Complex Predicates via Restriction

Miriam Butt, Tracy Holloway King, and John T. Maxwell III
Universität Konstanz, PARC, and PARC

Proceedings of the LFG03 Conference
University at Albany, State University of New York
Miriam Butt and Tracy Holloway King (Editors)

2003

CSLI Publications
http://csli-publications.stanford.edu/

Abstract

This paper continues the discussion of the RESTRICTION OPERATOR (Kaplan and Wedekind, 1993) and whether it can provide a linguistically adequate solution to the problem posed by syntactic complex predicate formation. The solution introduced here has been implemented as part of an on-going project aimed at the development of a computational grammar for Urdu and can be shown to model the linguistic facts of syntactic complex predicate formation as described by (Alsina, 1996) and (Butt, 1995). This also allows for a straightfoward extension to related phenomena in other languages such as German, Japanese, Norwegian, and French.

1 Introduction

(Alsina, 1996), (Butt, 1995) and (Mohanan, 1994) argue that complex predicate (CP) formation in Romance and Urdu/Hindi,[1] respectively, should take place in the syntax.[2] While the linguistic facts point towards the need to combine two or more predicational elements into one predicational domain within the syntax, a computational treatment which models this theoretical idea has so far proved elusive. (Kaplan and Wedekind, 1993) proposed a solution in terms of a RESTRICTION OPERATOR. However, their solution had the effect of lexically stipulating combinatoric possibilities on a verb-by-verb basis. Given that CP formation is regular and productive, and given that different types of CPs may be stacked, this essentially lexical approach proved to be inadequate.

In this paper, we report on an implementation of CPs for Urdu in the Parallel Grammar (ParGram) project (Butt et al., 1999; Butt et al., 2002) which models the original observations by Alsina and Butt more satisfactorily. The basic tool for analysis remains the restriction operator, but its application has been reformulated to model productive CP formation in the syntax, instead of the within the lexicon.

2 South Asian Complex Predicates

South Asian languages are known for the extensive and productive use of CPs. CPs combine a light verb with a verb, noun, or adjective to produce a new verb. For example, Urdu has a large class of light verbs which combine with verbs to modulate the event predication in terms of aktionsart properties and more subtle semantic effects such as suddenness or forcefulness. Examples are shown in (1b,c) for V-V CPs. (1a) shows a simplex use of the same main verb.[3]

(1) a. nAdyA AyI
Nadya.Nom come.Perf.F.Sg
'Nadya came.'

b. nAdyA A gayI
Nadya.Nom come go.Perf.F.Sg
'Nadya arrived.'

c. nAdyA A paRI
Nadya.Nom come fall.Perf.F.Sg
'Nadya arrived (suddenly, unexpectedly).'

[1] Mohanan discusses N-V complex predicate formation, whereas this paper focuses on V-V complex predicates. We therefore leave aside N-V complex predication for the purposes of this paper.

[2] (Frank, 1996) argues against this position and places complex predicate formation within the lexical component. However, while the subcategorization requirements and the restrictions on combinations are specified within the lexical entries (as is done in this paper), her approach must still crucially combine the separate pieces of the complex predicate in the syntax. That is, they are not assumed to be combined within the morphological component.

[3] The transcription here follows the ASCII transcription used for the implementation of the Urdu grammar. This is to allow maximum readability of the c-structure and f-structure output of the grammar, samples of which are to follow. Long vowels are marked with a capital letter, as are retroflex consonants. Nasalization is marked with an N, aspiration with an H.

The semantic effects of the light verbs in combination with a main verb are subtle and have been described in a wealth of literature (see, e.g., (Hook, 1974) for an overview, (Butt, 1995) for an initial analysis within LFG, (Butt and Geuder, 2001) and (Butt and Ramchand, 2003) for more recent work). Within the scope of this paper, only those effects which are necessary for an understanding of the implementation are discussed. We thus simplify the observed semantics somewhat (this is necessary because at this time the Urdu grammar lacks a proper semantic component) and differentiate between two major classes of light verbs: light verbs like ‘go’ signify the completion of an action, whereas light verbs like ‘fall’ signify inception. This in line with the observed linguistic facts, though this is not all that the individual light verbs are capable of.

