The Structure of Urdu — Case

Miriam Butt
Universität Konstanz
`miriam.butt@uni-konstanz.de`

1   The Language

Urdu is a South Asian language spoken in:

- Pakistan (national language)

- India (one of the 18 official languages)

- World-Wide due to South Asian Diaspora (big populations in U.K., U.S.A., Canada, etc.)

Urdu is closely related to Hindi.
Taken together these represent the second most spoken language in the world: **400 000 000**.

Properties: SOV, subject and object agreement, non-nominative case including an ergative.

2   Clause Structure

- SOV, fairly free word order.

- Everything head-final, except for some complementizers (cf. Bayer 1999), relative clauses.

- Word order determined by information structure (Butt and King 1996, 1997).

- Rampant Pro-drop. Also determined by information structure. (Butt and King 1997, Prasad 2003).

3   The Nominal Domain

- No definite determiners (demonstratives, one indefinite determiner). Specificity marked via the accusative case clitic *ko*, as in Turkish (Enç 1991).

- Adjectives agree in number and gender with the noun. Mostly preverbal.

- Quantifiers are prenominal, but can sometimes appear postnominally. *Anybody, nobody, somebody*, etc. are realized compositionally.

- There are non-nominative subjects (see Mohanan 1994 for subject tests).

- Correlatives (Dayal, Bhatt).

4   Case

This section is based on Butt and King (2005): The Status of Case. PDF file available at:
`http://ling.uni-konstanz.de/pages/home/butt/`

> The term case is from Latin *cāsus*, which is in turn a translation of the Greek *ptōsis*
> 'fall'. The term originally referred to verbs as well as nouns and the idea seems to
> have been of falling away from an assumed standard form . . . [Blake 2001:18]

(1)　　　Sample Sanskrit Declension

Number	Declination	Western name
1	*devas*	nominative
2	*devam*	accusative
3	*devena*	instrumental
4	*devāya*	dative
5	*devāt*	ablative
6	*devasya*	genitive
7	*deve*	locative

Case Forms in Middle Indo-Aryan

	Singular	Plural
Nominative	*-u, a, aṁ*	*-a, aĩ*
Accusative	[same as Nominative]	
Instrumental	*-eṁ, iṁ, he, hi*	*-e(h)ĩ, ehi, ahĩ*
Ablative	*-hu, ahu, aho*	*-hũ, ahũ*
Genitive/Dative	*-ho, aho, ha, su, ssu*	*-na, hã*
Locative	*-i, hi, hiṁ*	*-hĩ*

These forms were either lost or were collapsed into the modern oblique marking (Beames 1872:209).
The genitive and locative forms seem to have fallen together in Apabhraṃsa, and in old (or archaic)
Hindi the ablative, dative, and accusative singular were then also collapsed (Kellogg 1893:126).

The Modern Case Markers

Clitic	Case	Gram. Func.	Morph. Effect
∅	nominative	subj/obj	none
ne	ergative	subj	oblique marking on NP
ko	accusative	obj	oblique marking on NP
	dative	subj/ind. obj	oblique marking on NP
se	instrumental	subj/obl/adjunct	oblique marking on NP
k-	genitive	subj (infinitives)	agrees with head noun
		specifier	none
mẽ/pɑr/tɑk/∅	locative	obl/adjunct	oblique marking on NP

4.1 Ergative

The ergative is confined to subjects and must appear on transitive verbs in the perfect. Also notice the agreement pattern.

(2) a. yassin=ne   gari   xarid-i
     Yassin.M.Sg=Erg car.F.Sg.Nom buy-Perf.F.Sg
     'Yassin bought a car.'

     b. yassin   gari   xarid-e-g-a
     Yassin.M.Sg.Nom car.F.Sg.Nom buy-3.Sg-Fut-M.Sg
     'Yassin will buy a car.'

But, there are exceptions.

(3) nadya   kıtab   la-yi
     Nadya.F.Sg.Nom book.F.Sg.Nom bring-Perf.F.Sg
     'Nadya brought a book.'

Unergative verbs generally may appear with an ergative. This depends on the semantics to be expressed.

