A Paper to be presented at the International Symposium on 'Non-nominative Subjects' organized by ILCAA, Tokyo [18–21 December 2001]
{Do not quote without the permission of the authors}

NON-NOMINATIVE SUBJECTS IN URDU
A COMPUTATIONAL ANALYSIS

Miriam Butt and Tracy Holloway King
Universität Konstanz and Xerox PARC
Konstanz, Germany and Palo Alto, California, USA

1. Introduction

Within formal linguistic theory, case has traditionally been thought of in terms of a contrast between structural and inherent case, where inherent case is stipulated in the lexical entry of the verb. In computational terms, this division is easy to encode: structural case is the default case and is part of the grammar, while inherent case, which deviates from the nominative-subject/accusative-object pattern, is included in the lexical entries of individual verbs. In languages such as Georgian, where the choice of the case marker on subjects (nominative, ergative, or dative) depends on the particular tense/aspect morphology of the verb, the case disjunctions have to be encoded in terms of these morphological distinctions as well.

This division between structural and inherent case is true for standard approaches to case in South Asian languages as well. Mahajan (1990), for example, proposes that argument noun phrases in Hindi may have both structural and inherent Case. Structural Case is assigned in SpecAgrP, SpecIP, or in the complement to V position. The inherent Case of an argument is specified in the lexical entry of a particular verb form. Overt case morphology is always treated as an instance of inherent Case. However, noun phrases marked with inherent case may also be assigned structural Case if they are in the appropriate structural position. This dual system of Case assignment applies to ergative subjects and to direct objects marked with dative/accusative *ko* because these function as direct arguments which are also overtly marked with case clitics. Under this analysis, as with Georgian, case assignment (particularly structural Case) is dependent on the particular morphological form of the verb. For instance, verbs with perfect morphology are taken to lack the ability to assign structural Case. This in turn causes the object to move to a functional agreement position while the subject remains inside the VP and is marked inherently with the ergative case.

Davison (1999) provides the most complete account of the pattern of Hindi ergative marking to date. Her research is based on a careful survey of case marking across several verb classes. She treats the ergative as a structural Case which interacts with the specifications of the lexical entry of the verb. She proposes the licensing conditions in (1).

(1) a. *Verb condition:* the lexicon specifies which verbs have [ERG] external arguments.

b. *Aspect condition:* perfective Aspect licenses [ERG].

c. *Tense condition:* finite Tense licenses [ERG]

Davison's proposal is very close to the one presented here: information coming from the verbs' lexical entries interacts with information provided by the ergative. However, under her account, functional projections license the assignment of structural Case and the semantic factors involved in case-alternations

are associated with the lexical entry of a verb. We instead believe that while some inherent case marking information must be relegated to the lexicon, a more general story can be told with respect to the case alternations observed in Urdu/Hindi. In particular, our story involves a heavy reliance on information encoded by the case markers themselves, which interacts with semantic effects and a notion of grammatical functions (rather than functional projections).

In this paper, we present a number of case alternations in Urdu/Hindi and argue that they are governed by regular semantic alternations, rather than by idiosyncratic requirements of individual verbs. While these semantic alternations are also sensitive to structural and morphological conditions, they are not determined exclusively by them. We therefore argue for a three way division in the case system. In addition to structural and quirky case, we posit semantic case and allow case markers themselves to play an active role in the construction of the syntactic and semantic analysis of a clause. As we will see, this is necessary to account for case alternations in Urdu, including the appearance of non-nominative subjects.

After a theoretical consideration of the issues with respect to non-nominative subjects and case alternations, we turn to some implementational issues. The theoretical approach we posit requires a complex interaction between semantic features, argument structure, grammatical functions, and phrase structure. Lexical-Functional Grammar (LFG) provides the possibility of such a complex interaction via its system of mutually constraining levels of representation or *projections*. Due to its mathematically constrained nature, various computational implementations of LFG exist (`http://clwww.essex.ac.uk/LFG/`). We base this paper on the Xerox Linguistic Environment (XLE), perhaps the most extensive and comfortable implementation available to date (see Butt et al. 1999 for an overview and further references).

As part of our computational analysis, we discuss the representation of argument structure and the treatment of case markers in terms of "constructional case". We show that while not all of the tools assumed by the theoretical analysis are available within XLE, the basic architecture of LFG as implemented in XLE allows an insightful and precise modelling of the role of non-nominative subjects in South Asian languages.

2. Urdu Case

In this section, we first describe the basic case system of Urdu and then go on to describe the case alternations that we use to motivate our notion of semantic case.

2.1. The Basic System

Urdu has six cases, shown in table (2). We will be primarily concerned with the ergative, and dative/accusative cases and will take on a discussion of the instrumental in section 5.3. Note that the case markers are clitics, not inflections on the noun or postpositions (T. Mohanan 1994, Butt and King 1999).[1]

(2)

Urdu Cases	
Case	Clitic Form
nominative	∅
ergative	ne
dative/accusative	ko
instrumental	se
genitive	k-
locative	mẽ, par, tak, ∅

All the case markers may appear on subjects in Urdu (T. Mohanan 1994). However, there are a large number of verbs that show a split active pattern in the following sense. The ergative appears on on the subjects of transitive verbs when the verb carries perfect morphology (*-a/-i/-e*). In other tenses, these subjects are nominative. The subjects of unergatives are nominative in non-perfect tenses ((4b)); in the perfect, they are ergative if they express volitionality ((5b)) and nominative otherwise ((5c)). The subjects of perfect unaccusatives, in contrast, are always nominative ((4c) and (5d)).

Note that the case marking of the object does not depend on the case of the subject (as is often assumed in the accusative vs. ergative typologies posited by standard approaches to case). The case marking on the object (accusative vs. nominative) in Urdu/Hindi is orthogonal to subject case marking because the appearance of the accusative is tied to specificity (Butt 1993) in a manner very similar to that in Turkish (Enç 1991), even though Turkish does not have an ergative case.

(3)

Urdu Case Alternations		
	Non-perfect (Accusative Pattern)	Perfect (Active Pattern)
Transitives	NOM-ACC/NOM	ERG-ACC/NOM
Unacc. Intrans.	NOM	NOM
Unerg. Intrans.	NOM	ERG/NOM

Non-perfect examples of transitives, unergatives, and unaccusatives are shown in (4). In each sentence, the subject is nominative and the verb agrees with it.

(4) a. nadya gari cɑla-ti hɛ
Nadya.F.Nom car.F.Sg.Nom drive-Impf.F.Sg be.Pres.3.Sg
'Nadya drives a car.' (Transitive; Non-perfect)

b. nadya naha-e-g-i
Nadya.F.Nom bathe-3.Sg-Fut-F.Sg
'Nadya will bathe.' (Unergative; Non-perfect)

c. nadya ja-e-g-i
Nadya.F.Nom go-3.Sg-Fut-F.Sg
'Nadya will go.' (Unaccusative; Non-perfect)

When these appear in the perfect, the subject of a transitive verb must be ergative ((5a)), while the subject of an unergative may optionally take an ergative ((5b–c)). This optionality is correlated with a more volitional reading for the ergative. Unaccusative subjects remain in the nominative, as in (5d). Urdu verbs do not agree with ergative subjects; instead, they agree with their highest nominative argument (nominatives can only be subject or object, whereby subject is higher than object). The verb defaults to masculine singular third person agreement if there is no available nominative subject or object in its predicational domain, as in (5b) (i.e., the local clause).