Although these light verbs do not alter the subcategorization frame of the verb, they change the resulting functional structure of the sentence, providing new information about the kind of event/action that is being described. The light verb also determines case marking on the subject: light verbs based on intransitive main verbs like *paR* ‘fall’ require a nominative subject. Light verbs like *lE* ‘take’ or *dE* ‘give’, which are based on (di)transitive main verbs, require an ergative subject. For example, transitive main verbs in the perfect tense usually require an ergative subject, as in (2a). When combined with a light verb like *paR* ‘fall’, the subject must be nominative as in (2b). Case marking in Urdu is governed by a combination of structural and semantic factors which we do not discuss here (Butt and King, 2001; Butt and King, 2003). The light verb facts present an extension of the basic pattern.

(2) a. nAdyA=nE gAnA gayA
Nadya=Erg song.Nom sing.Perf.M.Sg
‘Nadya sang a song.’

b. nAdyA gAnA gA paRI
Nadya.Nom song.Nom sing fall.Perf.F.Sg
‘Nadya burst into song.’

c. nAdyA=nE gAnA gA lIyA
Nadya=Erg song.Nom sing take.Perf.M.Sg
‘Nadya sang a song (completely).’

As already mentioned, these CPs are very productive in Urdu: most verbal predication involves complex predicate formation of the kind in (1) and (2). A light verb is in principle compatible with any main verb; however, (mostly semantic) selectional restrictions do apply so that some combinations are ruled out completely, whereas others are subject to considerable dialectal variation. Furthermore, the CPs are not formed within the lexicon, but are the result of the *syntactic composition* of two predicational elements (Alsina, 1996; Butt, 1995). Within LFG (as well as other syntactic frameworks), predicational elements play a special role: it is over these that argument saturation is checked. The difficulties involved with CP formation are better illustrated by means of another type of CP, the Urdu permissive, which contriibutes its own arguments to the joint predication (Butt, 1995). The permissive light verb contributes a permitter (agent) which is realized as a subject. The highest argument of the main verb must therefore be realized as a non-subject function: it surfaces as a dative-marked thematic object, as in (3b), cf. (3a).

(3) a. nAdyA sOyI
Nadya.Nom sleep.Perf.F.Sg
‘Nadya slept.’

b. yassin=nE nAdyA=kO sOnE dIyA
Yassin=Erg Nadya=Dat sleep.Inf.Obl give.Perf.M.Sg
‘Yassin let Nadya sleep.’

Since both types of CPs are productive and occur frequently, an implementation that is both scalable and effi cientis necessary. Most verbs are compatible with several different light verbs and these combinations in turn are compatible with further sets of light verbs. It is therefore not feasible to have multiple lexical entries for each verb depending on which light verb they occur with. This is especially true since the CPs also combine with auxiliaries in predictable ways. The problem to be resolved is thus how two verbs with indepedent predicational information can be combined to form a single predicational domain which can then interact with other elements of the syntax such as auxiliaries in exactly the same manner that a single verb would.

3 The ParGram Project

The ParGram project (Butt et al., 1999; Butt et al., 2002) originally focused on three European languages: English, French, and German. Three other languages were added later: Japanese, Norwegian, and Urdu. The ParGram project uses the XLE parser and grammar development platform (Maxwell and Kaplan, 1993) to develop deep grammars, i.e., grammars which provide an in-depth analysis of a given sentence (as opposed to shallow parsing or chunk parsing, where a relatively rough analysis of a given sentence is returned).

All of the grammars in the ParGram project use the Lexical-Functional Grammar (LFG) formalism. Given that f-structures are assumed to encode a language universal level of analysis, one of the aims of ParGram is to see how far f-structure parallelism can be maintained across languages. In the project, analyses for similar constructions across languages are therefore held as similar as possible and the conventions developed within the ParGram grammars are extensive. The ParGram project dictates not only the form of the features used in the grammars (Butt et al., 2003), but also the types of analyses chosen for constructions. This parallelism requires the formulation of a rigid standard for linguistic analysis. This standardization has the computational advantage that the grammars can be used in similar applications, and it can simplify cross-language applications such as machine translation (Frank, 1999).

4 Implementation

In this section, we discuss the XLE implementation of restriction and how it is used to analyze Urdu complex predicates.

4.1 Without Restriction: Lexical Rules

The XLE implementation in use when Urdu joined ParGram allowed for basic modifi cationsof predicates. In particular, it had an implementation of lexical rules that was suffi cientto handle the English passive: argument grammatical functions could be renamed or deleted.

An example of this is shown in (4) for the Urdu passive; the template is practically identical to that of English. In this template, _SCHEMATA indicates the predicate with grammatical functions of the verb (e.g., for transitive 'open': $'\text{kHOl}<(\uparrow\text{SUBJ})(\uparrow\text{OBJ})>'$). In the active, nothing happens (left disjunct) except that a PASSIVE $-$ feature is added. In the passive, the object becomes the subject and the original subject is deleted (right disjunct). Example outputs are shown in (5) for the verb $'\text{kHOl}<(\uparrow\text{SUBJ})(\uparrow\text{OBJ})>'$.

