ARGUMENT SCRAMBLING WITHIN URDU NPS

Ghulam Raza   and   Tafseer Ahmed

Universität Konstanz   Universität Konstanz

Abstract

Clause-level discontinuity of NPs in different languages has been reported in the literature under different headings like discontinuity, extraposition, extraction, free topic, quantifier float and so on. However, discontinuity within the constituent-level, has only recently been noted (Raza, 2011) and is reported on in this paper. Urdu is a language in which both types of discontinuity, clause-level and constituent-level, have been found. In constituent-level discontinuous NPs, discontinuity occurs inside NPs at one structural position of a clause. In such phrases, the arguments of heads are non-contiguous to their respective heads inside NPs. The heads, however, cannot precede their arguments. This discontinuity and the constraints on the order of elements in NPs in Urdu pose a modeling problem. In contrast to what is usually assumed within ParGram (Butt et al., 1999; Dipper, 2003), a flat c-structure for Urdu NPs is therefore proposed and is modeled in the LFG framework.

1   Introduction

In simple terms, a constituent in which the individual words that make it up appear separately and are interleaved with other elements is called a discontinuous constituent. Languages that exhibit several of the following characteristics: free word order, discontinuous constituents, split-ergative case marking and null anaphora, are often referred to as non-configurational languages (Hale, 1980, 1983, 1992; Simpson, 1983, 1991). Warlpiri is considered a typical example of a non-configurational language which shows almost all the properties that have been associated with non-configurational languages. A sample sentence of Warlpiri is given in (1) to show the phenomenon of discontinuous NPs across the clause.

(1) **wita-jarra-rlu** ka-pala   wajili-pi-nyi *yalumpu*   (WarlPiri)
     small-Du-Erg   Pres-3DuSubj chase-NPast that.Abs

   kurdu-jarra-rlu *maliki.*
   child-Du-Erg   dog.Abs

   'The two small children are chasing that dog.'

                              (Austin and Bresnan 1996:217)

In (1) the two NPs 'two small children' and 'that dog' are not continuous. Some other elements are interleaved between the head nouns 'children' and 'dog' and their modifiers 'small' and 'that'. Other permutations of words in the sentence (provided the auxiliary *ka-pala* is always in the second position) can also be uttered spontaneously by a native speaker and the truth-conditional meaning of the sentence does not change. Austin and Bresnan (1996) have worked out a detailed analysis of non-configurationality in Australian aboriginal languages in the LFG framework.[1] Discontinuous noun phrases have also been investigated in detail for many other languages like German (Müller, 2004; van Riemsdijk, 1989;

---

[1]Legate (2002), however, has argued for a configurational analysis of Warlpiri.

Kuhn, 1998; Fanselow and Ćavar, 2002; Roehrs, 2006), Russian (Kazenin, 2005; Gouskova, 2001; Sekerina, 1997), Greek (Nthelitoes, 2004; Agbayani and Golston, 2005), Dutch (van Hoof, 1997), Ukrainian (Féry et al., 2007), Serbo-Croatian-Bosnian (Ćavar, 1999; Boškovič, 2005), etc.

Urdu/Hindi is a free-word-order language. The dependents of nouns and arguments of verbal predicates are usually marked for case by clitics. As part of providing a complete analysis and model of Urdu noun phrases in the context of a broad-coverage ParGram grammar of Urdu within the LFG framework (Butt et al., 1999), parallel to the syntax of other languages like English, German, French etc. (Butt et al., 1999), a corpus of Urdu newspapers, *Roznama Jang* and *Roznama Nawaiwaqt*, of Pakistan has been analyzed to establish patterns within Urdu noun phrases. Some of our observations about possible permutations of modifiers, specifiers and arguments in Urdu NPs point to an interesting NP organization of a kind not previously reported in the literature. For one, Urdu NPs can have two genitive marked arguments as specifier/complement of the head noun (Raza, 2010), just as in German and English. However the Urdu genitives are not always tied to a particular phrase structure position, unlike in German and English. And Urdu does not have alternate constructions for possessors in NPs as found in English and Hungarian (Laczkó, 2000), for example, although an *ezafe* construction is sometimes found with Persian loan words (Bögel and Butt, to appear). Furthermore, evidence for a general non-hierarchical nature of the Urdu NP comes from the fact that the arguments in NPs can be non-contiguous to their respective heads. In this paper we present and analyze the latter phenomenon and model it in LFG. In contrast to what is usually assumed within ParGram (Butt et al., 1999; Dipper, 2003), we propose a flat c-structure for Urdu NPs. Although Urdu NPs have been described at length in grammar books (e.g., Schmidt 1999; Platts 1967), the phenomenon described in this paper to our knowledge has not been noted before.

The paper is organized as follows. Section 2 provides an overview of the basic terminology of discontinuity. In section 3 we describe the argument-taking adjectives in Urdu. The argument-taking adjectives in Urdu interact with nouns to generate discontinuous constituents in noun phrases. The data of discontinuous noun phrases in Urdu is described in section 4. Theoretical implications of the data are described in section 5. In section 6, the constituent-level discontinuity in Urdu is implemented in the LFG framework and section 7 concludes the paper.

2 State of the Art

Different phenomena of discontinuous noun phrases have been discussed in the literature (see for example (Fanselow and Féry, 2006)) and are briefly described in the following subsections. Many examples quoted in these subsections are taken from Fanselow and Féry (2006).

2.1 Simple and inverted discontinuous noun phrases

If the order of elements in the discontinuous noun phrase is the same as the canonical order in the corresponding continuous noun phrase, then the discontinuous noun phrase is called a simple discontinuous noun phrase, else it is called an inverted discontinuous noun phrase.

