general rules of function application in Figure 12 allow a result x to be derived from the combination of a functor x/y with its argument, y occurring in either order.

For the sake of illustration, let us assign *Cecilia* and *Ross* the category n, *walks* the category s/n, and *bit* the category $(s/n)/n$. Then *Ross walks* will be of category s, the result of combining the functor *walks* with the argument *Ross*. The expression *bit Ross* will be of category s/n, the result of combining the $(s/n)/n$ functor *bit* with *Ross*. Combining this functor with the argument *Cecilia* yields the result *Cecilia bit Ross*, which is again of category s. These examples highlight one of the sources of complex slash notations. The simple convention adopted here does not specify the relative order of functors and arguments and thus fails to represent the fact that English verbs generally precede their objects and follow their subjects in declarative clauses.

There is a transparent correspondence between simple categorial systems and standard **phrase-structure grammars**. As a consequence,

(a)	$x/y\ y \Rightarrow x$	rightward or 'forward' application
(b)	$y\ x/y \Rightarrow x$	leftward or 'backward' application

Figure 12 Rules of function application.

categorial grammars were for a while regarded as notational variants of phrase-structure systems, and thought to suffer from the same descriptive limitations ascribed to standard phrase-structure systems. However, the various extended categorial formalisms have clarified some distinctive aspects of categorial systems and analyses. Reflecting their roots in logic and mathematics, categorial grammars represent a distinctively deductive approach to linguistic analysis. The derivation of a sentence is, in effect, a proof, in which lexical category assignments serve as premises and function application rules sanction the inference of a result. Although similar sorts of remarks apply, in a general way, to phrase-structure systems, the deductive structure of these systems plays no grammatical role. The grammatically significant output of a phrase-structure system consists of the trees that are defined, directly or indirectly, by its phrase-structure rules. In contrast, there is no 'native' notion of constituency defined by categorial systems, and it is often the inferential structure of such systems that is of primary importance.

This is especially true of the Lambek Calculus (Moortgat 1988; Morrill 1994), which represents one of the purest deductive systems applied to the task of linguistic description. Of particular importance in this system are rules that permit the inference of higher-order functors. The **type-raising** rule in Figure 13(a) raises an expression of any category *x* into a higher-order functor, which applies to an argument of category x/y and yields a result of category *y*. The rule of **division** in Figure 13(b) likewise divides the elements of a functor by a common category *z*.

To clarify the effect of such rules, let us apply type raising to the expression *Ross*, substituting *s* for *y* in Figure 13(a). Since *Ross* is initially assigned the category *n*, the raised functor is of category $s/(s/n)$, a functor from intransitive verbs to sentences. This analysis permits an alternative derivation of the sentence *Ross walks* in which *Ross* is the functor and *walks* is its argument. Moreover, *walks* can also undergo type raising, yielding the higher-order function $s/(s/(s/n))$. This functor applies to type-raised arguments like *Ross* and restores the function–argument relations determined by the original category assignments. The process of categorial 'ratcheting' can be continued indefinitely, yielding an infinite number of derivations of the sentence *Ross walks*. This property of categorial systems with flexible type-assignment rules is sometimes termed the 'spurious ambiguity' problem, since there is no semantic difference between analyses.

Nevertheless, higher-order types may permit new combinations, notably in conjunction with rules of **function composition**. The rules in Figure 14 allow two functors *f* and *g* to form a composed functor, *fog*, which applies to the argument of *g* and yields the result of *f*.

The interaction of type raising and composition is explored most systematically in combinatory categorial grammar (Steedman 1996), in which these devices form the basis of a variable-free treatment of extraction. The basic idea is that a chain of composed functors can 'pass along' information about an extracted element. This analysis can be illustrated with reference to the embedded question in *I wonder* [*who Cecilia has bitten*]. Let us first assign *bitten* the transitive verb category $(s/n)/n$, and provisionally assign the auxiliary *has* the category $(s/n)/(s/n)$, denoting a function from verb phrases to verb phrases. The function application rules in Figure 12 provide no means of combining these elements with the element *Cecilia*. However, if *Cecilia* is assigned the raised type category $s/(s/n)$, Figure 14(a) will sanction the composed functor *Cecilia has*, which is also of category $s/(s/n)$.