(4) PASS(_SCHEMATA) =

$$\begin{array}{lcl} \text{\_SCHEMATA} & \textbf{or} & \text{\_SCHEMATA} \\ (\uparrow \text{PASSIVE}) = - & & (\uparrow \text{OBJ}) \longrightarrow (\uparrow \text{SUBJ}) \\ & & (\uparrow \text{SUBJ}) \longrightarrow \text{NULL} \\ & & (\uparrow \text{PASSIVE}) = + \end{array}$$

(5) a. Active:
$'\text{kHOl}<(\uparrow\text{SUBJ})(\uparrow\text{OBJ})>'$
$(\uparrow\text{PASSIVE})=-$

b. Passive:
$'\text{kHOl}<\text{NULL},(\uparrow\text{SUBJ})>'$
$(\uparrow\text{PASSIVE})=+$

However, this operation over lexical items is not sufficient to cover Urdu CPs since the operations over predicate-argument structure that are necessary cannot be handled within the lexicon. (Kaplan and Wedekind, 1993) proposed that the problem of Urdu CPs is reminiscent of the head-switching type of structural mismatch discussed in the context of machine translation. Here, the predicate that is the head in one language (source language) must be rendered as an adverb or embedded predicate in another language (target language) and a different predicate is "elevated" to play the role of the head. (Kaplan and Wedekind, 1993) proposed that complex predication could be thought of in similar terms: although the main verb of a complex predicate provides the bulk of the predicational infomation, the light verb serves as the syntactic head of the construction in that it inflects for tense, etc. In order to solve the general problems associated with such structural mismatches, (Kaplan and Wedekind, 1993) introduced the notion of RESTRICTION. However, as first formulated for the Urdu permissive, the solution only allowed the application of the restriction operator within the lexicon and thus did not take into account the powerfully recursive nature of complex predication in Urdu, which allows different types of CPs to be stacked (Butt, 1994).

4.2 Restriction

A viable computational treatment of complex predication retreated into the background until the discussion arose again with respect to a special type of Norwegian passive. This brought the issue of complex predication into the forefront of the ParGram project discussions. As part of these, a solution was found in the implementation of restriction within XLE in which the restriction applies as part of the *syntactic composition* of two predicates.

Restriction allows f-structures and predicates to be manipulated in a controlled and detailed fashion. Given an f-structure like (6a), for example, it might be necessary to restrict out the case information (e.g., in order to assign some other case to the f-structure). In this situation, the restriction operator '/' can be applied to the current f-structure (↑/CASE) in order to arrive at the restricted f-structure in (6b). A restricted f-structure is thus identical to the original f-structure except that it does not contain the restricted attribute.

(6) a.
$$\begin{bmatrix} \text{PRED} & \text{'nAdyA'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{sg} \\ \text{CASE} & \text{erg} \end{bmatrix}$$
b.
$$\begin{bmatrix} \text{PRED} & \text{'nAdyA'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{sg} \end{bmatrix}$$

4.3 Event Modulating Complex Predicates

In this section, we discuss the rules necessary to analyze event modulating CPs with restriction in the Urdu grammar. A sample event modulating CP is shown in (7).

(7) nAdyA has paRI
Nadya.Nom laugh fall.Perf.F.Sg
'Nadya laughed (suddenly, unexpectedly).'

With these CPs, the light verb does not contribute an independent argument. Rather, it contributes information which pertains to the manner and type of the event (e.g., inceptive, telic) and which also thereby contributes information about the highest argument of the main verb (e.g., whether the action was deliberate or not) and which thus indirectly influences the case marking on the subject. In the current implementation, such essentially lexical semantic information is coded under the feature LEX-SEM. This feature must be restricted since CP formation can affect its values. Furthermore, the VTYPE must be restricted because the syntax needs to know that the verbal predication is that of a complex predicate, rather than that of a simplex verb.