Considering the prosodic properties, discontinuous noun phrases are divided into cohesive and non-cohesive discontinuous noun phrases. When the whole discontinuous noun phrase is integrated into a single intonational phrase, then it is a cohesive discontinuous noun phrase and if its two parts are separated into two intonational phrases then it is a non-cohesive discontinuous noun phrase. Simple discontinuous noun phrases are usually cohesive and inverted discontinuous noun phrases are usually non-cohesive. In Ukrainian (Féry et al., 2007) both types of discontinuous noun phrases have been reported.

(2) a. Marija maje **bahato krisel**. (Ukrainian)
Mary has.got many chairs.Gen.Pl
'Mary has got many chairs.'

b. **bahato** maje Marija **krisel.**

c. **krisel** Marija maje **bahato**.

The example sentence (2) from Ukrainian depicts the canonical order of a continuous noun phrase in (2a) and a simple discontinuous noun phrase in (2b) and an inverted discontinuous noun phrase in (2c).

2.2 Extraction from DP

Extraction from DP involves the dislocation of an argument or adjunct of the head noun to the left in the DP. For example in (3), *über Logik* 'about logic' is thematically dependent on the lexical noun *Bücher* 'books'. In (3b), although the DP is discontinuous, the adjunct of the noun is still adjacent to it. In (3c), however, the adjunct *über Logik* is taken out of the DP to the left and hence is an example of extraction.

(3) a. Er hat **viele Bücher über Logik** gekauft. (German)
He has many books on logic bought
'He has bought many books about logic.'

b. **Bücher über Logik** hat er **viele** gekauft.

c. **Über Logik** hat er **viele Bücher** gekauft.

This distinction between extraction and other discontinuous phrases was made by generative syntacticians (Haider, 1985). Extraction as in (3c) is generally explained by movement. Only a maximal projection is posited to move to a pre-auxiliary position. As *über Logik* is the maximal projection of a preposition, it

can be moved to the pre-auxiliary position. *Bücher über Logik*, however, is considered a submaximal projection of a noun. The maximal projection of a noun is assumed to have the specifier position filled by the determiner. So (3b) and (3c) are explained by different mechanisms and a distinction is made between (3b) and (3c). Müller (2004) has described various possible analyses for (dis)continuous constituents in German in HPSG with different assumptions and explanations.

2.3   Quantifier Float

Quantifier Float involves dislocation of the quantified expression away from the noun. This phenomenon has also been explained in terms of movement dependencies in that the DP can move to Spec,TP and the quantifier could be left *in situ* (Déprez, 2003). A further analysis for *all* is made in terms of adverbial quantification as it shares the distributional properties of adverbs like *ever*.

(4) a. **They all** have bought a car.

b. **They** have **all** bought a car.

In (4a) the base generated 'They all' has moved to Spec,TP and in (4b) the quantified part 'all' has been left *in situ* and only *They* has moved to Spec,TP.

McCloskey (2000) has observed another type of quantifier float in the context of *wh*-movement. The quantifier in (5b) is not bound with the subject, rather it is construed with the *wh*- question word referring to the object.

(5) a. **What all** did you get *t* for Christmas? (Irish English)

b. **What** did you get **all** for Christmas?

2.4   Free Topic structure

In a Free Topic structure, two semantically related elements forming a unique theme become discontinuous in the clause. Usually one element that is more abstract is made the topic and the other more specific element is placed in the canonical position.

(6) **Say-nun** ku-ka   **nightingale-man** a-n-ta.   (Korean)
bird-Top he-Nom nightingale-only   know-Pres-Dec
'As for birds, he only knows nightingales.'

In (6) *bird* and *nightingale* are both semantically related and form a unified theme in the clause, although they are separate from each other in the clause.

2.5 Extraposition

Extraposition is a phenomenon in which the dependent element of a noun is moved to the right in contrast with extraction where the dependent element is moved to the left.

(7) a. **A man** came in **who had a beard**.

b. **A book** came out **about logic**.

The relative clause in (7a) that describes the noun in the main clause is postposed and in (7b) the PP adjunct of the noun is postposed.

Although extraction, quantifier float, free topic and extraposition are all in some sense discontinuous noun phrases, in the generative framework these are generally distinguished from the absolute discontinuous noun phrases that only involve the separation of the head noun from its determiner, article or an adjective modifying it. Theoretically the phenomenon of a discontinuous noun phrase is licensed only if at least one of the heads involved appears in an A-bar position (Fanselow and Féry, 2006).

2.6 Clause-level discontinuity in Urdu

Almost all types of discontinuous NPs at clause level mentioned above have also been observed in Urdu. Although clause-level discontinuous NPs are not the main topic here, however, for contrast, evidence for such phrases is provided by the following examples.

(8) a. ندا نے منطق پر ایک کتاب خریدی ہے.
nɪda=ne **mantɪq=pɑr ek kɪtab** xɑrid-i hɛ.
Nida=Erg logic=Loc.on one book.F.3Sg buy-Perf be.Pres
'Nida has purchased a book on logic.'

b. منطق پر ندا نے ایک کتاب خریدی ہے.
mantɪq=pɑr nɪda=ne **ek kɪtab** xɑrid-i hɛ.
logic=Loc.on Nida=Erg one book.F.3Sg buy-Perf be.Pres

(Extraction from DP)

(9) a. علی نے بہت آم کھائے.

    ɑli=ne   **bɑhʊt am**   kʰa-e
    Ali.M.3Sg=Erg many   mango.M eat-Perf.M.3Pl
    'Ali ate many mangoes.'

b. آم علی نے بہت کھائے.

    am   ɑli=ne   **bɑhʊt** kʰa-e
    mango.M Ali.M.3Sg=Erg many   eat-Perf.M.3Pl

                                  (Quantifier Float)

(10) a. علی کو آم پسند ہے.