(5) a. nadya=ne gari cɑla-yi hɛ
Nadya.F=Erg car.F.Sg.Nom drive-Perf.M.Sg be.Pres.3.Sg
'Nadya has driven a car.' (Transitive; Perfect)

b. nadya=ne naha-ya
Nadya.F=Erg bathe-Perf.M.Sg
'Nadya bathed (on purpose).' (Unergative; Perfect)

c. nadya naha-yi
Nadya.F.Nom bathe-Perf.F.Sg
'Nadya bathed.' (Unergative; Perfect)

d. nadya/*=ne gɛ-yi
Nadya.F.Nom go-Perf.F.Sg
'Nadya went.' (Unaccusative; Perfect)

These relatively standard alternations related to the Urdu split active system (Butt and King 2001) are augmented by semantically motivated case alternations. Three of these are presented in the next section.

2.2. Case Alternations

Semantic correlates with case marking are the rule in Urdu and result in case alternations on a number of arguments. One of the best known alternations is on objects of perfective transitive verbs which can appear in the nominative or the accusative, as in (6). When the object is accusative, it must be specific (T. Mohanan 1994; Butt 1993; de Hoop 1992).

(6) a. ram=ne jiraf dekʰ-i
Ram=Erg giraffe.F.Nom see-Perf.F.Sg
'Ram saw a/some giraffe.'

b. ram=ne jiraf=ko dekʰ-a
Ram=Erg giraffe.F=Acc see-Perf.M.Sg
'Ram saw the (particular) giraffe.'

In this paper we concentrate on case alternations on subjects. These alternations show a combination of semantic and structural effects. First consider the ergative, whose use in many constructions is correlated with volitionality (Tuite, Agha, and Graczyk 1985; Butt and King 2001). This is seen with certain perfective unergatives whose subject can be either nominative (7a) or ergative (7b).

(7) a. ram kʰãs-a
Ram.M.Nom cough-Perf.M.Sg
'Ram coughed.'

b. ram=ne kʰãs-a
Ram.M=Erg cough-Perf.M.Sg
'Ram coughed (purposefully).'

Another interesting alternation that correlates with volitionality or control over an action concerns noun-verb complex predicates and modal readings with infinitives. The sentences in (8) illustrate an alternation with noun-verb complex predicates (T. Mohanan 1994). Here the case alternation interacts with a difference in the choice of light verb: agentive 'do' vs. unaccusative 'come'. The dative *ko* in (8b) marks a goal or experiencer in the manner of psych predicates, while the ergative *ne* marks agentivity or volitionality in (8a), thus confirming the semantic correlation between the ergative case and volitionality.

(8) a. nadya=ne kɑhani yad k-i
Nadya.F.Sg=Erg story.F.Sg.Nom memory do-Perf.F.Sg
'Nadya remembered the story (actively).'

b. nadya=ko kɑhani yad a-yi
Nadya.F.Sg=Dat story.F.Sg.Nom memory come-Perf.F.Sg
'Nadya remembered the story (the story came to Nadya).'

Furthermore, in a departure from the split-ergative pattern in which ergative case is tied to the presence of perfect morphology, Urdu allows the ergative to appear with an infinitive in combination with a present or past form of *ho* 'be'. This construction shows a systematic alternation between ergative and dative subjects, which coincides with a difference in modality, as shown in (9).

(9) a. nadya=ne zu ja-na hɛ
Nadya.F=Erg zoo.M.Loc go-Inf be.Pres.3.Sg
'Nadya wants to go to the zoo.'

b. nadya=ko zu ja-na hɛ
Nadya.F=Dat zoo.M.Loc go-Inf be.Pres.3.Sg
'Nadya wants/has to go to the zoo.'

In this infinitive construction, the ergative is the marked form and entails a subject who has control over the action. The dative is the unmarked form or elsewhere case: the dative subject may or may not have control over the action, the precise interpretation depends on the context (Bashir 1999). For a detailed LFG analysis of this construction see Butt and King (1999).

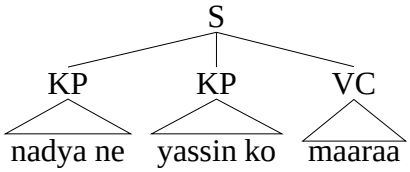
The generalization seems to be that semantic factors are behind these case alternations. Therefore, an analysis which requires brute-force listing of case assignments via lexical stipulation is unfeasible. Instead, the case markers interact with the lexical semantics of the verb via the argument-structure and some semantic features. Furthermore, work on Australian languages (Nordlinger 1998) has shown that precisely such an approach is needed for the case marking patterns in these languages, which also display productive use of non-nominative subjects.

3. Case in LFG

In LFG, information from different components combines to constrain one another and produce a consistent and coherent analysis of a given clause (see Bresnan 2001 for a recent overview of LFG). The differing modules of grammar (e.g., grammatical functions, semantics, and phonological information) are encoded in terms of projections from lexical entries and phrase structure rules, which in turn encode syntactic and morphological constituency. This is illustrated informally in (11) for the Urdu sentence in (10).

(10) nadya=ne yassin=ko mar-a
Nadya.F=Erg Yassin.M=Acc hit-Perf.M.Sg
'Nadya hit Yassin.'

(11) a. Constituent-structure:

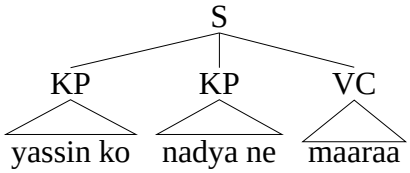


b. Functional-structure:

$$\begin{bmatrix} \text{PRED} & \text{'hit<SUBJ, OBJ>'} \\ \text{TENSE} & \text{PRES} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'Nadya'} \\ \text{CASE} & \text{ERG} \\ \text{GEND} & \text{FEM} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'Yassin'} \\ \text{CASE} & \text{ACC} \\ \text{GEND} & \text{MASC} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix}$$

A sentence like (10) has two syntactic structures associated with it. The first is a phrase structure tree, referred to as the c(onstituent)-structure.[2] LFG avoids the use of traces. The c-structure therefore closely reflects the actual string and contains a faithful representation of linear order and constituency information. The grammatical functions are encoded in the f(unctional)-structure as an attribute value matrix (AVM). Note that different word orders of (10), which are possible since Urdu is a "free" word order language, will have different c-structures but identical f-structures; the correlated differences in discourse-functions can be encoded in a separate projection (King 1997) or in the f-structure. Thus the c-structure in (12) also corresponds to the f-structure in (11b).

(12)



3.1. Theoretical Approach

Case phenomena have been extensively analyzed in LFG. Of particular interest here is the idea of Constructive Case, proposed by Nordlinger (1998). The basic idea behind Constructive Case is that nominal constituents with case morphology can define the larger syntactic context in which they appear.[3] Consider the Wambaya example in (13).

(13) 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, Nordlinger (1998:96))*

In (13), *galalarrinyi-ni bugayini-ni* 'big dog' is a discontinuous constituent, but both parts of the constituent are marked with ergative case. Under Nordlinger's analysis, the ergative case itself specifies that it is an ergative and that it must be part of a subject for the clause to be grammatical. This is outlined in the lexical entry in (14); the first line indicates that the noun phrase has ergative case, while the second states that it must be a subject.[4]

(14) *ni* $(\uparrow \text{CASE}) = \text{ERG}$
$(\text{SUBJ} \uparrow)$

Together with the predicate value for ‘dog’ supplied by the noun, this entry for the case marker results in the (simplified) syntactic f-structure in (15) for *galalarrinyi-ni* ‘dog-Erg’.

(15) $$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{‘dog’} \\ \text{CASE} & \text{ERG} \end{bmatrix} \end{bmatrix}$$

This structure can then be unified with the representation projected by the adjective (adjunct) *bugayini-ni* ‘big-Erg’ (16) to give a coherent analysis of the subject of the clause, as in (17).