(4) a. ram   kʰãs-a
     Ram.M.Sg.Nom cough-Perf.M.Sg
     'Ram coughed.' (Tuite, Agha and Graczyk 1985:264)
     b. ram=ne   kʰãs-a
     Ram.M.Sg=Erg cough-Perf.M.Sg
     'Ram coughed (purposefully).' (Tuite, Agha and Graczyk 1985:264)

The ergative also alternates with a dative Inf+be expressions (modality).

(5) a. nadya=ne   zu   ja-na   hε
     Nadya.F.Sg=Erg zoo.M.Sg.Obl go-Inf.M.Sg be.Pres.3.Sg
     'Nadya wants to go to the zoo.'
     b. nadya=ko   zu   ja-na   hε
     Nadya.F.Sg=Dat zoo.M.Sg.Obl go-Inf.M.Sg be.Pres.3.Sg
     'Nadya has to go to the zoo.'

Butt and King (1991) and Mohanan (1994) therefore argue that the ergative is associated with volitionality or the feature [+conscious choice]. Bashir (1999), based on an examination of current usage of the ergative in modern day Urdu TV dramas, concludes that the picture is not so simple. She observes the pattern below and proposes an explanation in terms of markedness.

The Ergative and Semantic Entailments

	Tense/Aspect	Valency	Ergative	Semantic Entailment
a.	Finite, Perfect	Intransitive Unergative	No	No entailment
b.			Yes	[+conscious choice]
c.	Finite, Perfect	Transitive	No	Exceptional, No entailment
d.			Yes	No entailment
e.	Infinitive	Any	No	[−source specified]
f.			Yes	[+source specified]

4.2   Nominative

The nominative is phonologically null. It can alternate with accusative *ko* to produce specificity effects. Notice the agreement patterns.

(6) a. nadya=ne   gaṛi   cɑla-yi   hɛ
     Nadya.F.Sg=Erg car.F.Sg.Nom drive-Perf.F.Sg be.Pres.3.Sg
     'Nadya has driven a car.'

    b. nadya=ne   gaṛi=ko   cɑla-ya   hɛ
     Nadya.F.Sg=Erg car.F.Sg=Acc drive-Perf.M.Sg be.Pres.3.Sg
     'Nadya has driven the car.'

A clause may contain more than one nominative argument.

(7) nadya   gaṛi   cɑla-ti   hɛ
    Nadya.F.Sg.Nom car.F.Sg.Nom drive-Impf.F.Sg be.Pres.3.Sg
    'Nadya drives a car.'

4.3   Accusative

The accusative is form-identical with the dative *ko*. Many approaches therefore assume that Urdu/Hindi lacks an accusative and that the *ko* is an inherent dative case (e.g., Mahajan 1990, Davison 1998). However, there are two distinct distributional patterns with regard to *ko*.
As shown in (6), the accusative alternates with the nominative on objects to express specificity. Furthermore, *ko* is associated with a notion of affectedness (Saksena 1982). Relevant examples come from causativization patterns.

(8) a. ɑnjʊm=ne   saddɑf=ko/*se   kʰana   kʰɪl-a-ya
     Anjum.F.Sg=Erg Saddaf.F.Sg=Acc/Inst food.M.Sg.Nom eat-Caus-Perf.M.Sg
     'Anjum made Saddaf eat food (gave Saddaf food to eat).'

    b. ɑnjʊm=ne   saddɑf=se/*ko   paoda   kaṭ-a-ya
     Anjum.F.Sg=Erg Saddaf.F.Sg=Inst/Acc plant.M.Sg.Nom cut-Caus-Perf.M.Sg
     'Anjum had Saddaf cut a/the plant.'

(9) a. ɑnjʊm=ne   saddɑf=ko   mɑsala   cɑkʰ-va-ya
     Anjum.F.Sg=Erg Saddaf.F.Sg=Acc spice.M.Sg.Nom taste-Caus-Perf.M.Sg
     'Anjum had Saddaf taste the seasoning.'

    b. ɑnjʊm=ne   saddɑf=se   mɑsala   cɑkʰ-va-ya
     Anjum.F.Sg=Erg Saddaf.F.Sg=Inst spice.M.Sg.Nom taste-Caus-Perf.M.Sg
     'Anjum had the seasoning tasted by Saddaf.'