    ɑli=ko   **am**   pɑsãd hɛ
    Ali.M.3Sg=Dat mango.M liked   be.Pres
    'Ali likes mango.'

b. پھل، علی کو آم پسند ہے.

    pʰɑl   ɑli=ko   **am**   pɑsãd hɛ
    fruit.3Sg Ali.M.3Sg=Dat mango.M liked   be.Pres
    'With respect to fruits, Ali likes mangoes'   (Free Topic)

(11) a. ایک شغص، جس کی داڑھی تھی، آیا.
    ek ʃɑxs   **jɪs=ki**   **daṛʰi**   **tʰi**   a-ya
    one person.M.3Sg Rel.3Sg=Gen.F beard.F.3Sg be.Past come-Perf.M.3Sg
    'One person, who had a beard, came.'

b. ایک شغص آیا جس کی داڑھی تھی.
    ek ʃɑxs   a-ya   **jɪs=ki**   **daṛʰi**   **tʰi**
    one person.M.3Sg come-Perf.M.3Sg Rel.3Sg=Gen.F beard.F.3Sg be.Past
    'One person came who had a beard.'   (Extraposition)

Extraposition with relative clauses and correlatives has been discussed in detail by Dayal (1994) and a relevant discussion is also made by Dwivedi (1994). Note that, the sentence in (9b) is an instance of an inverted discontinuous noun phrase in Urdu. Before moving to the within-constituent discontinuity in Urdu, argument taking adjectives are briefly described in section 3 to provide the necessary background.

3 Argument-taking adjectives

In Urdu, participial adjectives and some other adjectives originally derived from verbal stems of other languages take arguments. Some examples of the second type of argument-taking adjectives are listed in Table 1. The nouns in parenthesis are modified by the argument taking adjective.

Nr.	Type of Argument	Example of Adjective Phrase
(i)	Dative Marked	sɑdɑr=ko **hasıl** (ıxtıyarat) president=Dat attained (powers) '(The powers) attained by the president'
(ii)	Instrumental Marked	ɑdlıyɑh=se **xaıf** (hʊkmɑran) courts=Inst afraid (ruleres) '(The rulers) afraid of courts'
(iii)	Locative (in) Marked	bʊxar=mẽ **mʊbtɑla** (ʃɑxs) fever=Loc.in suffered (man) '(The man) suffered with fever'
(iv)	Locative (on) Marked	tɑqarir=pɑr **mʊʃtɑmıl** (kıtab) speech.Pl=Loc.on comprised (book) '(The book) comprised of speeches'
(v)	Adpositional	sıhɑt=ke lıye **mʊzır** (xurak) health=Gen for harmful (food) '(The food) harmful for health'
(vi)	Genitive Marked	sɑdɑr=ke **hami** (ɑfrad) president=Gen.Pl supporting (people) '(People) supporting to the president'

Table 1: Argument-taking adjectives

Adjectives listed in Table 1 are derived from Arabic verbal stems and can sometimes be replaced with native participial adjectives as shown in (12).

(12) a. صدر کو حاصل اختیارات
       sɑdɑr=ko   hasıl ıxtiyarat
       president=Dat attained power.M.3Pl
       'powers attained by the president'

b. اختیارات ہوئے ملے کو صدر

    sɑdɑr=ko   mɪl-e   hu-e   ɪxtiyarat

    president=Dat attain-Perf.M.3Pl be-Perf.M.3Pl power.M.3Pl

    'powers attained by the president'

The adjective *hasıl* in (12a) and the participial form *mıl-e hu-e* in (12b) are semantically equivalent with the meaning of 'attained'.[2] The examples of noun phrases shown in Table 1 are not discontinuous as the nouns here are without arguments. Examples of discontinuous constituents will come in the next section.

Some adjectives like *hami* 'supporting/supporter' in (vi) of Table 1 taking a genitive marked argument are also used as nouns in Urdu. In addition to these adjectives, participial forms of verbs are also used as adjectives. The subject of transitive participial form is genitive marked when the participial form is used as an adjective (see (Raza, 2011) for detail). Apart from the subject of participial adjectives, the theme of some degree adjectives in Urdu is also marked by genitive case, as is shown in (13) for the two adjectives *bʊra* 'bad' and *accʰa* 'good'.

(13) a. شغص برا کا دل

      dɪl=ka   bʊra   ʃɑxs

      heart.M.3Sg=Gen.M.3Sg bad.M.3Sg person.M.3Sg

      'a person of bad heart'

      b. لڑکی اچھی کی طبیعت

      tɑbi'ɑt=ki   accʰi   lɑrki

      nature.F.3Sg=Gen.F.3Sg good.F.3Sg girl.F.3Sg

      'a girl of good nature'

Ikeya (1995) showed that degree adjectives in English are one-place predicates and the contextual/semantic dimensions are in fact modifiers of these predicates. He reported three contextual dimensions of such predicates: Thematic Dimension (TD), Comparative Dimension (CD) and Degree Dimension (DD). In his example *He is very good at basketball for a short Japanese* all these dimensions are expressed: *at basketball* is TD, *for a short Japanese* is CD and *very* is DD. The first dimension TD was first reported by Bartsch (1986/87). These dimensions of degree adjectives are usually encoded by different case phrases or adpositional phrases. The genitive marked elements in (13) in fact are modifiers of adjectives encoding the thematic dimension of adjectives.

---

[2]The equivalent construct in Persian for this meaning is *hasıl ʃʊdah* which is the participial form of *hasıl ʃʊdan*. In Urdu, the adjective *hasıl* can also be considered as the reduced participle form of complex predicate *hasıl ho.*

Adjectives in Urdu can also allow for clausal complements. There is one class of adjectives that can appear in the copula constructions illustrated in (14). The morphemes *yıh* or *yıh bat* are equivalent to expletives in Urdu and can sometimes be dropped.