(16) $$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{CASE} & \text{ERG} \\ \text{ADJUNCT} & \begin{bmatrix} \text{PRED} & \text{‘big’} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

(17) $$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{CASE} & \text{ERG} \\ \text{PRED} & \text{‘dog’} \\ \text{ADJUNCT} & \begin{bmatrix} \text{PRED} & \text{‘big’} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

As can be seen from these examples, Constructive Case allows the case markers to play an active role in the clause. Not only do they assign case, but they can also specify information about the syntactic environment in which they occur, e.g., attaching to subject nominals in the example of the Wambaya ergative.

We argue for three types of case: QUIRKY CASE, SEMANTIC CASE, and STRUCTURAL CASE. Here we are concerned with semantic case and structural case.[5] Semantic case markers are characterized by selectional restrictions on semantic inferences over parameters such as volitionality; these are sometimes expressed as a restriction on thematic-roles. In addition, semantic case is also often sensitive to grammatical function, e.g., the particular case can only appear on a subject, or only on an oblique, etc. Finally, semantic case does not just affect the noun phrase to which it attaches; instead, it can affect the clausal semantics, e.g., through aspectual affectedness, specificity or partitivity.

An example of semantic case is the Urdu dative: the dative in Urdu and South Asian languages in general is associated with a goal argument (Verma and Mohanan 1990, cf. also Woolford 1997). The dative case is therefore associated with a constraint which captures this fact; an LFG formalization is shown in (18).

(18) ko $(\uparrow \text{CASE})=\text{DAT}$
$(\text{GOAL} \uparrow_{a-str})$

The first line of (18) associates the case clitic *ko* with the dative case. The second line stated that it must correspond to a goal in argument structure.[6] Nothing is said about the grammatical function of the dative: the fact that datives can be either subjects or obliques follows from constraints in other parts of the grammar.

Structural case in LFG is not necessarily associated with a particular phrase structure position; in fact, this is the least common way case is assigned. Instead, the relevant notion is the grammatical function of the case marked noun phrase. In our model, the case does not itself assign the grammatical function but instead helps to characterize the grammatical functions via wellformedness conditions. That is, the information contributed by the case marker provides further constraints on the grammatical functions. This is an important point as case markers do not define grammatical functions. Conversely, grammatical functions do not exclusively determine the case marking. For example, the ergative in Urdu is always associated with subjects, but does not define a subject. An LFG formalization is shown in (19).

(19) ne (↑ CASE)=ERG
(SUBJ ↑)

The first line in (19) states that the case of the noun with the clitic *ne* is ergative. The second requires it to be associated with a subject. More details on the ergative case assignment are discussed below.

3.2. The Relevance of a Computational Model

The theoretical approach to case sketched above differs markedly from standard analyses, which tie structural Case to functional categories and stipulate inherent/quirky Case in the lexical entry of a verb. To proponents of this standard type of approach, the view of case sketched above may seem to be worryingly unconstrained: how does one know that the various levels of representations indeed interact with one another in such a way as to give the right results? That is, can really only the wellformed sentences be generated and parsed? Can all the illformed sentences be ruled out so that they are neither parsed nor generated?

Given that LFG is a computationally viable theory of syntax, the questions above can be addressed via a methodology of proof via demonstration. In this paper we therefore report on a small grammar for Urdu which includes a treatment of non-nominative subjects. In particular, it will be seen that case entries and verb entries mutually constrain one another via linking rules that refer to the syntax and the argument structure.

A notion of argument structure has become well-established within theoretical linguistics (e.g., Grimshaw 1990); this level of representation has proven to be useful for capturing generalizations in a number of areas, e.g., nominalization, passivization, applicatives, causativization, locative inversion, and the formation of complex predicates. The role of argument structure as part of syntactic analyses has been explored in some detail within LFG and has led to the formulation of a Linking or Lexical-Mapping Theory (LMT) (see Bresnan and Zaenen 1990 for an overview).

However, argument structure has not found as secure a place within computational implementations of LFG despite extensive work on computational LFG. The main issue is one of grammar architecture. In "classical" LFG (see Bresnan 1982), the PRED is the cornerstone from which the f-structure is built. With the advent of LMT, the PRED was no longer seen as a primitive notion, but as a level of representation that could correspond either to a simple a(rgument)-structure as in (20), or a complex a-structure as in (21), which represents a language with a morphological causative.

(20) a-str *fall* < pt >
|
f-str PRED 'fall < SUBJ >'

(21) a-str CAUSE < ag *fall* < pt >>

f-str PRED

'fall-CAUS < SUBJ OBJ >'

For LFG implementations, this view of the relationship between argument structure and grammatical functions entails that PREDs should be able to be composed of separate pieces of information. However, at this point in the evolution of LFG implementations, we do not have a sufficient understanding of how to compose PREDs. Note that this issue is not merely an LFG-internal problem: Most computational implementations take the lemma of a verb and assign it a particular valency in the lexicon which can then only be changed via explicit rules, but not via predicate composition. Thus, while the argument frame of a lemma may subsequently be manipulated in the process of the derivation, it is often difficult to do so by means of generally applicable processes. Rather, argument alternations tend to be determined on a case-by-case basis, which obscures their productive nature and leads to a loss of empirical coverage and an increase in grammar and lexicon maintenance.

In this paper we report on an implemention of a-structure in an LFG grammar by relying only on minimal manipulation of the PRED, i.e. without an extension of the current implementation of the Xerox Linguistic Environment (XLE) grammar development platform. We first show where argument structure is situated in the basic architecture, then discuss the computational treatment of non-nominative subjects and case alternations in Urdu via a reliance on an interaction of information from various levels of grammar, including argument structure.

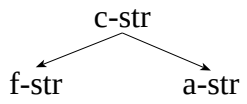
4. The Basic Implementation

The XLE system (see Butt, King, Niño and Segond 1999 for a description and further references) is a platform for large-scale LFG grammar development. XLE comprises interfaces to finite-state preprocessing modules for tokenization and morphological analysis (Karttunen et al. 1992; Kaplan and Kay 199), as well as an efficient parser and generator for LFG grammars (Maxwell and Kaplan 1991, 1993; Kaplan and Wedekind 2000). As mentioned previously, the LFG formalism assigns natural language sentences two levels of linguistic representation — a constituent phrase structure (c-structure) and an AVM functional structure (f-structure). The c-structure encodes constituency (dominance) and surface order (precedence). The f-structure is an attribute-value representation which encodes syntactic information in terms of morphosyntactic features (NUM, GEND, TENSE, etc.), as well as functional relations between predicates and their arguments or adjuncts. The two levels of representation are related via the correspondence function ϕ, which maps partial c-structures to partial f-structures.

4.1. Projection Architecture

The XLE system allows for the integration of other projections from the c-structure, such as semantic structure or discourse structure (Kaplan 1995). Here, we are interested in the c-structure, f-structure, and a-structure. Within our theory, the a-structure is a projection off the c-structure.[7]

(22)



4.2. *Lexical Entries*

The lexical entries for verbs (verb stems in our case) call two sets of templates:[8] one providing grammatical function information to the f-structure and one providing information about the argument structure roles. There is no direct link between grammatical functions and a-structure roles at this level, though a link may be established via the composition of the inverse of the projection (Halvorsen and Kaplan 1988) from c-structure to f-structure (ϕ) and the projection from c-structure to a-structure (α). Thus, in (23), the subject of the transitive verb *k*ʰ*ol* 'open' is not directly associated with the agent in a-structure.

(23) kʰol V @(TRANS open)
@ARG-AG @ARG-TH
@ARG-REL.

The template expansions for the relevant parts of this lexical entry are as follows. In (24), the expansion for TRANS provides a value for the PRED and specifies that there is a SUBJ and an OBJ; in addition, this structure can undergo passivization (section 5.3.1).

(24) TRANS(_P) = @(PASS (↑ PRED)='_P<(↑ SUBJ)(↑ OBJ)>').