The annotated phrase structure rule for these CPs is shown (8).[4]

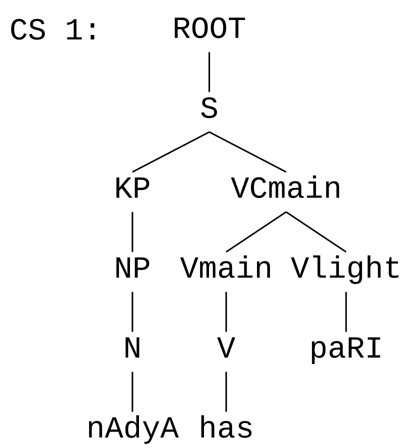
(8) Vasp $\longrightarrow$ Vmain Vlight

$\downarrow$/VTYPE/LEX-SEM=$\uparrow$/VTYPE/LEX-SEM $\downarrow$=$\uparrow$

The main verb and the light verb are co-heads in this construction: both are annotated $\downarrow$=$\uparrow$. The PRED is provided by the Vmain; Vlights have no PRED in their lexical entry. However, both Vmains and Vlights can have VTYPE values which may conflict. By restricting out the VTYPE, the final f-structure receives a VTYPE `complex-pred` which is provided by the lexical entry of the Vlight. This is shown in the f-structure in (10).

An example of the current analysis of the event modulating CP in (7) is shown in (9) and (10). The c-structure in (9) allows for a verbal complex which expands into a main verb followed by a light verb. There is no compelling evidence that Urdu has a VP (i.e., that a verb and its object are contained under one constituent); hence we do not assume one. Urdu is furthermore a language with fairly free word order, so the trees are quite flat: noun phrases are represented as sisters to one another under S (see the c-structures in (14) and (17)). We do assume KPs (Kase Phrases). Case markers in Urdu act as clitics to NPs (Butt and King, 2003), and as such have their own phrase structure node. In (9) the subject is nominative, which is phonologically null; so the KP has an empty head. A full KP can be seen in the c-structure analysis for the permissive in (14).

(9) C-structure tree for event modulating CP



[4]Some implementational details have been supressed here. For example, a grammar internal feature called CHECK is restricted out; the CHECK feature is used in CPs to make sure that constraining equations for the case markers are satisfied on the final f-structure and not on the restricted one. This can be seen in the sample f-structure in (10).

In addition, the rules are shown in cannonical LFG notation and not in the ASCII friendly XLE notation.

(10) F-structure AVM for event modulating CP
"nAdyA has paRI"

```
  PRED      'has<[0:Nadya]>'
            PRED      'Nadya'
            NTYPE     [PROPER name]
  SUBJ      SEM-PROP  [SPECIFIC +]
          0 CASE nom, GEND fem, NUM sg, PERS 3
  CHECK     [_VMORPH [_MTYPE infl]]
  LEX-SEM   [AGENTIVE -]
  TNS-ASP   AKTIONSART [INCEPTIVE +, TELIC +]
            ASPECT perf, MOOD indicative
19 PASSIVE -, STMT-TYPE decl, VTYPE complex-pred

  PRED      'has<[0:Nadya]>'
  SUBJ      [0:Nadya]
  CHECK     [_RESTRICTED +]
  TNS-ASP   [19-TNS-ASP]
14 LEX-SEM  unerg, PASSIVE -, STMT-TYPE decl, VFORM bare
```

The top f-structure in (10) represents the final analysis of the CP. The bottom f-structure shows the f-structure of the main verb *has* 'laugh'. The features which have been restricted from the main verb's f-structure are VTYPE and LEX-SEM because these are the features which the light verb can "overwrite". In the case of (10), the VTYPE feature is provided entirely by the light verb. Within the ParGram project, the feature X-TYPE is used to encode distinctions within a given category X which are useful at the f-structure level of analysis. The English grammar, for example, encodes different kinds of adverbs (sentential, degree modifiers, etc.) via the feature ADV-TYPE. In the Urdu grammar, we use the feature VTYPE to register the type of the verbal predication. So, in (10), the final top structure has VTYPE `complex-pred`. The lower structure for the main verb has LEX-SEM `unerg` because *has* 'laugh' by itself is an unergative verb; this is restricted and does not appear in the final CP analysis.

The light verb provides all of the TNS-ASP features in the final f-structure of this example. The AKTIONSART feature contains the information that the event is both inceptive and telic. The general effect of this type of complex predication is the denotation of a result state (a song is in the state of having been sung, a person is in the state of having arrived or having laughed). However, a result state can be interpreted in two differing ways depending on whether one wants to focus on the inception or the completion of the event (the two concepts seem to be orthogonal). The precise interpretation is lexically determined by the light verbs. For the purposes of the Urdu grammar, we mark light verbs like 'go' as signifying completion of an action, whereas light verbs like 'fall' signify inception.