(i) [Nominalized property] Adj Cop
(ii) [(*yıh/yıh bat*) 'it'] Adj Cop CP

(14) a. علی کا انعام جیتنا ممکن ہے۔

[ɑli=ka   ın'am   jit-na]   **mʊmkın** hɛ
[Ali=Gen.M.3Sg prize.M.3Sg win-Inf.M] possible be.Pres.3Sg
'Ali's winning of the prize is possible.'

b. (یہ (بات)) ممکن ہے کہ علی انعام جیتے۔

(yıh (bat))   **mʊmkın** hɛ   [kıh ɑli ın'am   jit-e]
(this (thing.F.3Sg)) possible be.Pres [that Ali prize.3Sg win-Subjn]
'It is possible that Ali will win the prize.'

In addition to *mʊmkın* 'possible', some other adjectives that fall in this class are *yaqini* 'sure', *zaruri* 'important', *sahih* 'true', etc.

Consider another class of adjectives which can take three alternate frames with the typical example *hɛran* 'surprised' illustrated in (15).

(i) NP-*par*
(ii) that-clause
(iii) NP-*par* that-clause

(15) a. ندا علی کے انعام جیتنے پر حیران ہے۔

nıda [ɑli=ke   ın'am jit-ne]=pɑr   **hɛran** hɛ
Nida.F [Ali=Gen.Obl prize.M win-Inf.Obl]=Loc.on surprised be.Pres.3Sg
'Nida is surprised at Ali's winning of the prize.'

b. ندا حیران ہے کہ علی نے انعام جیتا۔

nıda **hɛran** hɛ   [kıh ɑli=ne ın'am   jit-a]
Nida.F surprised be.Pres.Sg [that Ali=Erg prize.M win-Perf.M]
'Nida is surprised that Ali won the prize.'

c. ندا اس بات پر حیران ہے کہ علی نے انعام جیتا.

nɪda ɪs bat=pɑr **hɛran** hɛ [kɪh ɑli=ne ɪn'am jit-a]
Nida this thing=Loc.on surprised be.Pres [that Ali=Erg prize.M win-Perf.M]
'Nida is surprised that Ali won the prize.'

d. ندا علی پر حیران ہے کہ اس نے انعام جیتا.

nɪda ɑli=pɑr **hɛran** hɛ [kɪh ʊs=ne ɪn'am jit-a]
Nida.F Ali.M=Loc.on surprised be.Pres [that 3Sg=Erg prize.M win-Perf.M]
'Nida is surprised that Ali won the prize.'

It can be said that the canonical argument of the adjective *hɛran* 'surprised' is *par* marked NP. However, if the *par* marked element is some nominalization then the adjective can alternatively take *that*-clause arguments (15b–d). The adjectives *pareʃan* 'sad' and *xʊʃ* 'happy' show similar syntactic behavior.

4 NP-internal discontinuity

In NPs, both nouns and their arguments/modifiers can have their own arguments. The discontinuous constituents in NPs occur in Urdu when some argument-taking adjectives modify some argument taking noun or if the argument of the head noun licenses its own argument in the noun phrase.

Argument-taking adjectives are placed further away from the head noun in comparison with argument-less adjectives. Both the argument of the head noun and the argument of its modifier/argument can co-occur at the start of noun phrases giving rise to discontinuous constituents within a noun phrase. Although Urdu noun phrases have been described in grammar books (see Schmidt 1999; Platts 1967, etc.), the phenomenon of discontinuity within the bounds of noun phrases has not been noticed and discussed before. Consider first rather simple examples of noun phrases in (16)–(17).

(16) a. مقدّمات سے استثنٰی

mʊqaddɑmat=se ɪstɪsna
court-case.M.3Pl=Abl immunity.M.3Sg
'immunity from court-cases'

b. سلامتی پر بریفنگ

salamti=par   barifĩg
security.F.3Sg=Loc.on briefing.F.3Sg
'briefing on security'

c. آرمی چیف سے مطالبہ

armi-cif=se   muʈalbah
army-chief.M.3Sg=Abl demand.M.3Sg
'demand to the army-chief'

(17) a. مقدّمات سے آئینی استثنی

muʊqaddamat=se   aini   **ıstısna**
court-case.M.3Pl=Abl constitutional immunity.M.3Sg
'constitutional immunity from court-cases'

b. سلامتی پر تفصیلی بریفنگ

salamti=par   tafsili   **barifĩg**
securit.F.3Sgy=Loc detailed briefing.F.3Sg
'detailed briefing on security'

c. آرمی چیف سے قانونی مطالبہ

armi-cif=se   qanuni **muʈalbah**
army-chief.M.3Sg=Abl legal   demand.M.3Sg
'legal demand to the army-chief'

Example (16) contains just head nouns with a single argument. In (17), however, the head nouns are modified by argument-less adjectives. We see that the argument of the head noun in Urdu is separated from the noun when an adjective modifies the head noun. In English, on the other hand, the adjective modifying the noun is placed prenominally and the complement of the noun comes postnominally and so both remain contiguous to the head noun. A complex example of noun phrases in Urdu with different order of elements is given in (18).

(18) a. صدر کو حاصل مقدّمات سے آئینی استثنٰی

    sadar=ko$_1$ hasıl$_1$ mʊqaddamat=se$_2$ aini   ıstısna$_2$

    president=Dat possessed court-cases=Abl   constitutional immunity

    'constitutional immunity from court-cases possessed by the president'

b. مقدمات سے صدر کو حاصل آئینی استثنٰی

    mʊqaddamat=se$_2$ sadar=ko$_1$ hasıl$_1$ aini ıstısna$_2$

c. صدر کو مقدّمات سے حاصل آئینی استثنٰی

    sadar=ko$_1$ mʊqaddamat=se$_2$ hasıl$_1$ aini ıstısna$_2$

d. *حاصل مقدّمات سے صدر کو آئینی استثنٰی

    *hasıl$_1$ mʊqaddamat=se$_2$ sadar=ko$_1$ aini ıstısna$_2$

e. *حاصل صدر کو مقدّمات سے آئینی استثنٰی

    *hasıl$_1$ sadar=ko$_1$ mʊqaddamat=se$_2$ aini ıstısna$_2$

The subscripted numbers in (18) show which arguments belong to which heads. The order of elements in (18a) seems to be canonical where arguments of both noun and adjective are close to their heads. The bracketing structure for (18a) is given as:

