

Modeling complex adpositions in LFG: A case study from Persian and Urdu

Ghulam Raza

Department of Linguistics, Universität Konstanz, Germany

June 18, 2011

Fourth International Conference on Iranian Linguistics (ICIL4): Uppsala

Test Case Languages

Persian:

- ▶ An Indo-Iranian language
- ▶ Mainly spoken in Iran
- ▶ SOV
- ▶ Head-**initial** Adpositions

Urdu

- ▶ Indo-Aryan language
- ▶ Mainly spoken in Pakistan
- ▶ SOV
- ▶ Head-**final** Adpositions

Outline of the Talk

- ▶ What are complex adpositions?
- ▶ Syntactic distribution of nouns in complex adpositions
- ▶ Restricting complex adpositions in LFG paradigm

What are complex adpositions?

- ▶ In many languages, the basic adpositions together with some nouns seem to act as predicators taking other elements for their complements.
- ▶ Such predicators in linguistic literature are usually termed as complex adpositions.

Examples of complex prepositions

- ▶ French: en face de, en dépit de, au milieu de
- ▶ Spanish: al lado de, en casa de
- ▶ Swedish: i början av, med hjälp av, i stället för
- ▶ English: in view of, in spite of, by dint of
- ▶ Polish: bez względu na, w związku z, z uwagi na
- ▶ German: an Hand von, mit Hilfe von, in Bezug auf

(adapted from Trawinski 2003)

Some Examples in Persian/Urdu

(1) a.

az ru e tarix
from face Ez history
'based of history'

b.

be tor e sadr
to way Ez president
'as a president'

- ▶ The nouns like *ru* 'face' in complex adpositions are in fact relational nouns taking lexically their own complements.

Simple vs. Complex Adposition

(2) a.

u be xane e ali raft

he to home Ez Ali go-Perf

'He went to Ali's home.' (Simple adposition)

b.

u be illat e gürüsngi mürd

he to cause Ez hunger die-Perf

'He died of hunger.' (Complex Adposition)

Nouns in complex adpositions do not show their normal syntactic distribution

- ▶ The nouns in complex adpositions show different syntactic distribution as compared with their normal syntactic distribution.
- ▶ (Chime, 2006) has discussed such differences for Persian that actually apply to Urdu and many other languages too.
- ▶ These differences support the analysis that the two lexical categories *be* and *illat* project to a complex adposition *be illat* which heads the entire phrase.
- ▶ So, the assumption can be made that the lexically selected complement of the relational noun in such phrases is syntactically realized as the object of complex adposition.

Determiners/Quantifiers

(3) a.

bar (in/yık/har) bam e xane
 on (this/one/every) roof Ez house
 'on (this/one/every) roof of the house'

b.

bar (*in/*yık/*har) **asas** e motale'at e danıřmandan
 on (*this/*one/*every) base Ez research.PI Ez scientist.PI
 'based of scientists' researches'

Question Words/Exclamatories

(4) a.

bar (che) bam e xane
 on (what) roof Ez house
 'on (what) roof of the house'

b.

bar (*che) **asas** e motale'at e danıřmandan
 on (*what) base Ez research.PI Ez scientist.PI
 'based of scientists' researches'

Superlatives

(5) a.

bar (ziba-tarin) bam e xane
 on (the most beautiful) roof Ez house
 'on (the most beautiful) roof of the house'

b.

bar (*jadid-tarin) **asas** e motale'at e danıřmandan
 on (*the newest) base Ez research.PI Ez scientist.PI
 'based of scientists' researches'

Adjectives/Apositives

(6) a.

bar bam e (ziba e) xane
 on roof Ez (beautiful Ez) house
 'on the (beautiful) roof of the house'

b.

be illat e (*puʃ e) barandgi
 to reason Ez (*vain Ez) raining
 'because of the (*vain) reason of raining'

Plurality

(7) a.

az kitab-(ha) (y)e ali istifade kardam
 from book-(Pl) Ez Ali use do-Perf
 'I used Ali's books.'

b.

az janıb-(*ha) (*y)e dılat va mardum
 from side-(Pl) Ez government and nation
 'from the government and the nation'

Coordination/Reduplication

(8) a.

ali **ba** dost va dufman mudara mikunad
Ali with friend and enemy bearing do.Pres
'Ali bears everyone.'

b.

be illat (*va dalil) e barandgi
to reason (*and cause) Ez raining
'because of raining'

Dual syntax of PPs in Urdu

(9) a.

be qol e nida
to say Ez Nida
'according to Nida'

b1.

nida=ke be qol
Nida=Gen to say
'according to Nida'

b2.

be qol nida=ke
to say Nida=Gen
'according to Nida'

Restricting complex adpositions in LFG I

- ▶ The phenomenon of complex adposition formation is somehow similar to the verbal complex formation.
- ▶ Alsina (1996) and Butt (1995) have argued that complex predicate formation should take place in syntax.
- ▶ Frank (1996), however, have the opposite position to place complex predicate formation within the lexical component. Although in her approach subcategorization requirements and restrictions on combinations are specified within the lexical entries, still the separate pieces are combined in the syntax.