4.4 Permissive Complex Predicates

The restriction operation for permissive CPs is more interesting from a technical point of view because both the verbs contribute participants independently to the overall predicate-argument structure. A sample permissive with an intransitive verb is shown in (11).

(11) yassin=nE nAdyA=kO sOnE dIyA
Yassin=Erg Nadya=Dat sleep.Inf give.Perf.M.Sg
'Yassin let Nadya sleep.'

In this type of CP, the event denotations of the two verbs are not as closely intertwined. In the previous examples, the light verb serves to modulate the event structure of the main verb. In this example, a permissive event contains the information that another event was allowed to take place. In theoretical terms and in terms of our implementation, this can be modeled by specifying that the light verb *dE* has a PRED which takes as its second argument the main verb's PRED. The lexical entry for *dE*'s PRED is shown in (12).

(12) (↑ PRED)='dE<(↑ SUBJ),%PRED2>'

Since both the main verb and the light verb have a PRED, the PRED must be restricted out in the permissive CPs. The basic rule for this is in (13).

(13) V ⟶ V Vlight
↓/PRED/SUBJ/VTYPE/LEX-SEM=↑/PRED/SUBJ/OBJ-GO/VTYPE/LEX-SEM ↑=↓
(↑ PRED ARG2)=(↓ PRED)
(↑ OBJ-GO)=(↓ SUBJ)

First consider the similarities between the rule for the permissive CPs in (13) and that for the event modulating CPs in (8). In both cases, the verbs are co-heads, as indicated by the ↑=↓ annotations. Also, the main verb has the VTYPE and LEX-SEM restricted since the CP as a whole may have a different VTYPE and LEX-SEM than that of the main verb. Next consider the differences between the two rules. In the permissive, both the main verb and the light verb have a PRED. As such, the PRED is restricted out from the main verb. In addition, the SUBJ of the main verb is not the SUBJ of the CP. As such, the SUBJ is also restricted. The two equations under the restriction equation state how the grammatical functions of the main verb map onto the grammatical functions of the CP. First, the PRED of the main verb is assigned to the second argument of the CP. Second, the SUBJ of the main verb is assigned to the OBJ-GO (goal restricted object) of the CP. Any other grammatical functions of the main verb will remain the same in the CP. This will be illustrated below when the main verb is transitive; in this case, the main verb's OBJ is also the OBJ of the CP.

The results of the application of the rule in (13) are shown in the resulting f-structures in (15) for an intransitive main verb and in (18) for a transitive main verb. First consider the structures for the intransitive main verb in (11) which are shown in (14) and (15).

(14) C-structure tree for permissive CP

```
CS 1:     ROOT
           |
           S
    _______|________
   KP     KP      VCmain
  /  \   /  \    /      \
 NP  K  NP  K  Vmain Vlight-prec
 |   |  |   |    |        |
 N  nE  N  kO    V       dIyA
 |      |        |
yassin nAdyA   sOnE
```

(15) F-structure AVM for permissive CP

```
"yassin nE nAdyA kO sOnE dIyA"

   PRED      'dE<[0:Yassin], 'sO<[16:Nadya]>'>'
             PRED      'Yassin'
             NTYPE     [PROPER name]
   SUBJ      SEM-PROP  [SPECIFIC +]
           0 CASE erg, GEND masc, NUM sg, PERS 3
             PRED      'Nadya'
             NTYPE     [PROPER name]
   OBJ-GO    SEM-PROP  [SPECIFIC +]
          16 CASE dat, GEND fem, NUM sg, PERS 3
   CHECK     [_VMORPH [_MTYPE infl]]
   LEX-SEM   [AGENTIVE +, GOAL +]
   TNS-ASP   [ASPECT perf, MOOD indicative]
51 GEND masc, NUM sg, PASSIVE -, PERS 3, STMT-TYPE decl, VTYPE complex-pred

   PRED    'sO<[16:Nadya]>'
   SUBJ    [16:Nadya]
   CHECK   [_NMORPH obl, _RESTRICTED +]
32 GEND masc, LEX-SEM unerg, NUM sg, PASSIVE -, PERS 3, STMT-TYPE decl, VFORM inf
```