$[_{NP}[_{AP}[_{KP}$ sadar=ko$_1]$ hasıl$_1][_{KP}$ mʊqaddamat=se$_2]$ aini ıstısna$_2]$

The bracketed NP shows the logical structure and association of arguments in this noun phrase. The elements marked for case are called case phrases (Butt and King, 2005), hence the case marked arguments of nouns or adjectives are labelled as KP in the bracketed structure. All of the examples in (18a–c) are valid noun phrases of Urdu and are equivalent in meaning. The canonical order of elements in (18a), interestingly, is rarely found in newspaper corpora. Instead the orders in (18b)-(18c) are generally found in news corpora, with the latter one being the most common. In (18b) the argument-less adjective is adjacent to the head noun, then comes the argument-taking adjective with its argument to its left and the argument of the noun is at the left-most edge of the NP. The constituent AP in (18b) is contiguous; however, it becomes non-contiguous in (18c) where arguments of the adjective and the noun are in order on the left; then all the heads follow on the right. The orders in (18d–e) are ungrammatical due to a violation of the head-final

constraint in Urdu NPs. More examples of NP-internal discontinuity are shown in (19)–(20), where the head noun has a genitive marked argument and another ablative-marked argument.

(19) a. سلامتی پر بریفنگ کا آرمی چیف سے مطالبہ

    salamti=par$_1$ barif$\tilde{\text{i}}$g$_1$=ka armi-cif=se$_2$ mυtalbah$_2$
    security=Loc.on briefing=Gen army-chief=Abl demand
    'the demand to the army chief for a briefing on security'

   b. آرمی چیف سے سلامتی پر بریفنگ کا مطالبہ

    armi-cif=se$_2$ salamti=par$_1$ barif$\tilde{\text{i}}$g$_1$=ka mυtalbah$_2$

   c. سلامتی پر آرمی چیف سے بریفنگ کا مطالبہ

    salamti=par$_1$ armi-cif=se$_2$ barif$\tilde{\text{i}}$g$_1$=ka mυtalbah$_2$

(20) ملکی سلامتی پر آرمی چیف سے تفصیلی بریفنگ کا قانونی مطالبہ

   mυlki   salamti=par   armi-cif=se   tafsili
   of-country security.F.3Sg=Loc.on army.chief.M.3Sg=Abl detailed

   barif$\tilde{\text{i}}$g=ka   qanuni mυtalbah
   briefing.F.3Sg=Gen.M.3Sg legal   demand.M.3Sg

   'the legal demand to the army chief for a detailed briefing on the security of the country'

The argument *barif$\tilde{\text{i}}$g* 'briefing' of the head noun *mυtalbah* 'demand' in (19) takes its own case marked complement *salamti=par* 'on security'. All orders in (19a–c) are acceptable to native speakers and the last one is the most common in news corpora. Here, again we see that the argument of the genitive marked argument and the other argument of the head noun stack together on the left edge, and the genitive marked argument and the head noun are lumped together at the right edge. The noun phrase in (20) shows that the argument-less adjectives always have their positions adjacent to the head nouns.

5 Theoretical implications

In English, adjectives cannot take complements in their prenominal position. Consider the examples in (21).

(21) a. a proud mother

   b. a mother [proud of her daughter]

   c. *a [proud of her daughter] mother

   d. *a [yellow with age] manuscript (Maling 1983:284)

Emonds (1976) has proposed the 'Surface Recursion Restriction' according to which the modifiers themselves cannot be modified in prenominal position. Williams (1982) has formulated the 'Head-Final Constraint' in English: prenominal phrasal modifiers of nouns must be head-final. Due to this constraint no prepositional phrase (PP) can occur before nouns, as shown in (21c–d). The PP *of her daughter* is the complement of the adjective *proud* and *with age* is an adjunct of the adjective *yellow*.

Emond's restriction does not apply to Urdu as the modifiers at prenominal position can themselves be modified. William's Head final constraint only accounts for continuous constituents. So, it is also not applicable to Urdu as it does not account for the discontinuous AP within an NP.

Theoretical implications resulting from empirical observations of Urdu NPs can be stated in the form of two constraints. One is the adjacency constraint that the argument-less adjective is always contiguous to the head noun. In case there are more than one argument-less adjectives, their scrambling among each other is only possible in the vicinity of the head noun. The second constraint is the head-final constraint that the predicate heads cannot appear before their arguments or modifiers in NPs with the condition that the constituents within NPs may or may not be continuous. In Urdu NPs, the head noun appears at the right-most position (considering the order from left to right). The argument-less adjectives are placed just to the left of the head nouns and the rest of elements at the left edge can scramble among each other with the head final constraint in effect.

6   LFG implementation

As discussed above, examples of NPs from Urdu news corpora show that the arguments of the head noun and its modifiers/arguments can scramble inside the noun phrase, but that the heads must systematically follow their arguments. Non-continuous APs can appear inside a noun phrase. This evidence of discontinuous constituents within NPs implies a non-hierarchical structure of Urdu NPs. So we propose a flat structure for Urdu NPs. The pattern of Urdu NPs with the order of different elements is depicted in Figure 1. An excerpt from the grammar rules for the implementation of NPs in LFG is given in Figure 2.

