

Oxford Handbooks Online

Case in Lexical-Functional Grammar

Miriam Butt

The Oxford Handbook of Case

Edited by Andrej L. Malchukov and Andrew Spencer

Print Publication Date: Nov 2008 Subject: Linguistics, Morphology and Syntax

Online Publication Date: Sep 2012 DOI: 10.1093/oxfordhb/9780199206476.013.0005

Abstract and Keywords

The treatment of case has been one of the central concerns within lexical-functional grammar (LFG) since its inception in the late 1970s. LFG separates facts about linear word order and constituency from the functional dependency structure analysis of a clause. A sophisticated analysis of cross-linguistic case patterns only became available as LFG's linking theory (known as lexical mapping theory) evolved. In particular, once it was recognised on the basis of argumentation by Rappaport (1983) that argument structure needed to be posited as a level of representation that was independent of constituent structure, the way was paved for analyses of case to be stated in terms of generalisations over a(argument)-structure. That is, in terms of generalisations that take both semantic and syntactic factors into account. This article first presents some LFG basics, then briefly charts the development of linking theory and discusses current theories of case. It also considers grammatical relations and grammatical functions, constructive case, and differential case theory. It concludes with a consideration of LFG-based analyses that are couched within optimality theory.

Keywords: case, lexical-functional grammar, linking theory, optimality theory, argument structure, grammatical relations, grammatical functions, constructive case, differential case theory

4.1 Introduction

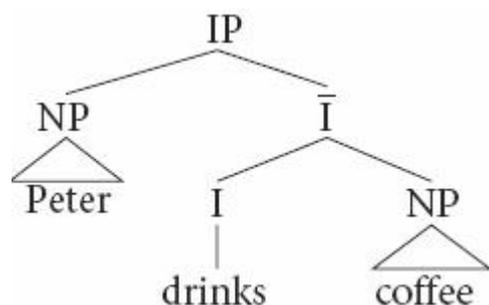
THE treatment of case has been one of the central concerns within Lexical-Functional Grammar (LFG) since its inception in the late 1970s. Several of the papers collected in the seminal book marking the emergence of LFG (Bresnan 1982) deal with case marking in particular (e.g. Neidle on Russian, K. P. Mohanan with respect to Malayalam). However, a sophisticated analysis of cross-linguistic case patterns only became available as LFG's Linking Theory (known as Lexical Mapping Theory) evolved. In particular, once it was recognized on the basis of argumentation by Rappaport 1983 that argument structure needed to be posited as a level of representation that was independent of constituent structure (e.g. very much unlike the assumptions of GB/MP, see Chapter 3), the way was paved for analyses of case to be stated in terms of generalizations over a(rgument)-structure. That is, in terms of generalizations that take both *semantic* and *syntactic* factors into account.

This chapter first presents some LFG basics in section 4.2, then briefly charts the development of Linking Theory in section 4.3 and presents current theories of case in section 4.4. The chapter closes with a consideration of LFG-based analyses that are couched within Optimality Theory (see Chapter 6 of this volume) in section 4.5.

(p. 60) 4.2 LFG basics

LFG separates facts about linear word order and constituency from the functional dependency structure analysis of a clause. Word order and constituency are represented at c(onstituent)-structure via tree representations, as shown in (1). LFG assumes a version of X'-theory that goes back to Bresnan (1977). For current assumptions about c-structural representations, see Bresnan 2001.

The f(unctional)-structure is represented in terms of an attribute-value matrix (AVM) and encodes functional syntactic information about grammatical relations, tense/aspect, case, number, person, etc. A sample f-structure for (1) is shown in (2).



(1) Peter drinks coffee.

Case in Lexical-Functional Grammar

In addition to the basic c- and f-structural representations, LFG's projection architecture potentially allows for several other projections. One standard additional projection is the s(ematic)-projection (e.g. Halvorsen and Kaplan 1988; Dalrymple 1999), which encodes the semantic analysis of the clause. The a-structure can also be conceived of as a projection, it can also be seen as an elaboration of the PRED value in f-structures such as (2) or (5) (Alsina 1996). A-structure in LFG in theoretical papers is generally represented as in (7), though it can also be formally represented as an AVM (Butt 1998), in line with f-structure representations.

(7) *pinch* (agent theme)

The a-structure encodes predicate-argument relationships in terms of thematic roles. These thematic roles are generally arranged in a thematic hierarchy, shown in (8) (based on Bresnan and Kanerva 1989). However, in practice very little reference is ever made to the hierarchy, except to the agent as being highest.