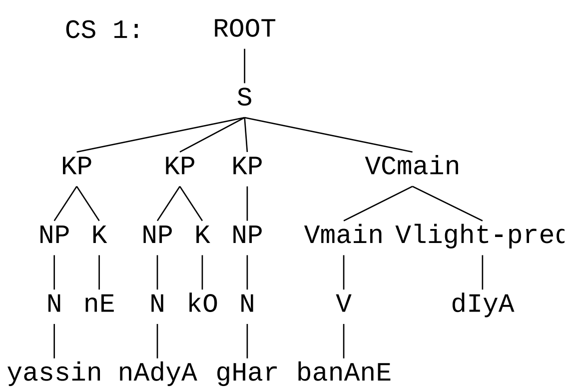
Recall that the light verb *dE* 'give' effectively adds a subject argument and demotes the subject of the main verb to an OBJ-GO. Restricting the PRED and SUBJ of the main verb's f-structure allows the final f-structure to assign new grammatical functions when necessary, i.e., to demote the SUBJ Nadya to an OBJ-GO and to inherit any remaining arguments of the main verb. The light verb *dE* 'give' subcategorizes for a subject (the permitter) and a predicate, as seen in the lexical entry in (12). In (15), the PRED feature has the value of a *composite* argument structure, namely a combination of the subcategorization frame of *dE* 'give' (subject and another predicate) and the subcategorization frame of *sO* 'sleep' modulo the operations licensed via the restriction operator.

In (15), the main verb was the intransitive *sO* 'sleep' and so there were no arguments for the CP to inherit other than the demoted subject. The analysis in (18) shows what happens with a transitive main verb like *banA* 'make'.

(16) yassin=nE nAdyA=kO gHar banAnE dIyA
Yassin=Erg Nadya=Dat house.Nom make.Inf give.Perf.M.Sg
'Yassin let Nadya build a house.'

yassin=nE	nAdyA=kO	gHar	banAnE	dIyA
Yassin=Erg	Nadya=Dat	house.Nom	make.Inf	give.Perf.M.Sg

(17) C-structure tree for permissive CP



(18) F-structure AVM for permissive CP

```
"yassin nE nAdyA kO gHar banAnE dIyA"

   PRED       'dE<[0:Yassin], 'banA<[16:Nadya], [32:gHar]>'>'
              PRED      'Yassin'
              NTYPE     [PROPER name]
   SUBJ       SEM-PROP  [SPECIFIC +]
            0 CASE erg, GEND masc, NUM sg, PERS 3
              PRED      'Nadya'
              NTYPE     [PROPER name]
   OBJ-GO     SEM-PROP  [SPECIFIC +]
           16 CASE dat, GEND fem, NUM sg, PERS 3
              PRED  'gHar'
   OBJ        NTYPE [GRAIN count]
           32 CASE nom, GEND masc, NUM sg, PERS 3
   CHECK      [_VMORPH [_MTYPE infl]]
   LEX-SEM    [AGENTIVE +, GOAL +]
   TNS-ASP    [ASPECT perf, MOOD indicative]
72 PASSIVE -, PERS 3, STMT-TYPE decl, VTYPE complex-pred

   PRED     'banA<[16:Nadya], [32:gHar]>'
   SUBJ     [16:Nadya]
   OBJ      [32:gHar]
   CHECK    [_NMORPH obl, _RESTRICTED +]
   LEX-SEM  [AGENTIVE +]
47 PASSIVE -, PERS 3, STMT-TYPE decl, VFORM inf
```

The main verb *banA* 'make' has two arguments: a subject and an object. This is indicated in the bottom f-structure in (18). The top f-structure represents the final analysis. Here the SUBJ, PRED, and VTYPE features of the main verb's f-structure have been restricted. The VTYPE feature now states that this is a `complex-pred`. As in the previous example, the PRED feature has the value of a composite argument structure. This results in an overall three-place CP which subcategorizes for a subject via the subcategorization frame of *dE* 'give', a restricted object (OBJ-GO) which is the demoted subject of *banA* 'make', and finally an object which is inherited from the subcategorization frame of *banA* 'make'. Despite the fact that the arguments come from different sources and that the predication is complex (as evidenced by the nesting inside the PRED value in the top f-structure), at the level of f-structure, the arguments function like those of a simplex predicate (cf. Butt 1995).

4.5 Stacked Complex Predicates

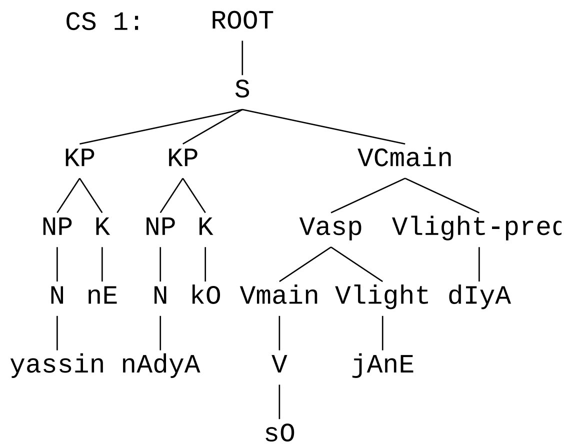
One of the main motivations for doing CP formation in the syntax in Urdu was their ability to stack. In particular, a permissive CP can be formed from an event modulating CP. A stacked CP from an intransitive and from a transitive verb are shown in (19).