Urdu NPs have been implemented in the XLE environment as part of a large scale Urdu grammar (Butt and King, 2007). To model discontinuous XPs at the constituent level within the LFG framework, use of several operators is made. The disjunction notation (|) has been used to assign various functional labels to the

```
                              NP
          ┌──────────┬────────┴────────┬──────────┐
       KP/PP        A₊                 A           N
          |          |                 |           |
    [Arg(N/A₊)   Arg-taking-modifiers]  Arg-less-modifiers   Head-noun
```
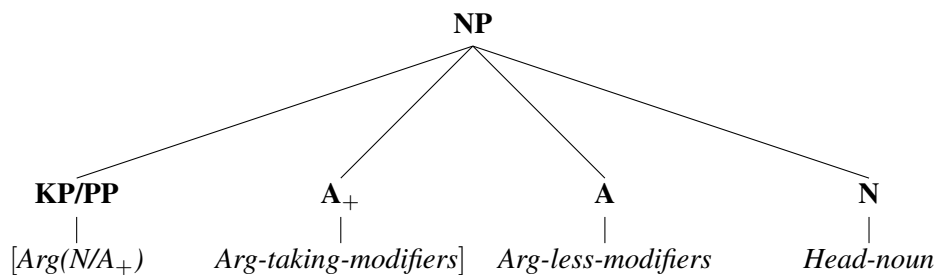
Figure 1: Word order in Urdu NPs: Elements in brackets can scramble among each other but the head-final constraint is in effect.

$$
\begin{aligned}
\text{NP} \rightarrow \text{KP*:} \quad &\{(\uparrow \text{ADJUNCT \$ OBL})=\downarrow \\
&|(\uparrow \text{ADJUNCT \$ OBJ-GO})=\downarrow \\
&|(\uparrow \text{OBL})=\downarrow \\
&|(\uparrow \text{OBJ-GO})=\downarrow\} \\
&, \text{`shuffle operator'} \\
\text{A}_+\text{*:} \quad &\downarrow \in (\uparrow \text{ADJUNCT}) \\
\text{A*:} \quad &\downarrow \in (\uparrow \text{ADJUNCT}) \\
\text{N:} \quad &\uparrow = \downarrow
\end{aligned}
$$

$$
\begin{aligned}
\text{NP} \rightarrow \text{KP*:} \quad &\{(\uparrow \text{ADJUNCT \$ OBL})=\downarrow \\
&(\uparrow \text{ADJUNCT}) >h\ (\uparrow \text{ADJUNCT \$ OBL}) \\
&| ....\} \\
&.....................................
\end{aligned}
$$

Figure 2: Grammar Rules

KP. The shuffle (,) operator establishes different word orders of the arguments in noun phrase. The ∈ sign has been used for two different purposes. It is used to add some element to an adjunct set. This is its general use. However, it is also used to assign some value nondeterministically to some feature of a member of the adjunct set. Both of its uses appear in the grammar rules for the NP. In the first two lines of the grammar rules in Figure 2, ∈ has been used to assign KP to the OBL function or OBJ-GO function of a member of the adjunct set. Another operator that has been taken advantage of is the head precedence operator (>h). The new rule with this operator is shown in the lower part of Figure 2. This operator is used for f-structure precedence and here it is used to make it sure that the head will not precede its arguments in the NP, thus implementing the Head-Final constraint. The possible c-structures for (18a–c) are shown in Figure 4. In (18c) the hierarchical structure of the AP inside the NP is not possible. So, a flat structure for Urdu NPs is assumed in general. The f-structure representation for each valid instance of (18)

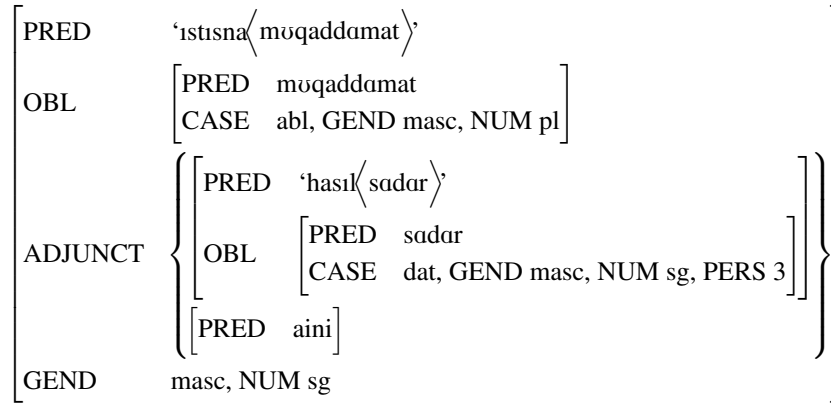is shown in Figure 3. In the f-structure we see that logical grouping of different elements is correctly captured.

$$
\begin{bmatrix}
\text{PRED} & \text{`ıstısna}\langle \text{mʊqaddɑmat} \rangle\text{'} \\[2ex]
\text{OBL} & \begin{bmatrix} \text{PRED} & \text{mʊqaddɑmat} \\ \text{CASE} & \text{abl, GEND masc, NUM pl} \end{bmatrix} \\[4ex]
\text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{`hasıl}\langle \text{sɑdɑr} \rangle\text{'} \\ \text{OBL} & \begin{bmatrix} \text{PRED} & \text{sɑdɑr} \\ \text{CASE} & \text{dat, GEND masc, NUM sg, PERS 3} \end{bmatrix} \end{bmatrix} \\ \begin{bmatrix} \text{PRED} & \text{aini} \end{bmatrix} \right\} \\[4ex]
\text{GEND} & \text{masc, NUM sg}
\end{bmatrix}
$$

Figure 3: f-structure representation of (18)