Restricting complex adpositions in LFG II

- ▶ Butt and King (2003) implemented complex predicates of Urdu via RESTRICTION (Kaplan and Wedekind, 1993)
- ▶ These proposals have been forwarded for the verbal complex formation in different languages. We extend the proposal for licensing the complex adpositions in LFG formalism.
- ▶ Trawinski (2003) proposed an HPSG-based linguistically motivated, formal treatment of CPs, applicable for computational platforms intended for developing typed feature structure grammars

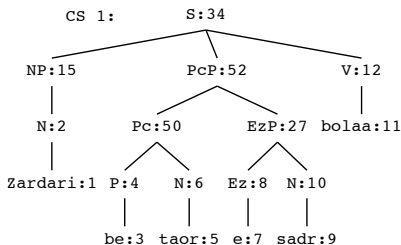
Restricting complex adpositions in LFG III

- ▶ In our case too, the subcategorization requirement and the features recognizing some specific combinations of adpositions and nouns are included in lexical entries and are checked at the syntax level.
- ▶ The syntactic composition of two predicates is done in LFG using RESTRICTION which allows f-structures and predicates to be manipulated in detail.
- ▶ In grammar rules the specifier/complement of the noun part is suppressed via RESTRICTION and is assigned as the object of the complex adpositions.

c-structure of complex adpositions

(10)

zardari **be** **taor** e sadr bol-a
 zardari to way Ez president speak-Perf.M.3Sg
 'Zardari spoke as a president.' (Urdu)



f-structure of complex adpositions

(11)

zardari **be** **tor** e sadr bol-a
zardari to way Ez president speak-Perf.M.3Sg
'Zardari spoke as a president.' (Urdu)

"Zardari be taor e sadr bolaa"

```

1[PRED 'bol<[1:Zardari]>']
SUBJ 2[PRED 'Zardari'
15[GEND masc, NUM sg, PERS 3]
ADJUNCT {
4[
11[PRED 'be<'taor<[7:sadr]>'>'
12[
34[
9[PRED 'sadr'
10[Case Gen-e, GEND masc, NUM sg, PERS 3]
7[
8[
50[
27[
52[GEND masc, NUM sg, PERS 3, after_ad +
TENSE past
[PRED 'taor<[7:sadr]>'
5 OBL [7:sadr]
6 GEND masc, NUM sg, PERS 3, after ad +

```

LFG grammar rules for Pc

Lexical Entry:

be $P * (\uparrow \text{ PRED}) = \text{'be} < \% \text{ ARG1} > \text{'}$;

$P * (\uparrow \text{ PRED}) = \text{'be} < (\uparrow \text{ OBJ}) > \text{'}$.

XLE Rule:

$P_c \rightarrow \{P: \uparrow = \downarrow;$

$N: \downarrow \backslash \text{PRED} \backslash \text{OBL} = \uparrow \backslash \text{PRED} \backslash \text{OBJ}$

$(\uparrow \text{ PRED ARG1}) = (\downarrow \text{ PRED})$

$(\uparrow \text{ OBJ}) = (\downarrow \text{ OBL})$

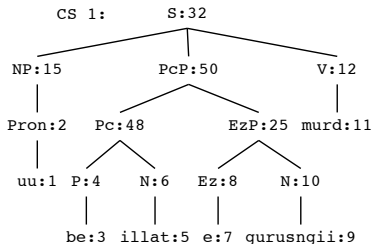
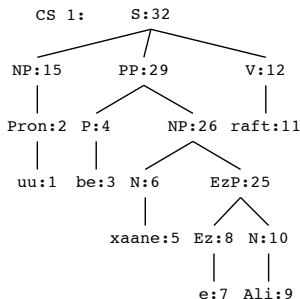
$(\downarrow \text{ after_ad}) = c +$

$|\dots$

$|\dots\}$

Simple vs. Complex Adposition (c-structure)

- (12) a. u **be** xane e ali raft. 'He went to Ali's house.'
b. u **be illat** e gurusngi murd. 'He died of hunger.'