(19) a. yassin=nE nAdyA=kO sO jAnE dIyA
Yassin=Erg Nadya=Dat sleep go.Inf give.Perf.M.Sg
'Yassin let Nadya sleep.'

b. yassin=nE nAdyA=kO gHar banA lEnE dIyA
Yassin=Erg Nadya=Dat house.Nom make take.Inf give.Perf.M.Sg
'Yassin let Nadya build a house.'

These stacked CPs follow straightforwardly from the analyses of the event modulating and permissive CPs. The only change that needs to be made is to allow the permissive to either take a main verb as its complement or an event modulating CP; no new restriction equations need to be applied. The c-structure and f-structure for (19a) are shown in (20) and (21).

(20) C-structure tree for permissive of event modulating CP



(21) F-structure AVM for permissive of event modulating CP

"yassin nE nAdyA kO sO jAnE dIyA"

```
  [PRED     'dE<[0:Yassin], 'sO<[16:Nadya]>'>'
  |         [PRED     'Yassin'
  |SUBJ     |NTYPE    [PROPER name]
  |         |SEM-PROP [SPECIFIC +]
  |        0[CASE erg, GEND masc, NUM sg, PERS 3]
  |         [PRED     'Nadya'
  |OBJ-GO   |NTYPE    [PROPER name]
  |         |SEM-PROP [SPECIFIC +]
  |       16[CASE dat, GEND fem, NUM sg, PERS 3]
  |CHECK    [_VMORPH [_MTYPE inf]]
  |LEX-SEM  [AGENTIVE +, GOAL +]
  |TNS-ASP  [ASPECT perf, MOOD indicative, TENSE past]
72[GEND masc, NUM sg, PASSIVE -, PERS 3, STMT-TYPE decl, VTYPE complex-pred]

  [PRED    'sO<[16:Nadya]>'
  |SUBJ    [16:Nadya]
  |CHECK   [_RESTRICTED +]
  |TNS-ASP [AKTIONSART [TELIC +]]
32[GEND masc, LEX-SEM unerg, NUM sg, PASSIVE -, PERS 3, STMT-TYPE decl, VFORM bare]

  [PRED    'sO<[16:Nadya]>'
  |SUBJ    [16:Nadya]
  |CHECK   [_NMORPH obl, _RESTRICTED +]
  |LEX-SEM [AGENTIVE -]
  |TNS-ASP [32-TNS-ASP]
46[GEND masc, NUM sg, PASSIVE -, PERS 3, STMT-TYPE decl, VFORM inf, VTYPE complex-pred]
```

In (21), there are three f-structures. The top f-structure is the final one with the permissive subject Yassin and the thematic goal object Nadya. The bottom two correspond to the main verb and the event modulating complex predicate's f-structures which are restricted.

5 Conclusions

The solution described above in terms of syntactic composition of arguments via the restriction operator allows the manipulation of subcategorization frames outside of the lexicon. This is particularly important as CPs in Urdu/Hindi and other languages are productive and separable in the syntax: they do not present instances of compounding or any other form of lexicalization. A sophisticated manipulation of subcategorization frames outside of the lexicon has not been previously possible, but finds clear applications for CPs crosslinguistically. The Urdu grammar has pioneered the use of restriction in the ParGram project. Since the implementation is recent, the exact details of the CP analysis within the Urdu grammar are subject to change. One issue which remains to be fully resolved is the interaction of different types of light verbs and the modeling of the verbal complex as a whole. Since the verbal complex includes different kinds of auxiliaries (passive, progressive), modals, and light verbs which combine with main verbs, adjectives, and nouns, the collection of interacting phenomena is complex.

A possible immediate application of the restriction operator within the ParGram project would be to the well known problem of *suru* ‘do’ and other CPs found in Japanese. In addition, the restriction operator opens up an innovative treatment of a subtype of the Norwegian passive, as in (22a), and allows for a potentially more satisfactory treatment of the German *lassen* ‘let’ construction, as in (22b), and the French causative *faire* ‘make’, as in (22c).