7 Conclusion

It has been shown that discontinuous constituents in Urdu can be found both at the clause level and at the noun phrase level. In Urdu NPs, discontinuous constituents arise when an argument taking noun is modified by an argument taking modifier or the argument of the head noun takes its own argument. The argument of the head noun and the arguments of its arguments/modifiers can scramble among each other with the head-final constraint in effect. This evidence suggests a flat structure for Urdu NPs. Alongside providing the theoretical implications of the phenomenon of constituent-level discontinuity, the syntax of Urdu NPs is implemented in the LFG framework. The existing theories about prenominal adjectives do not fit with the data of Urdu NPs. A constraint on adjacency of argument-less adjectives and a head final constraint allowing discontinuous constituents were posited to explain the syntax of Urdu NPs. Although existing apparatus in the LFG framework is sufficient to implement theoretical implications of syntax of Urdu NPs, the logical groupings of elements in these constructions are syntactically ambiguous and need semantics and pragmatics to disambiguate them.
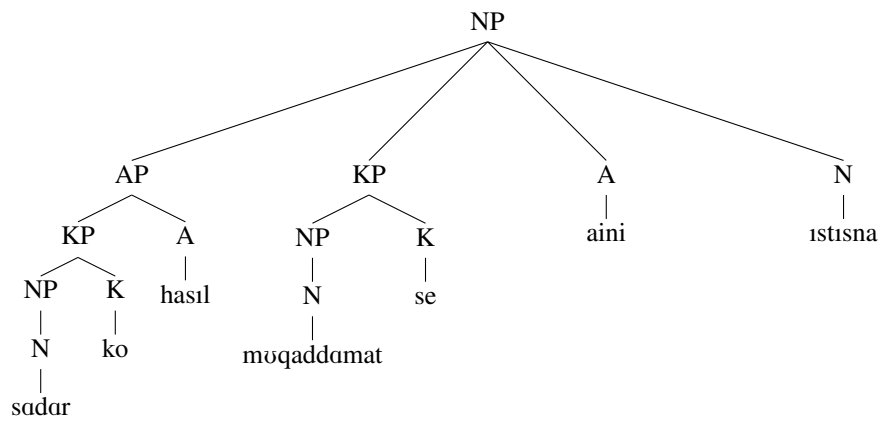
(a) C-structure for (18a)
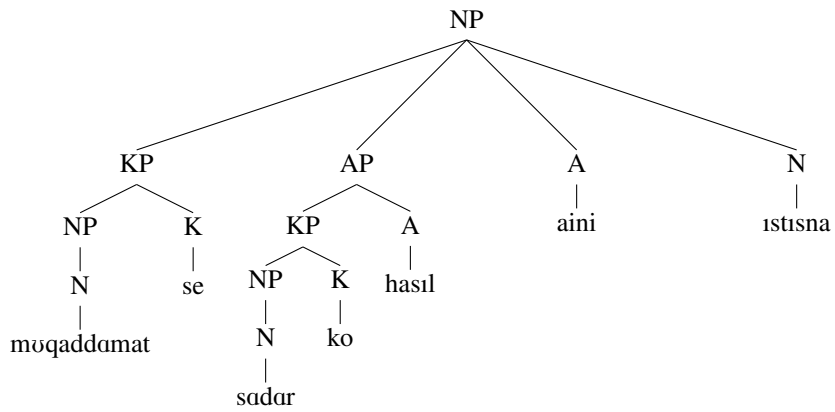
(b) C-structure for (18b)
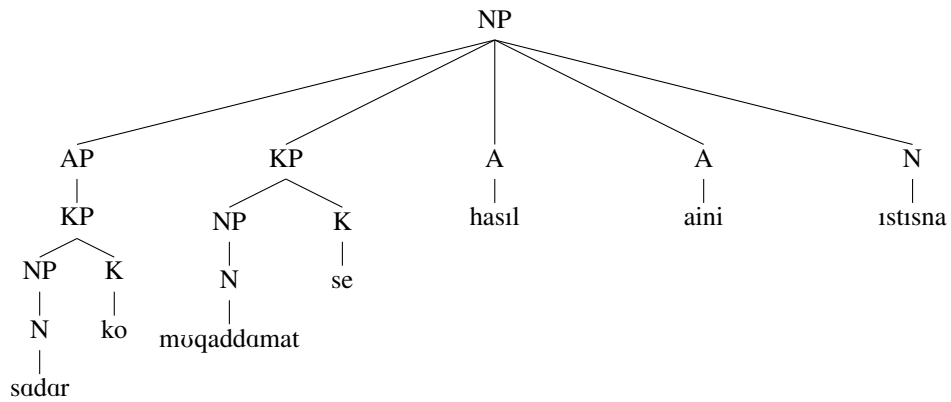
(c) C-structure for (18c)

Figure 4: C-structures for the instances in (18)

References

Agbayani, Bryan and Golston, Chris. 2005. Movements of Nonconstituents and Insensitivity to Islands in Ancient Greek. Rutgers Optimality Archive.

Austin, Peter and Bresnan, Joan. 1996. Non-configurationality in Australian Aboriginal Languages. *Natural Language and Linguistic Theory* (14.2), 15–68.

Bartsch, Renate. 1986/87. The Construction of Properties under Perspectives. *Journal of Semantics* (5), 293–320.

Bögel, Tina and Butt, Miraim. to appear. Possessive Clitics and Ezafe in Urdu. In Kersti Börjas and David Denison (eds.), *Morpho-syntactic categories and the expression of possession*, John Benjamins.

Boškovič, Željko. 2005. On the localitiy of left branch extraction and the structure of NP. *Studia Linguistica* (59), 1–45.

Butt, Miriam and King, Tracy Holloway. 2005. The Status of Case. In Dayal and Mahajan (eds.), *Clause Structure in South Asian Languages*, pages 153–198, Berlin: Springer Verlag.

Butt, Miriam and King, Tracy Holloway. 2007. Urdu in a Parallel Grammar Development Environment. *Language Resources and Evaluation, Special issue on Asian Language Processing: State of the Art Resources and Processing* pages 191–207.

Butt, Miriam, King, Tracy Holloway, Nino, María Eugenia and Segond, Frédérique. 1999. *A Grammar Writer's Cookbook*. CA: CSLI Publications.

Ćavar, Damir. 1999. *Aspects of the Syntax-Phonology Interface*. Ph. D.thesis, University of Potsdam.