The expansion for the agent argument is seen in (25) and is based on assumptions made by LFG's linking theory (Bresnan and Zaenen 1990), whereby both grammatical functions and argument roles are classified via the features [±r(estrictive)] and [±o(bjective)]. The AG attribute ROLE-O is assigned a − value, indicating that the AG can never have an objective function, i.e. it must be a subject or an oblique. There are then two options: (1) either it is further specified as ROLE-R − and requires no passivization and a SUBJ at f-structure (the agent is equated with the subject) or (2) the AG is marked as demoted and the f-structure is constrained to be a passive.

(25) ARG-AG =
$(\uparrow_{arg-str}$ AG ROLE-O) = −
{ $(\uparrow_{arg-str}$ AG ARG-FORM) = (↑SUBJ P-FORM) `option 1`
$(\uparrow_{arg-str}$ AG ROLE-R) = −
(↑SUBJ)
(↑PASSIVE) = −
|$(\uparrow_{arg-str}$ AG ARG-FORM)=NULL `option 2`
$(\uparrow_{arg-str}$ AG DEMOTED) =c +
(↑PASSIVE) =c +}.

Note that while there is no direct connection between the subject and the agent in the lexical entry for 'open' given in (23), the interplay of constraints within the template in (25) ensures that a connection between these two is established.

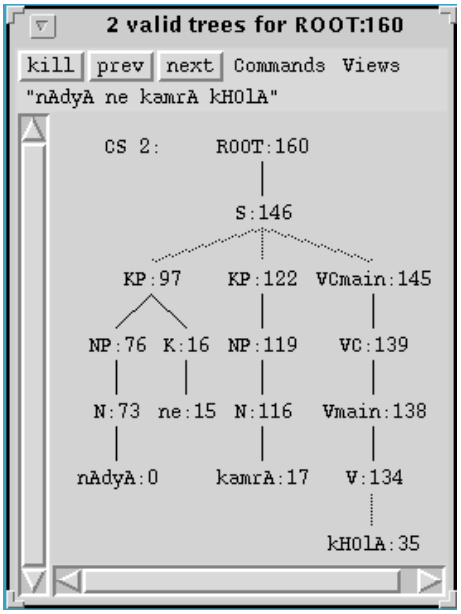
4.3. *Output*

The specfications illustrated in the previous section result in a triple of structures. These are shown in (27)–(29) for the transitive sentence in (26).

(26) nadya=ne kɑmra kʰol-a
Nadya.F=Erg room.M.Nom open-Perf.M.Sg
'Nadya opened the door.'

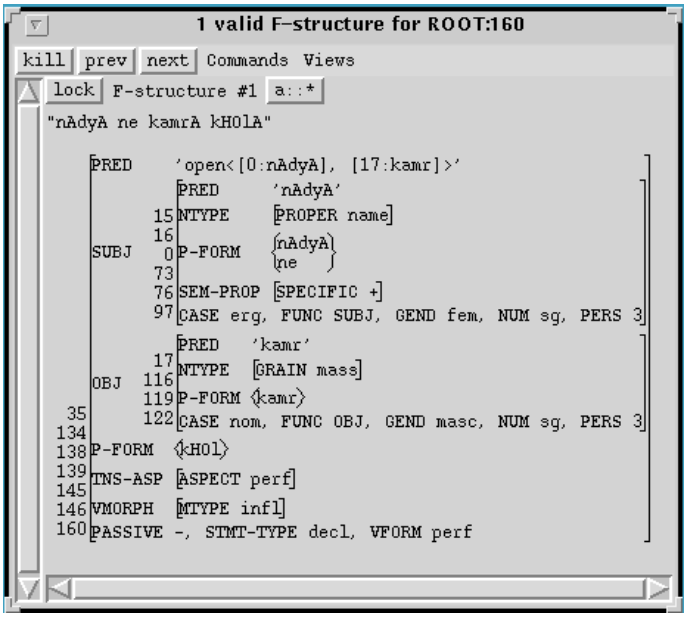
The c-structure in (27) contains an S which is comprised of three constituents, two KPs (case phrases) and a V(erbal)C(luster).

(27)

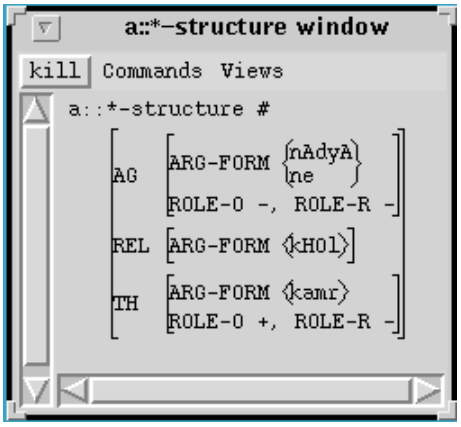


Two structures are projected from this c-structure: the f-structure in (28) and the a-structure in (29). The f-structure in (28) contains the expected SUBJ and OBJ arguments of the main PRED 'open'.

(28)



The a-structure in (29) contains an AG(ent), TH(eme), and REL(ation) which encodes the verb itself.

(29) 

Given this basic configuration of the system, we can now explore how semantic and syntactic constraints on case marking can be captured in a linguistically sophisticated and computationally viable way.

5. Theoretically Motivated Computational Analysis

This section goes through dative and ergative subject alternations, as well as instrumental subjects in ability passives, and shows how the approach to case sketched in section 3 is implemented successfully with respect to Urdu non-nominative subjects.

5.1. Dative Subjects

Recall that the dative in Urdu is roughly associated with the notion of 'goal' at argument structure. This is not unusual, as argument structure plays a role in constraining case assignment in many languages (see Zaenen, Maling and Thráinsson 1985 for Icelandic). Also recall that we propose to posit explicit entries for case markers. These entries interact with verbal entries such as in (23) for the verb 'open'.

The complete entry for the Urdu dative/accusative case marker *ko* is shown in (30). When this case marker is used as a dative (`option 1`), it marks an argument which is a goal (GO) in a-structure; this is accomplished by the existential equation ($\uparrow_{a-str}$ GO). Note that the dative can be associated either with subjects (FUNC SUBJ, `option 1a`)) or indirect objects (FUNC OBJth, `option 1b`) in the f-structure—either way it will still correspond to a goal argument. As an accusative (`option 2`), the *ko* denotes specificity or definiteness (Mohanan 1994) and is restricted to objects (FUNC OBJ).

(30) ko K * @(PWORD ko)

```
{ (↑CASE) = dat                    option 1
  { (↑FUNC) = SUBJ                 option 1a
    |(↑FUNC) = OBJth }             option 1b
  (↑a-str GO)
 |(↑CASE) = acc                    option 2
  (↑FUNC) = OBJ
  (↑SEM-PROP SPECIFIC) = + }.
```

The dative subject construction can now be analyzed on a par with nominative or ergative subject constructions. In each case, the basic subcategorization frame for transitive verbs is determined by the template TRANS and nothing special need be stipulated about the case marking. Compare the entry

needed for (31) in (32) with the entry for *kʰol* ‘open’ in (23). The only substantial difference is that a goal argument (ARG-GO) is specified instead of an agent (ARG-AG).[9]

(31) nadya=ko dɑr lɑg-a
Nadya.F=Dat fear.M.Nom be-attached-Perf.M.Sg
‘Nadya was afraid.’

(32) lag V @(TRANS be-attached)
@ARG-GO @ARG-TH
@ARG-REL.

The f-structure and a-structure analyses for (31) are shown in (33) and (34). The correct case marking falls out from the interaction of the information specified in the entry of the case marker ((30)), the subcategorization frame (f-structure), and the argument roles (a-structure). In the a-structure in (34) the GO(al) argument is associated with *Nadya*, the SUBJ of the f-structure in (33), satisfying the requirements of the dative case clitic *ko* shown in (30).