(8) Thematic Role Hierarchy

agent) *beneficiary*) *recipient/experiencer*) *instrumental*) *theme/patient*) *location*

4.3 Linking Theory

The first paper to associate case marking patterns with a-structure information was also the first paper to formulate some initial linking principles. Zaenen, Maling, (p. 63) and Thráinsson (1985) looked at case marking patterns in Icelandic. Besides devising tests that conclusively established the existence of non-nominative subjects, Zaenen, Maling, and Thráinsson (ZMT) formulated principles which governed a complex relationship between thematic roles, case, and grammatical functions. As an example, their association principles for Icelandic are shown in (9).

(9) Icelandic Association Principles

1. AGENTS are linked to SUBJ (Universal)
2. Case-marked THEMES are assigned to the lowest available GF. (Language-specific)
3. If there is only one thematic role, it is assigned to SUBJ; if there are two, they are assigned to subj and obj; if there are three, they are assigned to SUBJ, OBJ, 2OBJ [NB: the modern OBJ_θ]. This principle applies after principle 2 and after the assignment of restricted GFS. (Universal)
4. Default Case Marking: the highest available GF is assigned NOM case, the next highest ACC. (Universal)

The effect of the association principles is best illustrated via a concrete example. The Icelandic verb *óska* 'to wish' can be used either transitively or as a ditransitive so that the

Case in Lexical-Functional Grammar

goal 'her' in (10) is optional. When it is present, it is realized as the direct object (OBJ). When it is not present, the theme argument is instead linked to the direct object, as shown in (11).

(10) *pú hefuróskae (henni) pess*

You have wished her.DAT this.GEN

'You have wished this on/for her.' (ZMT 1986:470) Icelandic

(11) *óska:* (agent theme (goal))

[+gen] [+dat]

a. SUBJ 2OBJ OBJ

b. SUBJ oBJ

ZMT define inherent lexical case as being an idiosyncratic property of a lexical item, assigned by a verb, preposition, or adjective. In (11), the theme 'this' is assigned inherent genitive case by the verb 'wish'. If the theme were not marked with the genitive feature, then it would be linked to an OBJ in both the transitive and the ditransitive scenario by the association principles in (9). But given the special marking in conjunction with the language-specific principle in (9.2), it is assigned to a secondary object rather than the direct object in (11a). In (11b), the theme is linked to the direct object because that is the lowest available GF, given that there are no further arguments to be accommodated.

ZMT's notion of inherent case came to be known as quirky case. The term 'quirky' suggests a random lawlessness, but a close inspection of ZMT's original paper shows that inherent case assignment actually proceeds in a very regulated manner. The 'quirky' genitive or dative cases are always regularly associated with a (p. 64) given thematic role. Genitives regularly occur on themes (11) and datives mark goals as well as themes. There seem to be no instances of truly idiosyncratic case, rather case assignment seems to be principled and follows from lexical semantic factors.

Today's standard Linking Theory relates GFs to thematic roles via two abstract linking features, [\pm o](bjective) and [\pm r](estriuctive), by which both thematic roles and GFs can be classified. Additionally, a number of principles govern the association of GFs and thematic roles. Furthermore, argument changing operations such as passives (argument deletion) or applicatives (argument addition) are taken into account (see Butt 2006 for a detailed discussion). The role of case in most accounts has stayed much as it was in ZMT's analysis of Icelandic: an extra piece of information that helps determine the mapping between GFs and thematic roles. However, there are also some approaches that take the theory of case within LFG a bit further. Two of these are sketched briefly in the next section.