4.4   Dative

The dative is identical in form to the accusative. It differs from the accusative in that it marks indirect objects, as in (10), and subjects, as in (11), and never alternates with nominative objects. The dative indirect object in (10) never becomes subject under passivization, unlike the accusative.

(10) ɑnjʊm=ne   sɑddɑf=ko   cɪtʰi   d-i
      Anjum.F.Sg=Erg Saddaf.F.Sg=Dat letter.F.Sg.Nom give-Perf.F.Sg
      'Anjum gave Saddaf a letter.'

The dative is associated with the *θ*-role *goal/experiencer* (cf. Verma and K.P. Mohanan (1990) on experiencer subjects, and Mohanan 1994 and references therein). Experiencer subjects encompass modal contexts such as in (11a), psych predicates as in (11b), the alternation with the ergative as in (5), and subjects of N-V complex predicates, as in (11c).

(11) a. nadya=ko   skul   ja-na   pɑṛ-a
         Nadya.F.Sg=Dat school.F.Sg.Obl go-Inf.M.Sg fall-Perf.M.Sg
         'Nadya had to go to school.'
      b. nadya=ko   ḍɑr   lɑg-a
         Nadya.F.Sg=Dat fear.M.Sg.Nom be attached-Perf.M.Sg
         'Nadya was afraid.'
      c. nadya=ko   kɑhani   yad   a-yi
         Nadya.F.Sg=Dat story.F.Sg.Nom memory come-Perf.F.Sg
         'Nadya remembered the story.'

Historical Development

The dative/accusative *ko* is generally traced to the Sanskrit locative noun *kákshe* 'armpit, side' (Kellogg 1893:130).

4.5   Instrumental

The instrumental *se* is extremely versatile. It may be used for instrumental adjuncts as in (12a), for source expressions, both locative, as in (12b), and material, as in (12c), as well as for comitatives, as shown in (12d), and for causees, as in (9).

(12) a. nadya=ne   dɑrvaza   cabi=se   kʰol-a
         Nadya.F.Sg=Erg door.M.Sg.Nom key.F.Sg=Inst open-Perf.M.Sg
         'Nadya opened the door with a key.'

      b. nadya=ne   aj   lahor=se   fon   ki-ya
         Nadya.F.Sg=Erg today Lahore=Inst phone do-Perf.M.Sg
         'Nadya called from Lahore today.'
      c. sʊnar=ne   sone=se   har   bɑna-ya
         goldsmith.M.Sg=Erg gold.M.Sg.Obl=Inst necklace.M.Sg.Nom make-Perf.M.Sg
         'The goldsmith made a necklace out of the gold.'
      d. nadya   sɑddɑf=se   bat   kɑr rɑh-i   hɛ
         Nadya.F.Sg.Nom Saddaf.F.Sg=Inst talk.F.Sg.Nom do stay-Perf.F.Sg be.Pres.3.Sg
         'Nadya is talking to Saddaf.'

The instrumental is also used on adjuncts that express the demoted agent (logical subject), as in the standard passive in (13). Standard passives are formed with the verb *ja* 'go' (in all tenses) in combination with perfect morphology on the main verb.

(13) cor   (pʊlɪs=se) pakṛ-a   gɛ-ya/ja-ta
      thief.M.Sg.Nom police=Inst catch-Perf.M.Sg go-Perf.M.Sg/go-Impf.M.Sg
      'The thief was caught by the police.' (adapted from Mohanan (1994:183))

The instrumental also occurs in a construction described as a passive of disability in some grammars of Urdu/Hindi (e.g., Glassman 1976, Van Olphen 1980), as in (14).

(14) a. nadya=se   yə ʊrdu=k-i   cɪṭʰi   paṛʰ-i   nɑhĩ
       Nadya.F.Sg=Inst this Urdu=Gen-F.Sg letter.F.Sg.Nom read-Impf.F.Sg not
       ja-ti
       go-Impf.F.Sg
       'Nadya does not have the ability to read this Urdu letter.'