(33)

```
1 valid F-structure for ROOT:145
kill prev next Commands Views
lock F-structure #1 a::*
"nAdyA ko dar lagA"

    PRED     'be-attached<[0:nAdyA], [17:dar]>'
             PRED      'nAdyA'
        15   NTYPE     [PROPER name]
        16
SUBJ     0   P-FORM    {ko, nAdyA}
        62
        65   SEM-PROP  [SPECIFIC +]
        86   CASE dat, FUNC SUBJ, GEND fem, NUM sg, PERS 3
        17   PRED   'dar'
       103   P-FORM {dar}
OBJ    106
       109   CASE nom, FUNC OBJ, GEND masc, NUM sg, PERS 3
 29
119
123 P-FORM  {lag}
124
130 TNS-ASP [ASPECT perf]
131 VMORPH  [MTYPE infl]
145 PASSIVE -, STMT-TYPE decl, VFORM perf
```

(34)

```
a::*-structure window
kill Commands Views
a::*-structure #
GO   [ARG-FORM {ko, nAdyA}
      ROLE-O   -]
REL  [ARG-FORM {lag}]
TH   [ARG-FORM {dar}
      ROLE-O +, ROLE-R -]
```

5.2. *Ergative Subject Alternations*

The two case alternations shown in section 2.2 also fall out from this analysis of the Urdu ergative and dative. Here we consider the dative-ergative alternation in the infinitive construction, repeated in (35). The alternation in (8) essentially works the same way as this one, but as it involves complex predicate formation, we leave it aside for the purposes of this paper.

Recall that in this infinitive construction, the ergative is the marked form and entails a subject who has control over the action. The dative is the unmarked form or elsewhere case: the dative subject may or may not have control over the action, the precise interpretation depends on the context.

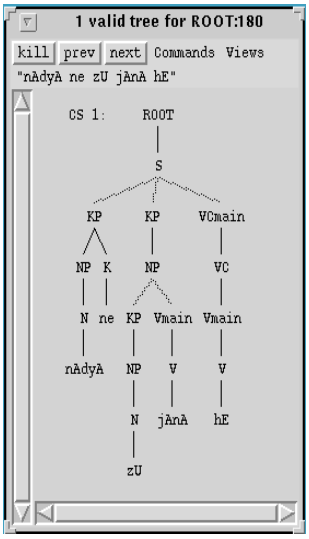
(35) a. nadya=ne zu ja-na hϵ
Nadya.F=Erg zoo.M.Loc go-Inf be.Pres.3.Sg
'Nadya wants to go to the zoo.'

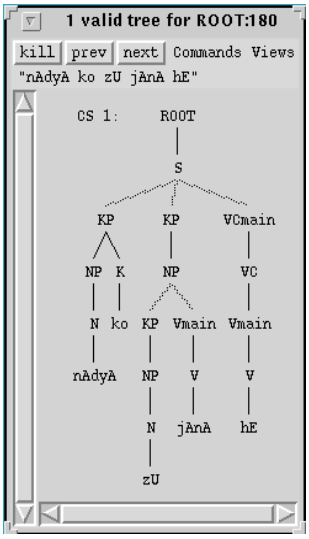
b. nadya=ko zu ja-na hϵ
Nadya.F=Dat zoo.M.Loc go-Inf be.Pres.3.Sg
'Nadya wants/has to go to the zoo.'

The complete entry for the ergative case marker *ne* is shown in (36). As can be seen, it is much simpler than the entry for the dative/accusative *ko* in (30). This is primarily because ergatives can only appear on subjects.

(36) ne K * @(PWORD ne)
(↑CASE) = erg
(↑FUNC) = SUBJ
$(\uparrow_{a-str}$ AG)

The semantic contrast in the case alternation in (35) is accounted for in our approach by an association of the ergative with volitionality or internal control over an action (SEM-PROP CONTROL internal), (38), while the dative is underspecified with respect to this semantic property, (39). Note that the phrase structures for the two constructions are strikingly similar, as shown in (37), differing only in the case marker.

(37) a. 

b. 

(38)

```
2 valid F-structures for ROOT
kill prev next Commands Views
lock F-structure #1 a::*
"nAdyA ne zU jAnA hE"

PRED    'be<[0:Nadya], [31:go]>'
        PRED     'Nadya'
        NTYPE    [PROPER name]
SUBJ    P-FORM   {Nadya, ne}
        SEM-PROP [CONTROL internal, SPECIFIC +]
      0 CASE erg, FUNC SUBJ, GEND fem, NUM sg, PERS 3
        PRED   'go<[0:Nadya], [16:zoo]>'
        SUBJ   [0:Nadya]
                PRED     'zoo'
                NTYPE    [GRAIN mass]
        OBL     P-FORM   <zoo>
XCOMP           SEM-PROP [LOCATION inherent]
     16         CASE loc, FUNC OBL, GEND masc, NMORPH obl, PERS 3
        P-FORM <jA>
        VMORPH [MTYPE infl]
     31 CASE nom, GEND masc, NUM sg, PASSIVE -, VFORM inf
P-FORM  <hO>
TNS-ASP [TENSE pres]
VMORPH  [MTYPE infl]
 51 FIN +, NUM sg, PASSIVE -, PERS 3, STMT-TYPE decl, VFORM pres
```

(39)

```
2 valid F-structures for ROOT
kill prev next Commands Views
lock F-structure #1 a::*
"nAdyA ko zU jAnA hE"

PRED    'be<[0:Nadya], [31:go]>'
        PRED     'Nadya'
        NTYPE    [PROPER name]
SUBJ    P-FORM   {Nadya, ko}
        SEM-PROP [SPECIFIC +]
      0 CASE dat, FUNC SUBJ, GEND fem, NUM sg, PERS 3
        PRED   'go<[0:Nadya], [16:zoo]>'
        SUBJ   [0:Nadya]
                PRED     'zoo'
                NTYPE    [GRAIN mass]
        OBL     P-FORM   <zoo>
XCOMP           SEM-PROP [LOCATION inherent]
     16         CASE loc, FUNC OBL, GEND masc, NMORPH obl, PERS 3
        P-FORM <jA>
        VMORPH [MTYPE infl]
     31 CASE nom, GEND masc, NUM sg, PASSIVE -, VFORM inf
P-FORM  <hO>
TNS-ASP [TENSE pres]
VMORPH  [MTYPE infl]
 51 FIN +, NUM sg, PASSIVE -, PERS 3, STMT-TYPE decl, VFORM pres
```

The verb ‘be’ is treated as a main verb in the analyses above, it takes a subject and an infinitival complement (XCOMP), the subject of the matrix ‘be’ (‘Nadya’) controls the subject of the infintival. As infinitivals in Urdu can themselves be case marked and inflect for number and gender, the XCOMP carries this information as well as other information.

Now, consider the appearance of the ergative on certain (unergative) intransitive verbs. Here, the ergative alternates with a nominative subject. However, as with the examples above, the ergative here

denotes volitionality on the part of the subject.

(40) a. ram kʰãs-a
Ram.M.Nom cough-Perf.M.Sg
'Ram coughed.'

b. ram=ne kʰãs-a
Ram.M=Erg cough-Perf.M.Sg
'Ram coughed (purposefully).'

The alternation here is possible for unergatives, but not for unaccusatives because only unergatives are specified for an agent in their lexical entries. An unaccusative verb such as 'fall' will not contain an agent in the lexical entry and therefore fail in combination with an ergative case marker.

The constraint that the ergative is required in conjunction with perfect morphology for transitives and unergatives must be stipulated as part of the verbal morphology. If perfect morphology is encountered and the verb is not unaccusative, then the subject must be in the ergative.

The interaction between various modules of the grammar and the information in the lexical entries is therefore complex. However, in exchange for this complexity, the lexical entries for verbs and case markers can be kept simple, and be expressed in very general terms, thereby doing justice to linguistic generalizations while offering greater grammar maintainability.