4.4 Some theories of case

4.4.1 Constructive Case

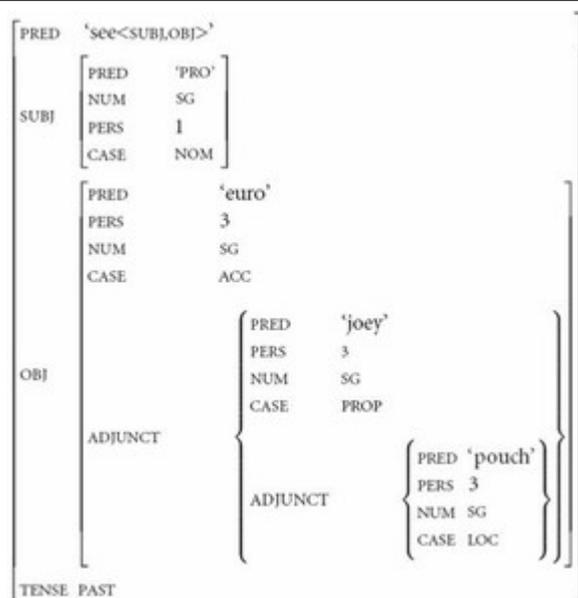
Nordlinger (1998b, 2000) takes on the phenomenon of *case stacking* in Australian languages and develops a theory of *Constructive Case*. The phenomenon is exemplified by (12), where the word ‘pouch’ is marked with three cases: one to show that it is signalling a location, one to show that it is part of a possessive or accompanying relation to another word (the proprietive case), and one to show that it is part of (modifying) an accusative case-marked noun (see Chapter 52 for more discussion of this type of case marking). The word ‘joeys’ (a baby euro – a type of kangaroo) has two cases. The proprietive shows that it stands in an accompanying relationship with another (with the euro), and the accusative shows that it is part of (modifying) an accusative case-marked noun. Finally, the ‘euro’ is accusative as the direct object, while the ‘I’ is nominative (unmarked).

(12) *Ngayu nhawu-lha ngurnu tharnta-a mirtily-marta-a*
I saw-PAST that.ACC euro-ACC joeys-PROP-ACC
thara-ngka-marta-a.
pouch-LOC-PROP-ACC
‘I saw the euro with a joeys in (its) pouch.’ Martuthunira
(Dench 1995: 60)

The f-structure analysis in (13) shows that the case markers indicate which functional layers of analysis the case-marked word is embedded in. That is, the three case markers on ‘pouch’ signal that it is a locative adjunct embedded under a proprietive adjunct that in turn modifies an accusative direct object.

(p. 65) Nordlinger (1998b) sees the case morphology itself as playing a large role in *constructing* the syntax of the clause. An ergative case in her analysis, for example, carries the following pieces of syntactic information: i) that there is a subject; ii) that it is ergative (ergatives are always subjects, but subjects are not always ergative). These pieces of information are encoded as part of the lexical entry of the ergative, as shown in (14a). Similarly, the abstract entries for the accusative, proprietive, and locative in Martuthunira specify the grammatical function they indicate, as well as the case feature, as shown in (14b-d).

Note the special use of the ↑ in the lexical entries of the case markers. The specification of the case feature is standard: each case marker specifies that the attribute case is assigned a certain value (ergative, accusative, etc.). This ensures that whatever constituent carries the case marker will be analysed as ergative, or accusative, or locative, etc. The second line in each entry involves *inside-out functional designation* (Dalrymple 1993, 2001). The ↑ following the specification of a grammatical function formulates a requirement that, come what may, the constituent should be analysed as a subject in (14a), an object in (14b), and an adjunct in (14c-d).²



[Click to view larger](#)

(13)

(p. 66) (14)

- a. ERGATIVE:
(↑CASE) = ERG
(SUBJ ↑)
- b. ACCUSATIVE:
(↑CASE) = ACC
(OBJ ↑)
- c. ERGATIVE:
(↑CASE) = LOC
(ADJUNCT ↑)
- d. PROPRIETIVE:
(↑CASE) = PROP
(ADJUNCT ↑)

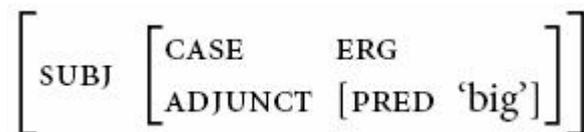
Now consider the example in (15) from

Wambaya. Here the NP 'big dog' is a discontinuous constituent. This poses no problem for Nordlinger's view of case as the effect of the analysis is that the combination of information from the lexical entries of 'big', 'dog', and the ergative case in (14a) results in the two partial f-structures shown in (16) and (17). Both the ergative 'dog' and the 'big' specify that they are parts of the subject. The 'dog' serves as the head of the phrase and the 'big' as an adjunct which modifies it.