Dayal, Vaneeta. 1994. Scope-marking as Indirect Wh-Dependency. *Natural Language Semantics* (2), 137–170.

Déprez, Viviane. 2003. *On the Typology of Syntactic Positions and the Nature of Chains*. Ph. D.thesis, MIT, Cambridge, Mass.

Dipper, Stefanie. 2003. *Implementing and Documenting Large-Scale Grammars — German LFG*. Ph. D.thesis, IMS, University of Stuttgart.

Dwivedi, Veena. 1994. *Syntactic Dependencies and Relative Phrases in Hindi*. Ph. D.thesis, University of Massachusetts, Amherst.

Emonds, Joseph. 1976. *A Transformational Approach to English Syntax*. New York: Academic Press.

Fanselow, Gisbert and Ćavar, Damir. 2002. Distributed Deletion. In Artemis Alexiadou (ed.), *Theoretical Approaches to Universals*, pages 65–107, Amsterdam, Benjamins.

Fanselow, Gisbert and Féry, Caroline. 2006. Prosodic and Morphosyntactic Aspects of Discontinuous Noun Phrases: a Comparative Perspective. Ms. Potsdam University.

Féry, Caroline, Paslawska, Alla and Fanselow, Gisbert. 2007. Discontinuous Constructions in Ukrainian. *Journal of Slavic Linguistics* (15.1).

Gouskova, Maria. 2001. Split Scrambling: Barriers as Violabale Constraints. In Karine Megerdoomian and Leona Anne Bar-el (eds.), *WCCFL Proceedings*, pages 220–233, Somerville, MA: Cascadilla Press.

Haider, Hubert. 1985. Über sein oder nicht sein: zur Grammatik des Pronomens sich. In *Erklärende Syntax des Deutschen. [Studien zur deutschen Grammatick 25]*, pages 223–254, Tübingen: Narr.

Hale, Kenneth. 1980. *On the position of Warlpiri in a Typology of the Base*. Bloomington: Indiana University Linguistics Club.

Hale, Kenneth. 1983. Warlpiri and the Grammar of Non-configurational Languages. *NLLT* (1), 5–74.

Hale, Kenneth. 1992. Basic Word Order in Two "Free Word Order" Languages. In *Pragmatics of Word Order Flexibility*, pages 63–82, John Benjamins, Amsterdam.

Ikeya, Arika. 1995. Predicate-argument Structure of English Adjectives. In Benjamin K T'sou and Tom BY Lai (eds.), *Proceedings of the 10th Pacific Asia Conference on Language, Information and Computation*, pages 149–156.

Kazenin, Konstantin. 2005. A Syntactic Analysis of Split Constructions in Russian. Ms., University of Potsdam.

Kuhn, Jonas. 1998. Resource Sensitivity in the Syntax-semantics Interface and the German Split NP Construction. In *Proceedings of the ESSLLI X Workshop on CurrentTopics in Constraint Based Theories of Germanic Syntax*, Saarbrücken.

Laczkó, Tibor. 2000. Derived Nominals, Possessors and Lexical Mapping Theory in Hungarian DPs. In Miriam Butt and Tracy H. King (eds.), *Argument Realization*, pages 189–227, Stanford: CSLI Publications.

Legate, Julie. 2002. *Warlpiri: Theoretical Implications*. Ph. D.thesis, MIT, Cambridge, Mass.

Maling, Joan. 1983. Transitive adjectives: a case of categorial reanalysis. In F. Henry and B. Richards (eds.), *Linguistic Categories: Auxiliaries and Realted Puzzles*, volume 1, pages 253–289, Dordrecht: D Reidel.

McCloskey, James. 2000. Quantifier Float and Wh-Movement in an Irish English. *Linguistic Inquiry* 31(1), 57–84.

Müller, Stefan. 2004. Continuous or Discontinuous Constituents? A Comparison between Syntactic Analyses for Constituent Order and Their Processing Systems. *Research on Language and Computation* (2), 209–257.

Nthelitoes, Dimitris. 2004. *Syntax of Elliptical and Discontinuous Nominals*. Masters Thesis, UCLA.

Platts, John. 1967. *A Grammar of the Hindustani or Urdu Language*. London: Crosby Lockwood and Son, 1909, republished in 2002 by Sang-e-Meel Publications, Lahore.

Raza, Ghulam. 2010. Analyzing the stucture of Urdu NPs with multiple genitives. In *Proceedings of the Conference on Language and Technology 2010 (CLT10)*, Islamabad.

Raza, Ghulam. 2011. *Subcategorization Acquisition and Classes of Predication in Urdu*. Ph. D.thesis, University of Konstanz.

Roehrs, Dorian. 2006. Split-NP as an Instance of Sideward Movement. University of Pennsylvania Working Papers in Linguistics 13.1, 1-14.

Schmidt, Ruth Laila. 1999. *Urdu: An Essential Grammar*. London: Routelege.

Sekerina, Irina. 1997. *The Syntax and Processing of Scrambling Constructions in Russian*. Ph. D.thesis, CUNY.

Simpson, Jane. 1983. *Aspects of Warlpiri Morphology and Syntax*. Ph. D.thesis, MIT.

Simpson, Jane. 1991. *Warlpiri Morpho-Syntax: A Lexicalist Approach*. Dordrect: Kluwer.

van Hoof, Hanneke. 1997. Left Dislocation and Split Topics. In Elena Anagnostopoulou, Henk van Riemsdijk and Frans Zwarts (eds.), *Materials on Left Dislocation*, volume 14 of Liguistic Aktuell, pages 275–305, Amsterdam and Philadelphia: John Benjamins.

van Riemsdijk, Henk. 1989. Movement and Regeneration. In *Dialectal Variation and Theory of Grammar*, pages 105–136, Dordrecht: Foris.

Williams, Edwin. 1982. Another Argument that Passive is Transformational. *Linguistic Inquiry* (13), 160–163.