Simple vs. Complex Adposition (f-structure)

- (13) a. u **be** xane e ali raft. 'He went to Ali's house.'
b. u **be ıllat** e gürüsngi mürd. 'He died of hunger.'

"uu be xaane e Ali raft"

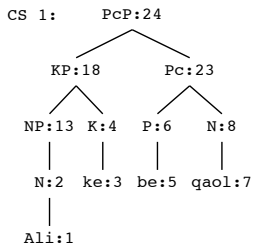
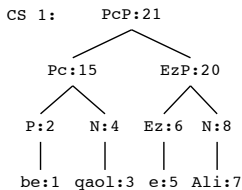
11 PRED 'raftan<[1:uu]>'
12 1[PRED 'uu'
13 SUBJ 2[NUM sg, PERS 3
14 15[
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839

"uu be illat e gurusngii murd"

1[PRED 'murdan<[1:uu]>'
 2 1[PRED 'uu'
 3 2[SUBJ 2[
 4 15[
 5 15[
 6 15[
 7 15[
 8 15[
 9 15[
 10 15[
 11 15[
 12 15[
 13 15[
 14 15[
 15 15[
 16 15[
 17 15[
 18 15[
 19 15[
 20 15[
 21 15[
 22 15[
 23 15[
 24 15[
 25 15[
 26 15[
 27 15[
 28 15[
 29 15[
 30 15[
 31 15[
 32 15[
 33 15[
 34 15[
 35 15[
 36 15[
 37 15[
 38 15[
 39 15[
 40 15[
 41 15[
 42 15[
 43 15[
 44 15[
 45 15[
 46 15[
 47 15[
 48 15[
 49 15[
 50 15[
 51 15[
 52 15[
 53 15[
 54 15[
 55 15[
 56 15[
 57 15[
 58 15[
 59 15[
 60 15[
 61 15[
 62 15[
 63 15[
 64 15[
 65 15[
 66 15[
 67 15[
 68 15[
 69 15[
 70 15[
 71 15[
 72 15[
 73 15[
 74 15[
 75 15[
 76 15[
 77 15[
 78 15[
 79 15[
 80 15[
 81 15[
 82 15[
 83 15[
 84 15[
 85 15[
 86 15[
 87 15[
 88 15[
 89 15[
 90 15[
 91 15[
 92 15[
 93 15[
 94 15[
 95 15[
 96 15[
 97 15[
 98 15[
 99 15[
 100 15[
 101 15[
 102 15[
 103 15[
 104 15[
 105 15[
 106 15[
 107 15[
 108 15[
 109 15[
 110 15[
 111 15[
 112 15[
 113 15[
 114 15[
 115 15[
 116 15[
 117 15[
 118 15[
 119 15[
 120 15[
 121 15[
 122 15[
 123 15[
 124 15[
 125 15[
 126 15[
 127 15[
 128 15[
 129 15[
 130 15[
 131 15[
 132 15[
 133 15[
 134 15[
 135 15[
 136 15[
 137 15[
 138 15[
 139 15[
 140 15[
 141 15[
 142 15[
 143 15[
 144 15[
 145 15[
 146 15[
 147 15[
 148 15[
 149 15[
 150 15[
 151 15[
 152 15[
 153 15[
 154 15[
 155 15[
 156 15[
 157 15[
 158 15[
 159 15[
 160 15[
 161 15[
 162 15[
 163 15[
 164 15[
 165 15[
 166 15[
 167 15[
 168 15[
 169 15[
 170 15[
 171 15[
 172 15[
 173 15[
 174 15[
 175 15[
 176 15[
 177 15[
 178 15[
 179 15[
 180 15[
 181 15[
 182 15[
 183 15[
 184 15[
 185 15[
 186 15[
 187 15[
 188 15[
 189 15[
 190 15[
 191 15[
 192 15[
 193 15[
 194 15[
 195 15[
 196 15[
 197 15[
 198 15[
 199 15[
 200 15[
 201 15[
 202 15[
 203 15[
 204 15[
 205 15[
 206 15[
 207 15[
 208 15[
 209 15[
 210 15[
 211 15[
 212 15[
 213 15[
 214 15[
 215 15[
 216 15[
 217 15[
 218 15[
 219 15[
 220 15[
 221 15[
 222 15[
 223 15[
 224 15[
 225 15[
 226 15[
 227 15[
 228 15[
 229 15[
 230 15[
 231 15[
 232 15[
 233 15[
 234 15[
 235 15[
 236 15[
 237 15[
 238 15[
 239 15[
 240 15[
 241 15[
 242 15[
 243 15[
 244 15[
 245 15[
 246 15[
 247 15[
 248 15[
 249 15[
 250 15[
 251 15[
 252 15[
 253 15[
 254 15[
 255 15[
 256 15[
 257 15[
 258 15[
 259 15[
 260 15[
 261 15[
 262 15[
 263 15[
 264 15[
 265 15[
 266 15[
 267 15[
 268 15[
 269 15[
 270 15[
 271 15[
 272 15[
 273 15[
 274 15[
 275 15[
 276 15[
 277 15[
 278 15[
 279 15[
 280 15[
 281 15[
 282 15[
 283 15[
 284 15[
 285 15[
 286 15[
 287 15[
 288 15[
 289 15[
 290 15[
 291 15[
 292 15[
 293 15[
 294 15[
 295 15[
 296 15[
 297 15[
 298 15[
 299 15[
 300 15[
 301 15[
 302 15[
 303 15[
 304 15[
 305 15[
 306 15[
 307 15[
 308 15[
 309 15[
 310 15[
 311 15[
 312 15[
 313 15[
 314 15[
 315 15[
 316 15[
 317 15[
 318 15[
 319 15[
 320 15[
 321 15[
 322 15[