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

This (dis)ability "passive" differs syntactically from the standard passive. It is possible with intransitives, as in (14b), unlike the standard passive. The instrumental NP is obligatory and exhibits subject properties with regard to control and anaphora (see Mohanan 1994 for a list of diagnostics for grammatical subjecthood in Hindi).[1]

Historical Development

The instrumental *se* may either be connected with Sanskrit *sam* 'with' or with the locative singular noun *sañge* 'in attachment to' (Kellogg 1893:132).

4.6 Genitive

The genitive may be roughly characterized as marking subjects of nonfinite clauses, as in (15a), subjects of finite copula constructions, as in (15b), and specifiers of nominals, as in (15c). Like other case marked nominals in Urdu/Hindi, genitives may be scrambled. As such, they are functionally, but not phrase structurally determined. Genitives are not semantically motivated: Mohanan (1994:177) considers and discards an analysis in terms of the semantic notion possession.

(15) a. ram=ke   bɛṭʰ-ne=pɑr   mã=ne   ʊs=ko   kʰana
       Ram.M.Sg=Gen.M.Sg.Obl sit-Inf.Obl=on mother.F.Sg=Erg Pron=Dat food.M.Sg.Nom
       di-ya
       give-Perf.M.Sg
       'On Ram's sitting down, the mother gave him food.' (Adapted from Mohanan 1994:78)

---

[1] See Butt (1997), attached to the end of this handout, for an analysis of this construction as a complex predicate with dispositional semantics whose subject is an instrumental NP. Bhatt (1998) proposes an alternative analysis in terms of negative polarity. Davison (1990) examines this construction within a larger discussion on "peculiar passives".

    b. ram=ka   ek beṭa   hɛ

       Ram.M.Sg=Gen.M.Sg one son.M.Sg.Nom be.Pres.3.Sg

       'Ram has one/a son.' (Adapted from Mohanan 1994:177)

    c. rani=ka   bʰai

       Rani.F.Sg=Gen.M.Sg brother.M.Sg.Nom

       'Rani's brother'

In general, the genitive can be analyzed as marking specifiers. Were it not for the rather compli-
cated agreement pattern associated with it, this case would be crosslinguistically unremarkable.
As can be seen from the examples in (15), the genitive inflects to agree with the head noun in
terms of gender, number, and obliqueness. Payne (1995) discusses the Hindi genitive and views
the agreement pattern as an instance of *Suffixaufnahme*.

Historical Development
The fact that the genitive inflects can be traced directly to its historical origin. After a fierce
debate in the last century, the view espoused by Hoernle won out and was taken over by Beames
(1872:285) and Kellogg (1893:129). Under this view, the genitive is analyzed as having arisen
from *kṛita* 'done by', the Sanskrit past participle of *kṛi* 'do' as follows. Sanskrit *kṛita* > Prakrit
kerita > *keriai* > modern Urdu/Hindi *k-*. The original participal inflected for agreement and the
genitive case marker has not lost this property.

4.7 Pronouns

Pronoun Paradigm

	NOM	ERG	ACC/DAT	INST	LOC	GEN
1.Sg	mɛ̃	mɛ̃=ne	mʊjʰ=ko mʊjʰe	mʊjʰ=se	mʊjʰ=pαr	mer-a/i/e
1.Pl	hαm	hαm=ne	hαm=ko hαme	hαm=se	hαm=pαr	hαmar-a/i/e
2.Disresp.	tu	tu=ne	tʊjʰ=ko tʊjʰe	tʊjʰ=se	tʊjʰ=pαr	ter-a/i/e
2.Familiar	tʊm	tʊm=ne	tʊm=ko tʊmhe	tʊm=se	tʊm=pαr	tʊmhar-a/i/e
2.Resp.	ap	ap=ne	ap=ko	ap=se	ap=pαr	ap=k-a/i/e
3.Prox.Sg	ye	ıs=ne	ıs=ko ıse	ıs=se	ıs=pαr	ıs=k-a/i/e
3.Prox.Pl	ye	ın=ne ınhõ=ne	ın=ko ınhõ=ko ıne	ın=se ınhõ=se	ın=pαr ınhõ=pαr	ın=k-a/i/e ınhõ=k-a/i/e
3.Dist.Sg	vo	ʊs=ne	ʊs=ko ʊse	ʊs=se	ʊs=pαr	ʊs=k-a/i/e
3.Dist.Pl	vo	ʊn=ne ʊnhõ=ne	ʊn=ko ʊnhõ=ko ʊnhɛ̃	ʊn=se ʊnhõ=se	ʊn=pαr ʊnhõ=pαr	ʊn=k-a/i/e ʊnhõ=k-a/i/e