5.3. Instrumental Subjects — Ability "Passives"

The approach detailed above also extends to other types of non-nominative subjects such as the examples found with constructions generally labeled "(dis)ability passives" in grammars (e.g., Glassman 1976, Van Olphen 1980). Some examples are shown in (41).

(41) a. nadya=se yɛ ʊrdu=k-i cɪt̪ʰi pɑr̪ʰ-i nahĩ ja-ti
Nadya.F.Sg=Inst this Urdu=Gen-F.Sg letter.F.Sg.Nom read-Impf.F.Sg not go-Impf.F.Sg
'Nadya does not have the ability to read this Urdu letter.'

b. ʊs=se cɑl-a nahĩ ja-e-g-a
Pron=Inst walk-Perf.M.Sg not go-3-Fut-M.Sg
'She/he can't possibly walk.' (in the context of a broken leg) (Glassman 1976:275)

The instrumentals can be shown to be subjects on the basis of such tests as control and binding (Butt 1997). The construction has the effect of predicating of the subject the ability or disability to perform a certain action and has been termed a "passive" because it resembles the standard form of the passive in terms of surface morphology. Compare the passive of (42a) in (42b) with the examples in (41). Note that in both types of examples, there is a "perfect" morpheme on the main verb (*-a*) and that the verb *ja* 'go' is used as the second verb in the sequence. However, there is a difference in that the 'go' in (41) appears to be functioning as a light verb (see Butt 1995, 1997), while the 'go' in (42) is a passive auxiliary. Also note the difference in preferred word order: the instrumental is an adjunct in (42) and is preferentially placed immediately before the verb. In (41) the instrumental is a subject and preferentially occurs as the first noun phrase in the clause.

(42) a. nadya=ne yassin=ko mar-a
Nadya.F=Erg Yassin.M=Acc hit-Perf.M.Sg
'Nadya hit Yassin.'

b. yassin nadya=se mar-a gɛ-ya
Yassin.M.Nom Nadya.F.Inst hit-Perf.M.Sg go-Perf.M.Sg
'Yassin was hit by Nadya.'

Because the standard passive and the ability construction are so closely related in form, we detail a treatment of the standard passive first, and then contrast that with an analysis of the instrumental subject of the ability construction.

5.3.1. Traditional Passives

The passive is the classical example of a lexical rule in LFG (Bresnan 1982). The passive rule is taken to demote the subject and then to promotes the object to subject. This is shown schematically in (43) in terms of LFG notation.

(43) (↑PASSIVE)=+
(↑SUBJ)⟶NULL
(↑OBJ)⟶(↑SUBJ)

The effect of this lexical rule is illustrated in (44) for an English example. The input to the lexical rule is shown in (44b), the output of the lexical rule is shown in (45b). Note that any other arguments the verb has are unaffected by the lexical rule. This is true for the thematic roles as well — the underlying thematic role specification remains unchanged.

(44) a. They told me that Mary would arrive today.

b. PRED = 'tell<SUBJ,OBJ,COMP>'

(45) a. I was told that Mary would arrive today.

b. PRED = 'tell<NULL,SUBJ,COMP>'

Thus, Lexical rules relate different subcategorization frames to one another by a regular process of argument alternation (Kaplan and Bresnan 1982).

Now, let us return to our implementation for Urdu. The basic rule used in the Urdu grammar is shown in (46). This lexical rule is called by all the verbs in the lexicon which allow passivization (i.e., transitives and ditransitives).

(46) PASS(_SCHEMATA) =
{ _SCHEMATA `option 1`
(↑PASSIVE) = –
|_SCHEMATA `option 2`
(↑PASSIVE) = +
(↑OBJ)⟶(↑SUBJ)
(↑SUBJ)⟶NULL
($\uparrow_{arg-str}$ AG DEMOTED) = +}.

The PASS lexical rule takes a subcategorization frame, a SCHEMATA, for its argument; this frame can have any number of arguments of any type. The first disjunct (`option 1`) does nothing to the subcategorization frame, but adds the information that the f-structure is not passive. The second disjunct (`option 2`) changes the OBJ to the SUBJ and demotes the original SUBJ to NULL; in addition, it states that the f-structure is passive and that the AG in the a-structure is demoted. The structures for the passive in (47) are shown in (48) and (49).

(47) yassin nadya=se mar-a gɛ-ya
yassin.M.Nom Nadya.F=Inst hit-Perf.M.Sg go-Perf.M.Sg
'Yassin was beaten by Nadya.'

(48)

```
1 valid F-structure for ROOT:198
kill prev next Commands Views
lock F-structure #1 a::*
"yassin nAdyA se mArA gayA"

    PRED     'hit<NULL, [0:yassin]>'
             PRED     'yassin'
             NTYPE    PROPER name
    SUBJ   0 P-FORM   {yassin}
          89
          92 SEM-PROP SPECIFIC +
          97 CASE nom, FUNC SUBJ, GEND masc, NUM sg, PERS 3
             PRED     'nAdyA'
          30 NTYPE    PROPER name
          31
    ADJUNCT 15 P-FORM {nAdyA}
             {se    }
         115
 51      118 SEM-PROP SPECIFIC +
256      137 ADJUNCT-TYPE directional, CASE inst, GEND fem, NUM sg, PERS 3
 32
154 P-FORM {jA }
156        {mAr}
249 TNS-ASP ASPECT perf
251
184 VMORPH  MTYPE infl
198 PASSIVE +, STMT-TYPE decl, VFORM perf
```

(49)

```
a::*-structure window
kill Commands Views
a::*-structure #
    AG   ARG-FORM NULL, DEMOTED +, ROLE-O -
    REL  ARG-FORM {jA }
                  {mAr}
    TH   ARG-FORM {yassin}
         ROLE-O -, ROLE-R -
```

Note that although the original SUBJ is no longer an argument in the f-structure, but an adjunct, it is still the AG, albeit demoted, in the a-structure. Demoted agents in Urdu generally are case-marked with the instrumental. This is encoded in the lexical entry of the instrumental case marker in (50). The template FUNC-ADJUNCT marks the grammatical function associated with this case marker as an adjunct of the clause and thus distinguishes it from core arguments such as subject and object.

(50) se K * @(PWORD se)
(↑CASE) = inst
@FUNC-ADJUNCT

This entry for the instrumental *se* does not as yet account for the ability readings introduced above. The next section details an account of this type of non-nominative subject.

5.3.2. Analysis of Ability Readings

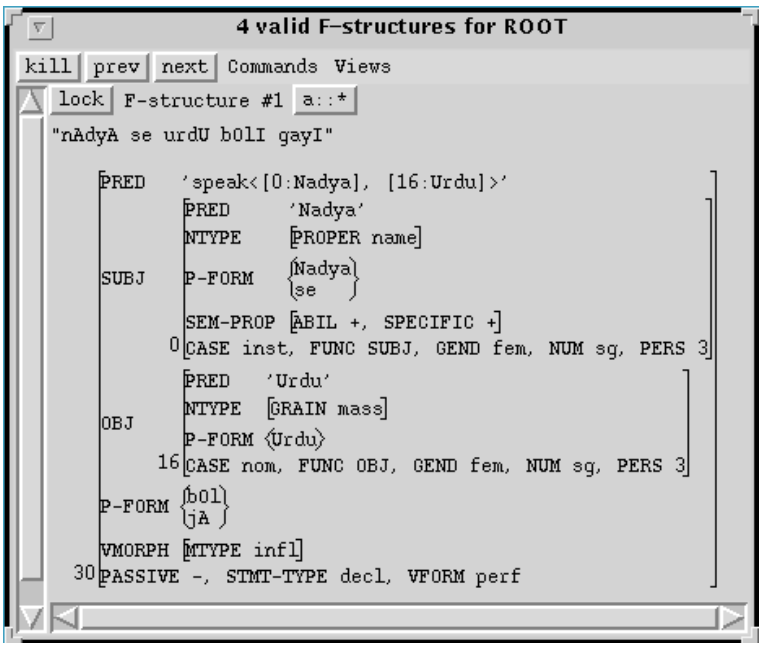
In order to be able to treat the ability construction, the instrumental case marker needs to be modified. The entry in (51) has been expanded by a disjunct which allows instrumental subjects in addition to adjuncts. However, if the subject option is chosen (`option 2`), the semantics associated with the subject must be that of an ability reading. The appearance of the instrumental on subjects is thus constrained by means of the semantic specification because the information in this disjunct interacts with verbal lexical entries (argument structure, further semantic and structural specifications) in a manner parallel to that of the dative and the ergative.