(22) a. Kniven blir skåret kjøtt med.
the-knife is cut meat with
‘The knife cut the meat.’

b. Der Fahrer hat den Traktor reparieren lassen.
the.Nom driver has the.Acc tractor repair let
‘The driver had the tractor repaired.’

c. Paul fera passer son examen à Jean.
Paul make.Fut take his exam to Jean
‘Paul made Jean take his exam.’

The current ParGram analyses treat these phenomena as instances of basic complement taking verbs, a solution which is not supported by the linguistic evidence and discussions amassed within theoretical linguistics.

In addition, (Wedekind and Ørsnes, 2003) are exploring using restriction as a mechanism for capturing well-formedness conditions on verbal complexes. There have been a number of approaches to this problem in LFG (see (Falk, 2003) for an overview). Analyses have included treating each verb/auxiliary as heading its own XCOMP and having a separate projection in which to state these restrictions. (Wedekind and Ørsnes, 2003) check for the restrictions as the c-structure is built up, but restrict out the features so that the final f-structure is very simple.

Syntactic restriction provides a scalable and efficient solution for the general phenomenon of complex predication in LFG gramamrs. Our implementation makes use of the restriction operator originally proposed by (Kaplan and Wedekind, 1993), but is in line with the original observations (e.g., (Alsina, 1996; Butt, 1995)) as to the nature and function of complex predication.

References

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

Miriam Butt and Wilhelm Geuder. 2001. On the (semi)lexical status of light verbs. In Norbert Corver and Henk van Riemsdijk, editors, *Semi-lexical Categories: On the content of function words and the function of content words*, pages 323–370. Mouton de Gruyter, Berlin.

Miriam Butt and Tracy Holloway King. 2001. Non-nominative subjects in Urdu: A computational analysis. In *Proceedings of the International Symposium on Non-nominative Subjects*, pages 525–548, Tokyo. ILCAA.

Miriam Butt and Tracy Holloway King. 2002. Urdu and the Parallel Grammar project. In *Proceedings of COLING 2002*. Workshop on Asian Language Resources and International Standardization.

Miriam Butt and Tracy Holloway King. 2003. The status of case. In Veneeta Dayal and Anoop Mahajan, editors, *Clause Structure in South Asian Languages*. Kluwer Academic Publishers, Dordrecht. To Appear.

Miriam Butt and Gillian Ramchand. 2003. Building complex events in Hindi/Urdu. In Nomi Ertischik-Shir and Tova Rapoport, editors, *The Syntax of Aspect*. Oxford University Press, Oxford. Submitted.

Miriam Butt, Tracy Holloway King, Marí a-EugeniaNiño, and Frédérique Segond. 1999. *A Grammar Writer's Cookbook*. CSLI Publications.

Miriam Butt, Helge Dyvik, Tracy Holloway King, Hiroshi Masuichi, and Christian Rohrer. 2002. The Parallel Grammar project. In *Proceedings of COLING 2002*. Workshop on Grammar Engineering and Evaluation.

Miriam Butt, Martin Forst, Tracy Holloway King, and Jonas Kuhn. 2003. The feature space in parallel grammar writing. In *ESSLLI 2003 Workshop on Ideas and Strategies for Multilingual Grammar Development*.

Miriam Butt. 1994. Machine translation and complex predicates. In *Proceedings of KONVENS 94*, pages 62–71.

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

Yehuda Falk. 2003. The English auxiliary system revisted. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of LFG03*. CSLI On-line Publications.

Anette Frank. 1996. A note on complex predicate formation: Evidence from auxiliary selection, reflexivization, and past participle agreement in French and Italian. In Miriam Butt and Tracy Holloway King, editors, *The Proceedings of the LFG96 Conference*. CSLI On-line Publications.

Anette Frank. 1999. From parallel grammar development towards machine translation. In *Proceedings of MT Summit VII*, pages 134–142.

Peter Edwin Hook. 1974. *The Compound Verb in Hindi*. The University of Michigan, Center for South and Southeast Asian Studies.

Ron Kaplan and Jürgen Wedekind. 1993. Restriction and correspondence-based translation. In *Proceedings of the Sixth European Conference of the Association for Computational Linguistics*, pages 193–202.

John T. Maxwell, III and Ron Kaplan. 1993. The interface between phrasal and functional constraints. *Computational Lingusitics*, 19:571–589.

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

Jürgen Wedekind and Bjarne Ørsnes. 2003. Restriction and verbal complexes in LFG: A case study for Danish. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of LFG03*. CSLI On-line Publications.