4.8   Clitics vs. Affixes

The Urdu case markers are clitics.

- They do not pattern like affixes with respect to coordination.

- Focus clitics can separate the case markers from the head noun.

(16) a. *[[kʊtt or gʰoɽ]-e]=ko   b. *[[kʊtt-a   or gʰoɽ]-e]=ko
        dog and horse-M.Sg.Obl=Acc   dog-M.Sg and horse.M.Sg-Obl=Acc

(17) a. yasin=ne   [kʊtt-e   or gʰoɽ-e]=ko   dekʰ-a   hɛ
        Yassin.M.Sg=Erg dog-M.Sg.Obl and horse-M.Sg.Obl=Acc see-Perf.M.Sg be.Pres.3.Sg
        'Yassin saw the dog and the horse.'

   b. nadya   [lahor or karacʰi]=se   hɛ
      Nadya.F.Sg.Nom Lahore and Karachi=Inst be.Pres.3.Sg
      'Nadya is from Lahore and Karachi.'

(18) a. ʊs=hi=ne   kam   ki-ya
        Pron.3.Sg=Foc=Erg work.M.Sg.Nom do-Perf.M.Sg
        'That one himself/only did (the) work.'

   b. tʊjʰ=hi=ko   di-ya
      you.Obl=Foc=Dat give-Perf.M.Sg
      'I gave it to you (and not to anyone else).'   (Platts 1967:300)

   c. mɛ̃   vahã saikɑl=hi=se   pɑhṹcʰ sɑk-ti   hṹ
      I.Nom there bicycle=Foc=Inst reach   able-Impf.F.Sg be.Pres.1.Sg
      'I can get there with just a bike.'   (Sharma 1999)

   d. Gauri: to ʊs=se   mɑtlɑb?
            so Pron.3.Sg.Obl=Inst meaning.M.Sg.Nom

      Lakʰa: hɑm=hi=se   sara   mɑtlɑb
             Pron.3.Pl=Foc=Inst all.M.Sg.Nom meaning.M.Sg.Nom
      Gauri: 'So, what meaning of that [what's the meaning of that]?'
      Laka: 'All meaning is from us/me alone [the meaning is all of my knowing].'
      (Lɑgan 'Tax', Hindi Movie)

4.9   Clitics vs. Postpositions

Due to the above properties, and due to the fact that the case markers attach postnominally, case endings have been described as postpositions in many accounts of Urdu/Hindi. Again following Mohanan (1994), we wish to make clear that postpositions differ from the case markers in terms of form and distribution. Consider the typical postpositions in (19) and example in (20).

(19)   ke picʰe   'behind'   ke pɛhle   'before'
       ke nicʰe   'under'   ke pas   'next to'
       ke upɑr   'over'   ke satʰ   'with'
       ke ɑndɑr   'inside'   ke liye   'for'
       ke samne   'in front of'   ke tɑrɑf   'in the direction of'
       ke age   'in front of (further along)' ke bad   'after'

(20) bɪlli   bɪstar   [ke   nicʰe] so   rah-i   hɛ
      cat.F.Sg.Nom bed.M.Sg Gen.Obl under sleep stay-Perf.F.Sg be.Pres.3.Sg
      'The cat is sleeping under the bed.'