(15) *galalarrinyi-ni gini-ng-a dawu bugayini-ni*
 dog.i-ERG 3SG.MASC.A-1.O-NFUT bite big.i-ERG
 'The big dog bit me.' Wambaya



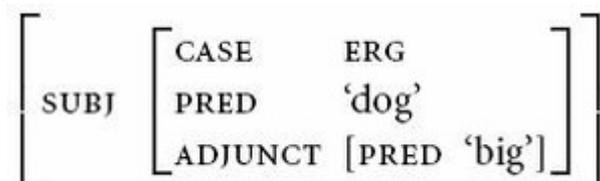
(16)



(17)

These two sets of information are unified into the structure shown in (18) as a routine part of the clausal analysis within the LFG formalism. The problem

of discontinuous constituents is solved by using the case morphology as a primary source of information about clausal structure.



(18)

However, note that Nordlinger's system makes no mention of a-structure and, in

contrast to ZMT on Icelandic, draws no connection between a-structural information and case marking. The next section introduces a model of case within LFG that assumes a complex interaction between a-structure, GFs, and case marking.

4.4.2 Differential Case Theory

Urdu exhibits alternations in which the only difference between two clauses is the case morphology on one of the noun phrases. This morphological difference (p. 67) signals a difference in semantic interpretation. Some examples are shown in (19) and (20). In (19a), the ergative subject indicates that the participant ('Nadya') has some control over the action, whereas the dative case in (19b) implies more of an obligation to perform the action (see Bashir 1999 for a more differentiated analysis). This is an example of case marking involving a *subject alternation*, or *differential subject marking* as it has come to be called in the Optimality Theoretic literature (see Chapter 6).

(19)

a. *nadya=ne zu ja-na hε*

Nadya.F.SG=ERG ZOO.M.SG.OBL GO-INF.M.SG BE.PRES.3.SG

'Nadya wants to go to the zoo.' Urdu

b. *nadya=ko zu ja-na hε*

Nadya.F.SG=DAT ZOO.M.SG.OBL go-INF.M.SG be.PRES.3.SG

'Nadya has to go to the zoo.' Urdu

Example (20) involves an *object alternation*, or *differential object marking* in Optimality Theoretic terms. Here the only difference between the two sentences is the case marker on the object in (20b). In addition to marking dative case, as in (19b), the *ko* also functions as a marker of specificity/definiteness on direct objects in Urdu. Because it appears only on direct objects in this semantic context, and because it is not retained under passivization, this case marker can be analysed as an accusative. The effect of *ko* in (20b) is that Nadya must be interpreted as having a particular giraffe in mind that she wants to go to see. In (20a), in contrast, it could be some generic giraffe or giraffes that Nadya would like to see (at the zoo, for example). The 'giraffe' in (20a) is glossed as nominative. This case has no overt morphophonological realization in Urdu, something which is cross-linguistically quite common.

(20)

a. *nadya=ne jiraf dek^h-na hε*

Nadya.F.SG=ERG giraffe.M.SG.NOM see-INF.M.SG be.PRES.3.SG

'Nadya wants to see a giraffe/giraffes.' Urdu

b. *nadya=ne jiraf=ko jiraf=ko dek^h-na hε*

Nadya.F.SG=ERG giraffe.M.SG=ACC see-INF.M.SG be.PRES.3.SG

'Nadya wants to see the giraffe.' Urdu

Parallels to this alternation between nominative and accusative can be found in Turkish (Enç 1991) and other South Asian languages. In this example, as in (19), the only difference between the a and b versions is the case marking. This, and the clear connection to a semantic difference, prompted Butt and King (1991) to begin formulating an approach to case that included a notion of *semantic case*. This term has generally been used to refer to the case marking of adjuncts such as locatives or temporal expressions. Butt and King used this term to apply to those case markers of core arguments which also contribute information that is relevant for the final semantic interpretation of the clause. In order to allow the semantic information contributed by the case markers to flow directly into the analysis of the clause, (p. 68) Butt and King (1991), exactly like Nordlinger (1998b), proposed explicit lexical entries for case markers. An example taken from later work (Butt and King 2003), is shown in (21) for the use of accusative *ko* in (20b).

(21) *ko* (↑ CASE) = ACC
(OBJ ↑)
(↑_{sem-str} SPECIFICITY) = +

Butt and King's (1991) proposals for semantic case foreshadowed Nordlinger's ideas about *constructive case* in that the case markers themselves are considered to be active components which contribute to the analysis of a clause. The lexical entry for the accusative use of *ko* states that: i) the case is accusative; ii) the relevant NP should be a direct object; iii) the NP should be interpreted as specific at the level of s(ematic)-structure.