(51) se K * @(PWORD se)
(↑CASE) = inst
{ @FUNC-ADJUNCT `option 1`
|(↑FUNC) = SUBJ `option 2`
(↑SEM-PROP ABIL) = +}.

The structures in (53) and (54) show an analysis of the ability construction in (52).

(52) nadya=se ʊrdu bol-i gɛ-yi
Nadya.F=Inst Urdu.F.Nom speak-Perf.F.Sg go-Perf.F.Sg
'Nadya was able to speak Urdu.'

(53)



(54)

```
a::*-structure window
kill  Commands  Views
a::*-structure #
[ AG        [ARG-FORM {Nadya, se}
            ROLE-O -, ROLE-R -]
  EVENT-T   [REL [ARG-FORM jA]
            TH  [???MOTHER-a::-EVENT-T-TH]]
  REL       [ARG-FORM {bol, jA}]
  TH        [ARG-FORM {Urdu}
            ROLE-O +, ROLE-R -]
-2 ]
```

Note that although the argument structure shows an agent and a theme argument, an ergative, for example, would not be compatible with this construction. This is because the ergative would have to interact with the specifications provided by the verb *ja* 'go', which only license an ability reading when used in this construction, i.e. as a light verb in conjunction with a main verb with "perfect" morphology.[10]

Thus, in this section we have shown how certain non-nominative subjects are accounted for in our LFG implementation of a small-scale Urdu grammar. Since the case markers themselves contain syntactic and semantic information, as well as case information, they interact with the requirements of various constructions, resulting in the extensive system of non-nominatives subjects found in Urdu.

6. Conclusion

In this paper, we have presented a theoretical approach to non-nominative subjects in Urdu which involves a "constructive" treatment of case. Under our analysis, the case markers themselves are specified for structural and semantic information. This information interacts with information specified in other parts of the grammar (primarily the verbal lexical entries) in order to produce wellformed analyses. The existence of several semantically motivated case alternations in Urdu points to the need for incorporating semantic information into any approach to Urdu case marking. Under our analysis, Urdu case markers such as the ergative *ne*, the dative/accusative *ko* and the instrumental *se* can be seen as *semantic cases* in the sense that they help express semantically motivated alternations. However, they are also subject to structural considerations, such as the precise type of morphology or light verb they may appear with.

The paper also showed that this theoretically motivated approach to case is computationally viable in the sense that the resulting analyses are constrained in just the right way. We discussed the integration of argument structure into the implementation and showed how this level of representation could facilitate the formulation of a non-stipulative and generally applicable account of case-marking and semantically conditioned case alternations. The account is both linguistically satisfying and computationally feasible. In particular, no special templates need to be written for the Urdu verbs beyond specifying basic information about valency such as transitivity or intransitivity. Nor do the templates need to contain explicit information about possible case marking patterns. This is a desirable result as subjects in Urdu can potentially be marked with all of the available case markers in the language.

Endnotes

[1] Pronouns behave exceptionally as they may take an inflectional dative/accusative *-e* instead of the dative/accusative clitic *ko*.

[2] Because Urdu case markers are clitics and hence semi-independent elements, we use the notion of KP (KasePhrase) to encode case-marked noun phrases. The VC stands for "verbal complex".

[3] Technically, this is accomplished by inside-out functional uncertainty constraints associated with the cases (Halvorsen and Kaplan 1988; Dalrymple 1993; Andrews 1996). However, as inside-out functional uncertainty is not yet implemented in XLE (Xerox Linguistic Environment), the computational platform for our implementation, we mimic this behavior in case marking with the FUNC attribute.

[4] We do not discuss the details of the LFG formalism here; these can be found in Bresnan 2001 and references therein. Basically, the up arrows (↑) encode mappings between nodes of the phrase structure tree and the functional-structure. The '↑' refers to the particular AVM that the phrase structure node in question corresponds to. So, in the examples in this paper, the '↑' refers to the functional-structure of the noun phrase containing the case marker. For example, in (55) the '↑' refers to the AVM with PRED 'dog' in it in (56); thus, the first line of (55) states that this part of the functional-structure contains the pair CASE ERG, as is seen in (56), while the second states that this part of the functional-structure is contained within the SUBJ of the next bigger AVM, as is also seen in (56).

[5] Our notion of semantic case must be distinguished from a more generally known usage in which semantic case is taken to refer to case on adjuncts, but not structural arguments. Under our approach, semantic case represents a mixture of structural and semantic constraints, all generalizable properties of the language. Quirky case then is restricted to lexically, inherently stipulated case that does not follow from any generalizable property in the language. An example of this is the Urdu verb *la* 'bring' which, being a transitive verb, should take an ergative subject with perfect morphology but instead takes a nominative.

(i) nadya kɩtab la-yi
Nadya.F.Nom book.F.Nom bring-Perf.F.Sg
'Nadya brought a book.' (Transitive; Perfect)

(ii) *nadya=ne kɩtab la-yi
Nadya.F=Erg book.F.Nom bring-Perf.F.Sg
'Nadya brought a book.' (Transitive; Perfect)

As there is no semantic or syntactic reason for this difference, the nominative case must be stipulated in the lexical entry of this particular verb.

[6] See footnote 4 on the meaning of the up arrows; in this paper, subscripted arrows refer to projections other than the functional-structure, namely the argument-structure.

[7] This is in contrast to Butt, Dalrymple and Frank (1997), who assume the projection architecture in (i).

(i) c-str ⟶ a-str ⟶ f-str ⟶ s(emantic)-str

[8] Examples of templates are seen below. Templates are a implementational device that facilitate the grouping of information. For example, it would be possible to put all the information in (25) in the lexical entry of every verb with an agent argument. However, this would lead to maintenance problems whenever a minor change in how agents are treated was made, as well as increase the changes of making typographical errors. So, instead a template is created containing all the relevant information, and the lexical entries call the template. See Butt et al. 1999 for more on templates in XLE.

[9] For completeness, the templates called in (32) are provided in (i)–(iii) (TRANS was shown in (24)).

i. ARG-GO = { (↑$_{arg-str}$ GO ARG-FORM) = (↑SUBJ P-FORM) `option 1`
(↑$_{arg-str}$ GO ROLE-O) = −
(↑SUBJ CASE) =c dat
(↑$_{arg-str}$ AG)
|(↑$_{arg-str}$ GO ARG-FORM) = (↑OBJth P-FORM) `option 2`
(↑$_{arg-str}$ GO ROLE-O)=+
(↑$_{arg-str}$ GO ROLE-R)=+
(↑OBJth CASE)=c dat}.

ii. ARG-TH = (↑$_{arg-str}$ TH ROLE-R) = −
{ (↑$_{arg-str}$ TH ARG-FORM)=(↑SUBJ P-FORM) `option 1`
(↑SUBJ)
(↑$_{arg-str}$ TH ROLE-O) = −
|(↑$_{arg-str}$ TH ARG-FORM) = (↑OBJ P-FORM) `option 2`
(↑$_{arg-str}$ TH ROLE-O) = +
(↑OBJ)
|(↑$_{arg-str}$ TH)=(↑$_{arg-str}$ EVENT) `option 3`
(↑COMP)}.

iii. ARG-REL = ($\uparrow_{arg-str}$ REL ARG-FORM)=($\uparrow$P-FORM).