All the items in (19) correspond to a relational preposition in English. In contrast to English, they appear after the NP. The *ke* in each of these postpositions is the oblique form of the genitive, which is now invariant for most postpositions. For the native speaker, this invariant *ke* is not associated with the genitive, but is taken to be part of the postposition. Similarly, the final *-e* on some of the postpositions in (19) is presumably a reflex of the oblique marking, which indicated a locative. As shown in (21), some postpositions allow inflection and establish the link to an original genitive construction (also see Masica 1991:234) in which the postpositions of today were nouns linked to another noun via the genitive (cf. English *because of*).

(21) mɛ̃=ne bɪlli=ka   picʰa   ki-ya
      I=Erg   cat.F.Sg=Gen.M.Sg behind.M.Sg.Nom do-Perf.M.Sg
      'I went after the cat.'

In addition, the contentful part of the postposition may appear by itself, as in (22). This is not the case for any of the case clitics, including the locatives *mẽ, par,* and *tak* listed in the case table.

(22) upɑr ao
      up   come.Impf
      'Come up!'

There is thus a clear difference in distribution and form between the case markers and the *ke* postpositions. However, the locative case markers (*mẽ* 'in', *par* 'on', *tak* 'towards') do pattern with the postpositions in one respect. Although Urdu/Hindi does not normally exhibit case stacking, *se* 'from' (and *mẽ* 'in') may stack on top of locatives, be they case markers, as in (23a), or postpositions, as in (23b).

(23) a. ʋn   lõgõ=mẽ=se   tin
         that.Pl.Obl people.Obl.Pl=in=from three
         'three from among those people'
     b. ɑlmari   [ke   picʰe]=se
         cupboard Gen.Obl behind=from
         'from behind the cupboard'

However, the relevant generalization is over locatives, not over a particular syntactic class (case clitics vs. *ke* postpositions). It would therefore be a mistake to base the identification of case clitics with postpositions on this one argument, especially as the case clitics can all appear on subject noun phrases, while noun phrases with *ke* postpositions do not mark subjects (Mohanan 1994). As such, postpositions must be distinguished from case clitics. In the following section, we propose that case clitics are heads of a KP, whereas postpositions are Ps which head a PP.

4.10 Structural Representation of Case Markers

Crosslinguistically it has long been noticed that only certain types of words become clitics (see Sadock 1991 for discussion). In order to capture this generalization, it has been proposed that functional heads can be clitics, while lexical categories such a nouns cannot, unless they undergo historical development and change category along with the change in prosodic status (see Franks 2000 and references therein).

The Urdu/Hindi data support this idea, and we propose that the case endings are functional heads of a KP (KaseP). We assume that the oblique marking on masculine nouns ending in -a (singular -e in (24)) is synchronically the result of the complement-head relationship between the K and the NP. This marking is obligatory when there is an overt K head.

(24) a. larke=ne   b.
       boy.M.Sg.Obl=Erg

4.10.1 Genitives

We analyze genitives as occuring in the SpecNP position.[2] For a similar analysis of genitives see Davison (1998).

The agreement between the genitive *k-* and the head noun is due to the fact that this case marker originated from a participial construction. This agreement can be viewed synchronically as the result of NP internal agreement.

(25) a. asım=ke   b.
       Asim.M.Sg=Gen.Obl

       larke=ne
       boy.M.Sg.Obl=Erg

4.10.2 Bare Nominatives

Bare nominatives, i.e., the nominals which have no overt case ending and no oblique inflection, distribute syntactically like KPs with overt case marking. We therefore assume that these project a KP, albeit one without an overt K head, as in (26b). Since LFG does not posit empty categories, the K head of the KP is not projected in the structure in (25b) (see King 1995 and Bresnan 2001 for

---

[2]This structure allows for coordination data such as that in (i).

(i) asım=ke   pıtta ji   or   amir=ke   dada ji   ne
   Asim.M.Sg=Gen.Obl father Resp and Amir.M.Sg=Gen.Obl grandfather Resp Erg
   'Asim's father and Amir's grandfather'

constraints on LFG phrase structure). Under the assumption that the oblique endings are the result of the overt K head's requirements on its complement, no oblique ending occurs in the nominative. The nominative case comes from default rules (section 4.13) which state that subject and object KPs in Urdu/Hindi require case and that if there is no other case, the nominative is assigned.