But the information carried by case markers is only part of a more complex system that accounts for differential case marking patterns. Butt and King 2003, 2005 assume the version of linking theory proposed in Butt 1998. In this version of linking theory, no explicit thematic hierarchy is assumed and the thematic roles are restricted to a very basic set: *agent, goal, theme, locative*. Beneficiaries, recipients, experiencers, etc. are all assumed to be an instance of a goal, more or less abstract.

As in standard linking theory, case as such is not integrated directly into the linking between GFs and a-structure. However, case marking becomes relevant for linking when different linking possibilities exist, just as was the case for ZMT. The information provided by the case morphology can be used to select just one of the possibilities.

But, as the examples in (19) and (20) show, the function of case marking must go beyond just the determination of grammatical functions. In fact, under Butt and King's Differential Case Theory (DCT), most instances of case work out to be instances of semantic case, which is defined as being simultaneously: i) semantically predictable; ii) subject to syntactic restrictions, such as being confined to certain grammatical functions.

Case in Lexical-Functional Grammar

Indeed, the bulk of the cross-linguistic case marking phenomena involve an interaction between syntactic and semantic constraints.

Take the ergative case in Urdu, for example. This case marker can only appear on subjects and so must obey a particular syntactic restriction. Because it is an agentive case, the ergative does not just play a structural role, it also gives rise to some semantic effects. As is well known from a host of studies on Hindi/Urdu, the ergative alternates with the unmarked nominative on unergative intransitives (the most comprehensive study to date of the distribution of the ergative is Davison 1999). As shown in (22), this alternation correlates with an expression of control/ volitionality.

(22)

a. *yassin k^hãs-a*

Yassin.M.SG.NOM cough-PRF.M.SG

'Yassin coughed.' Urdu

(p. 69) **b.** *yassin=ne k^hãs-a*

Yassin.M.SG=ERG cough-PRF.M.SG

'Yassin coughed (purposefully).' Urdu

Another example is the accusative/dative *ko* in Urdu. As was seen in connection with example (20), when it appears on direct objects, it signals specificity. Again, there is a combination of syntactic (direct objects only) and semantic factors (specificity). The *ko* can also appear on subjects (19b) and on indirect objects, as in (23).

(23) *nadya=ne billi=ko dud di-ya he*

Nadya.F=ERG CAT.F.SG=DAT milk.M.NOM give-PRF.M.SG be.PRES.3. SG

'Nadya has given milk to the cat.' Urdu

In both cases, the dative is associated with a more or less abstract goal. In (23) the 'cat' is the goal of the giving. In *experiencer* constructions as in (19b) or (24a-c), the experiencer of the event can be thought of as a kind of abstract goal (cf. Landau 2002).

(24)

a. *nadya=ko skul ja-na* *Case in Localist Case Grammar 121*

Nadya.F. SG=DAT school.F.SG.OBL go-INF.M.SG fall-PRF.M.SG

'Nadya had to go to school.' Urdu

b. *nadya=ko*

Nadya.F.SG=DAT fear.M.SG.NOM be attached-PRF.M.SG

'Nadya was afraid.' Urdu

c. *nadya=ko kahani yad a-yi*

Nadya.F.SG=DAT story.F.SG.NOM memory come-PRF.F.SG

'Nadya remembered the story.' Urdu

d. *nadya=ne kahani yad yad k-i*

Nadya.F.SG=ERG story.F.SG.NOM memory do-PRF.F.SG

'Nadya remembered the story (actively).' Urdu

With psych predicates as in (24c) there is again an alternation with the ergative, though in this case it correlates with the use of the agentive 'do' in (24d) vs. the non-agentive 'come' in (24c). The dative use of *ko* is also governed by a combination of syntactic and semantic factors. It is restricted to indirect objects and subjects, but is subject to a coherent *goal* semantics in Urdu.

In DCT, there is thus explicit reference to a-structure concepts such as *goal* (datives) or *agent* (ergative) in addition to further semantic concepts such as specificity as part of the analysis of case. DCT recognizes case as an extremely complicated and complex part of the morphosyntactic and semantic interface, particularly as there are also some instances of case marking which appear to be tied to purely positional/structural or idiosyncratic factors. In DCT, therefore, examples such as the English adnominal genitive (e.g. *John's hat*) are analysed as purely positional/ structural case.