[10]The relevant lexical entry for ‘go’ is shown in (i). It is `option 2b` which assigns the subject the required ability semantics.

i.	jA	V	*. . .main verb uses. . .*	`option 1`
		Vlight	@(PWORD jA)	`option 2`
			{ ($\uparrow$TNS-ASP COMPLETIVE) = +	`option 2a`
			($\uparrow$VTYPE) =c unacc	
			| @(ARG-EV-T jA)	`option 2b`
			@ARG-TH-AG-MERGER	
			($\uparrow$VFORM) =c perf	
			($\uparrow$SUBJ SEM-PROP ABIL) = +}	
		AUXpass	*. . .passive auxiliary use. . .*	`option 3`

REFERENCES

Alex Alsina. 1996. *The Role of Argument Structure in Grammar*. CSLI Publications, Stanford, California.

Avery Andrews. 1996. Semantic case-stacking and inside-out unification. *Australian Journal of Linguistics*, 16:1–55.

Elena Bashir. 1999. The Urdu postposition *ne*: its changing role in the grammar. *The Yearbook of South Asian Languages and Linguistics*. To appear.

Joan Bresnan and Annie Zaenen. 1990. Deep unaccusativity in LFG. In Katarzyna Dziwirek, Patrick Farrel, and Errapel Mejias-Bekandi, editors, *Grammatical Relations: A Cross-Theoretical Perspective*, pages 45–57. CSLI Publications, Stanford, California.

Joan Bresnan, editor. 1982. *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge, Massachusetts.

Joan Bresnan. 2001. *Lexical-Functional Syntax*. Blackwell, Oxford.

Miriam Butt and Tracy Holloway King. 2001a. The status of case. In Veneeta Dayal and Anoop Mahajan, editors, *Clause Structure in South Asian Languages*. To appear.

Miriam Butt and Tracy Holloway King. 2001b. Case systems: Beyond structural distinctions. In Ellen Brandner and Heike Zinsmeister, editors, *New Perspectives on Case Theory*. CSLI Publications, Stanford, California. To Appear.

Miriam Butt, Mary Dalrymple, and Anette Frank. 1997. An architecture for linking theory in LFG. In *Proceedings of the LFG97 Conference*, San Diego, California, July. `http://csli-publications.stanford.edu/LFG2/lfg97.html`.

Miriam Butt, Tracy Holloway King, Maria-Eugenia Niño, and Frédérique Segond. 1999. *A Grammar Writer’s Cookbook*. CSLI Publications, Stanford, California.

Miriam Butt. 1993. Object specificity and agreement in Hindi/Urdu. In *Papers from the 29th Regional Meeting of the Chicgo Linguistic Society*, volume 1, pages 80–103.

Miriam Butt. 1995. *The Structure of Complex Predicates in Urdu*. CSLI Publications, Stanford, California.

Miriam Butt. 1997. Aspectual complex predicates, passives and disposition/ability. Talk given at the LAGB97. Extensive handout available at:
`http://ling.uni-konstanz.de/pages/home/butt`.

Mary Dalrymple. 1993. *The Syntax of Anaphoric Binding*. CSLI Publications, Stanford, California.

Alice Davison. 1999. Ergativity: Functional and formal issues. In Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan, and Kathleen Wheatley, editors, *Functionalism and Formalism in Linguistics, Volume I: General Papers*. John Benjamins, Amsterdam.

Helen de Hoop. 1992. *Case Configuration and Noun Phrase Interpretation*. Ph.D. thesis, Rijksuniversiteit Groningen.

Mürvet Enç. 1991. The semantics of specificity. *Linguistic Inquiry*, 22(1):1–25.

Eugene Glassman. 1976. *Spoken Urdu*. Nirali Kitaben, Lahore.

Jane Grimshaw. 1990. *Argument Structure*. The MIT Press, Cambridge, Massachusetts.

Per-Kristian Halvorsen and Ronald M. Kaplan. 1988. Projections and semantic description in Lexical-Functional grammar. In *Proceedings of the International Conference on Fifth Generation Computer Systems (FGCS-88)*, pages 1116–1122, Tokyo, Japan. Reprinted in Mary Dalrymple, Ronald M. Kaplan, John Maxwell, and Annie Zaenen, eds., *Formal Issues in Lexical-Functional Grammar*, 279–292. CSLI Publications, Stanford, California. 1995.

John T. Maxwell III and Ron Kaplan. 1991. A method for disjunctive constraint satisfaction. In M. Tomita, editor, *Current Issues in Parsing Technology*, pages 173–190. Kluwer Academic Publishers, Boston.

John T. Maxwell III and Ron Kaplan. 1993. The interface between phrasal and functional constraints. *Computational Linguistics*, 19:571–590.

Alana Johns. 2000. Ergativity: A perspective on recent work. In Lisa Cheng and Rint Sybesma, editors, *The First* Glot International *State-of-the-Article Book*, pages 47–73. Mouton de Gruyter, Berlin.

Ronald Kaplan and Joan Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*, pages 173–281. The MIT Press, Cambridge, Massachussetts. Reprinted in Mary Dalrymple, Ronald M. Kaplan, John Maxwell, and Annie Zaenen, eds., *Formal Issues in Lexical-Functional Grammar*, 29–130. CSLI Publications, Stanford, California. 1995.

Ron Kaplan and Martin Kay. 1994. Regular models of phonological rule systems. *Computational Linguistics*, 20:331–378.

Ron Kaplan and Juergen Wedekind. 2000. LFG generation produces context-free languages. In *Proceedings of COLING 2000, Saarbrueken*, pages 141–148.

Ronald M. Kaplan. 1995. Three seductions of computational psycholinguistics. In Mary Dalrymple, Ronald M. Kaplan, John Maxwell, and Annie Zaenen, editors, *Formal Issues in Lexical-Functional Grammar*, pages 339–368. CSLI Publications, Stanford, California.

Lauri Karttunen, Ron Kaplan, and Annie Zaenen. 1992. Two level morphology with composition. In *Proceedings of COLING 92, Nantes*, volume I, pages 141–148.

Tracy Holloway King. 1997. Focus domains and Information-Structure. In *Proceedings of the LFG 1997 Conference*. CSLI On-line Publications.

Anoop Mahajan. 1990. *The A/A-Bar Distinction and Movement Theory*. Ph.D. thesis, MIT.

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

Tara Mohanan. 1994. *Argument Structure in Hindi*. CSLI Publications, Stanford, California.

Rachel Nordlinger. 1998. *Constructive Case: Evidence from Australian Languages*. CSLI Publications, Stanford, California.

Kevin Tuite, Asif Agha, and Randolph Graczyk. 1985. Agentivity, transitivity, and the question of active typology. In W. H. Eilfort, P. D. Kroeber, and K. L. Peterson, editors, *Papers from the Parasession on Causatives and Agentivity at the 21st Regional Meeting of the Chicago Linguistic Society*, pages 252–270.

Herman Van Olphen. 1980. *First-Year Hindi Course*. Department of Oriental and African Languages and Literatures, University of Texas, Austin, Texas, Austin.

M. K. Verma and K.P.Mohanan, editors. 1990. *Experiencer Subjects in South Asian Languages*. CSLI Publications, Stanford, California.

Ellen Woolford. 1997. Four-way case systems: Ergative, nominative, objective and accusative. *Natural Language and Linguistic Theory*, 15:181–227.

Annie Zaenen, Joan Maling, and Höskuldur Thráinsson. 1985. Case and grammatical functions: the Icelandic passive. *Natural Language and Linguistic Theory*, 3:441–483.