Dual syntax of PcPs in Urdu (c-structure)

- (14) a. **be qpl** e ali. 'according to Ali.'
b. ali ke **be qpl**. 'according to Ali.'



Dual syntax of PcPs in Urdu (f-structure)

- (15) a. **be qol** e ali. 'according to Ali.'
b. ali ke **be qol**. 'according to Ali.'

"be qaol e Ali"

	[PRED	'be<'qol<[5:Ali]>'>']
		7	PRED 'Ali'	
		8	Case Gen-e, GEND masc, NUM sg, PERS 3	
	OBJ	5		
2		6		
15	20			
21		GEND masc, NUM sg, PERS 3, after_ad +		

	[PRED	'qol<[5:Ali]>']
3	OBL	[5:Ali]		
4		GEND masc, NUM sg, PERS 3, after_ad +		

"Ali ke be qaol"

	[PRED	'be<'qol<[1:Ali]>'>']
		3	PRED 'Ali'	
		4	Case Gen-k, GEND masc, NUM sg, PERS 3	
	OBJ	1		
		2		
6		13		
23		18		
24		GEND masc, NUM sg, PERS 3, after_ad +		

	[PRED	'qol<[1:Ali]>']
7	OBL	[1:Ali]		
8		GEND masc, NUM sg, PERS 3, after_ad +		

(N + P) complex adpositions in Persian

- ▶ bana bar
- ▶ nisbat be
- ▶ raje be

(*zer* + N) complex adpositions in Urdu

- ▶ police ki *zer e harasat* 'arrested by the police'
- ▶ adliyah ki *zer e nigrani* 'under the vigilance of courts'
- ▶ Zardari ki *zer e sadarat* 'headed by Zardari'
- ▶ Ali ke *zer e istemal* 'used by Ali'
- ▶ faoj ke *zer e qabzah* 'captured by the Army'
- ▶ doctor ke *zer e ilaj* 'cured by the doctor'

Lexical features and constraints

- ▶ Which adposition can form complex adposition with some noun?
- ▶ Which noun can actually be part of some complex adposition?
- ▶ Which specific combinatorics preferences are?
- ▶ The order of two parts of a complex adposition
- ▶ The constraints on position and case of the object of complex adpositions

Conclusion

- ▶ Many complex adpositions developed in Persian are also used in Urdu
- ▶ Complex adpositions in Urdu formed with Persian simple adposition *zer* are found in Urdu but not in Persian
- ▶ The phenomenon of complex adposition formation resembles with the verbal complex formation
- ▶ The available apparatus in XLE environment makes it possible to model the linguistically motivated analysis of complex adpositions in LFG

Thanks!

References

- Alsina, A. 1996. *The Role of Argument Structure in Grammar*. CSLI Publications.
- Butt, M. 1995. *The Structure of Complex Predicates in Urdu*. CSLI Publications.
- Chime, Z. A. 2006. *An Account for Compound Prepositions in Farsi*. Proceedings of COLING/AICL, Sydney.
- Frank, A. 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.
- Trawinski, B. 2003. Licensing Complex Prepositions via Lexical Constraints. In Francis Bond, Anna Korhonen, Diana McCarthy, and Aline Villavicencio (Eds.), *Proceedings of the ACL 2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, Sapporo Convention Center, Sapporo, Japan, pp. 97-104. Association for Computational Linguistics.