This functor can in turn compose with *bitten*, yielding the functor *Cecilia has bitten*. This functor is of category s/n, i.e. a sentence with a missing argument. By combining type raising and composition in this way it is possible to propagate information about a missing element across an unbounded domain, to the point at which its 'filler' – *who*, in this case – occurs.

(a)	$x \Rightarrow y/(y/x)$	type raising or 'lifting'
(b)	$x/y \Rightarrow (z/x)/(z/y)$	division or 'Geach's rule'

Figure 13 Category inference rules.

(a)	$x/y \;\; y/z \Rightarrow x/z$	rightward or 'forward' composition
(b)	$y/z \;\; x/y \Rightarrow x/z$	leftward or 'backward' composition

Figure 14 Rules of function composition.

This simple example illustrates the important point that categorial derivations may contain sequences that do not correspond directly to units in constituency-based grammars, though analogues of composition are employed in some versions of HPSG. A rather different departure from standard models of constituency is characteristic of the syntactic component of Montague grammars. In contrast to the rigidly concatenative Lambek and combinatory systems, the syntactic fragments developed within the Montague tradition (Bach 1980; Dowty 1982; Jacobson 1987) propose non-concatenative 'wrap' operations to describe syntactically discontinuous constructions. For example, wrap operations permit an analysis of the verb–particle construction *put the rabbit out*, in which the object *the rabbit* is interposed between the parts of a complex transitive verb, *put out.* Similar analyses are applied to resultatives, ditransitives and various other constructions that, in one way or another, resist analysis in terms of a rigidly continuous syntactic description.

These analyses exploit a general distinction between syntactic rules and the combinatory operations that they perform. The function application rules in Figure 12 concatenate adjacent functors and arguments, though they could just as well be formulated to attach an argument to the head of a complex functor. The categorial effect of the rule would be the same; only the form of the derived expression would change. Although this distinction is of considerable linguistic interest, it is largely independent of the core deductive properties of categorial systems. Hence, contemporary categorial approaches have tended to standardise on Lambek or combinatory systems. On the other hand, the contrast between rules and operations corresponds to an important distinction between dominance and precedence constraints in GPSG and HPSG. Hence it is in linearisation approaches that one sees the clearest development of syntactic insights from Montague grammar.

General remarks

Space constraints preclude a comprehensive discussion or even an exhaustive list of related approaches. Nevertheless, it is appropriate to mention a couple of frameworks that are of particular relevance to those described above. **Tree Adjoining Grammars** (**TAGs**; Joshi and Schabes 1996) introduce a distinction between initial and auxiliary trees that effectively isolate the recursive component of a phrase-structure grammar. In addition to their use for primary description and analysis, TAGs provide a 'normal form' for investigating other grammar formalisms. For example, the formal properties of 'weakly context sensitive' formalisms, such as head grammars or combinatory categorial grammars, can often be determined by translating or 'compiling' these formalisms into a corresponding TAG whose properties have been or can be established. Models of construction grammar (Kay and Filmore 1999) can also be seen to complement other constraint-based approaches, though in a more empirical way, by supplying fine-grained lexical analyses that extend high-level descriptions of possible constructions or construction inventories.

The literature on non-transformational approaches now includes basic text books for each of the major feature-based grammars (Borsley 1996; Sag and Wasow 1999; Bresnan 2001), along with overviews (Sells 1985) and compilations (Borsley and Börjars forthcoming). These sources provide a useful entry point for linguists looking to investigate this family of approaches.

J. P. B.

Suggestions for further reading

Borsley, R.D. (1996) *Modern Phrase Structure Grammar*, Oxford: Blackwell.

Borsley, R.D. and Börjars, K. (eds) (forthcoming) *Nontransformational Grammar*, Oxford: Blackwell.

Bresnan, J.W. (2001) *Lexical-Functional Syntax*, Oxford: Blackwell.

Sag, I.A. and Wasow, T. (1999) *Syntactic Theory: A Formal Introduction*, Stanford, Calif.: CSLI Publications.

Sells, P. (1985) *Lectures on Contemporary Syntactic Theories*, Stanford, Calif.: CSLI Publications.