DCT also assumes a notion of default case. Indeed, it can be observed that structural/ positional case is often an instance of default case (the Elsewhere Case, (p. 70) Kiparsky 1973). There are some languages which require that all NPs have case. For these languages default case satisfies something that has become known as the Case Filter (Rouveret and Vergnaud 1980). That is, if a given NP is not already associated with case due to some specification in some other part of the grammar, then default case assignment principles can apply. Typical cross-linguistic default cases are nominative or genitive. Default case only applies to the core grammatical relations subject and object. The other grammatical relations tend to involve some kind of specialized semantics and therefore do not involve defaults. Note that this is essentially the insight behind GB/ Minimalism's postulation of exactly two structurally Case-marked arguments, which are associated with nominative and accusative case.

Finally, DCT also acknowledges that some instances of case marking are due to truly coincidental historical developments: no generalization can be made and the case marking has to be stated as an exception to the system. These instances of case marking are labelled *quirky*.

An example of truly idiosyncratic marking is shown in (25). Urdu requires that subjects of agentive transitive perfect verbs be marked with the ergative case. However, the verb 'bring' in (25) is a transitive verb with perfect morphology and the bringer can be unproblematically classified as an agent. But the subject 'Nadya' in (25) is nominative and not ergative.

(25) *nadya kitab la-yi*

Nadya.F.SG.NOM book.F.SG.NOM bring-PRF.F.SG

'Nadya brought a book.' Urdu

The reasons for this aberrant case marking are not clear. Also, there are no other agentive transitive verbs which behave like this. Therefore this bizarre nominative case marking must be stipulated as part of the lexical entry of 'bring' (the nominative on objects is regular in Urdu).

On the whole, however, DCT assumes that such instances of quirky case are relatively rare. Rather, most case marking phenomena should fall under the rubrik of *semantic case*, that is, as being constrained simultaneously by both syntactic and semantic conditions. This stands in marked contrast to many theories of case, but does seem to reflect the rather unique cross-linguistic positioning of case at the interface between lexical and clausal semantics, syntax and morphophonology.

4.5 Incorporating Optimality Theory

The advent of Optimality Theory (OT, see Chapter 6) was observed with a great deal of interest in LFG and adopted into the theory very quickly (Bresnan 2000). (p. 71) The input to an evaluation by OT constraints is assumed to be f-structure and c-structure pairings. The task of the OT constraints is to pick the most optimal pairing.

Several OT-LFG analyses exist with respect to case and they tend to fall in line with OT analyses of case in general. That is, they adopt the insights with respect to differential case marking advanced by Aissen 1999, 2003). Some examples are Sharma's (2001) analysis of Kashmiri case clitics, Deo and Sharma's (2006) analysis of typological and diachronic variation in Indo-Aryan ergative patterns, and Asudeh's (2001) analysis of optional case patterns in Marathi. These analyses have all the advantages and disadvantages of the general OT analyses that are discussed in Chapter 6 of this volume. Asudeh's analysis, however, proposed an interesting extension to OT, namely, the availability of optionality in the evaluation of the most optimal candidate. In addition, Lee's (2001a, b; 2003) work on bidirectional OT and output-output correspondences provides a new perspective on Hindi and Korean case as well as word order freezing effects, which are due to identically case-marked arguments in a clause.

4.6 Summary

In sum, LFG contains a range of differing approaches to case marking. No account would claim to have developed a complete theory of case marking cross-linguistically. Rather, new ideas are continually being developed and tried out in light of new empirical evidence. This does not mean, however, that each account discards already established

Case in Lexical-Functional Grammar

insights. Rather, each new account builds on the strong points of previous insights within LFG, but is simultaneously not willing to let itself be blinded by the previously established ideas.

Notes:

(1) Within LFG, only minimal constituents are assumed for which there is actual distributional evidence in the language. The difference between a language like English and a language like Urdu is taken to be a difference between endocentricity (all heads project maximally) and exocentricity (heads do not necessarily maximally project). See Bresnan 2001: ch. 7) and Dalrymple (2001: ch. 3) for further discussion.

(2) See also Andrews (1996).

Miriam Butt

Miriam Butt is Professor of Theoretical and Computational Linguistics at the University of Konstanz. She works primarily on Urdu (syntax, morphology, semantics, some phonology) and is currently concentrating on the history and distribution of case. She is also working on building a computational grammar (parser/generator) and finite-state morphology for Urdu.