(26) a. laṛka
        boy.M.Sg.Nom

b. KP
   |
   NP
   |
   N
   laṛka

4.10.3 Bare Locatives

Another type of bare nominal exists in Urdu/Hindi: locatives as in (27). Since these distribute like the overtly case marked nominals, we again assume a KP in which the K head is not projected. Again, the locative case and the feature structure associated with it is associated with the KP via default rules (section 4.13).

(27) adnan   ḍakxane/zu   gɛ-ya   hɛ
     Adnan.M.Sg.Nom post office.M.Sg.Obl/zoo.M.Sg.Obl go-Perf.M.Sg be.Pres.3.Sg
     'Adnan has gone to the post office/zoo.'

(28) a. zu   b. KP   c. ḍakxane   d. KP
        zoo.M.Sg.Obl | post office.M.Sg.Obl |
                             NP   NP
                             | |
                             N   N
                             zu   *ḍakxane*

We now need to account for the presence of the oblique inflection. In this instance, we analyze the oblique *e* as a case marker that is a bound morpheme whose surface realization is governed by the morphophonological properties of the nominal (e.g., masculine nouns ending in *a* overtly realize this morpheme, as in (28c)).
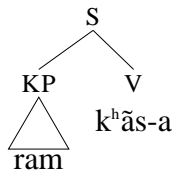
4.11 Case and Clausal Structure

In this section we provide an overview of the LFG architecture as relevant to case assignment. Section 4.16 provides analyses of particular cases, focusing on the ergative and dative.
In LFG, information from different components combines to produce a consistent and coherent analysis. The different modules of grammar (e.g., grammatical functions, semantics, and phonological information) are encoded in terms of projections from lexical entries and phrase structure rules, which in turn encode syntactic and morphological constituency. This is informally illustrated in (30) for (29).

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

(30) a. Constituent-structure: b. Functional-structure:

$$
\begin{array}{c}
S \\
\diagup\diagdown \\
KPV
\end{array}
\qquad
\begin{bmatrix}
\text{PRED} & '\text{cough}<\text{SUBJ}>' \\
\text{TNS-ASP} & \begin{bmatrix} \text{TENSE} & \text{PAST} \end{bmatrix} \\
\text{SUBJ} & \begin{bmatrix} \text{PRED} & '\text{Ram}' \\ \text{CASE} & \text{NOM} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \end{bmatrix}
\end{bmatrix}
$$

k^{h}ãs-a

ram

A priori the role of case marking can be handled in various ways in LFG. For example, the architecture does not presuppose an intimate connection between case, agreement and structural position, although this would be one possibility.

4.12 Grammatical Functions and Mapping Theory

The association of grammatical functions with thematic roles is handled via a flexible, yet constrained theory of *Mapping* (see Bresnan and Zaenen 1990 for an overview). Arguments of a predicate are specified in the lexicon with the features [± r(estricted)] and/or [± o(bjective)]. Roughly, patient-like roles are [−r], secondary patient-like roles are [+o], and other roles are [−o]. For example, the a(rgument)-structure of the English verb *pound* would look as in (31).[3] These specifications constrain the way arguments are associated with grammatical functions, which are also classified by means of these features as shown in (32).

(31) a-structure *pound* <   ag   pt   >
                                [−o] [−r]

(32)

Gram. Functions	Features	Gram. Functions	Features
SUBJ	[−r, −o]	OBL$_\theta$	[+r, −o]
OBJ	[−r, +o]	OBJ$_\theta$	[+r, +o]

The intrinsic role classifications of the argument structure are related to the fully specified grammatical functions by mapping principles (not discussed here), as in (33).

(33) a-structure *pound* <   ag   pt   >
                                [−o] [−r]
                                 | |
      f-structure   SUBJ OBJ

4.13 Structural Case

Structural case involves case assigned on the basis of syntactic information. It is usually correlated with grammatical function. It may also be associated with phrase structure position.
Structural case is often an instance of default case and hence functions as the Elsewhere Case (cf. Zaenen, Maling, and Thraínsson's (1985) notion of default vs. lexically stipulated case). For languages which require that all NPs have case, this can be stated as in (34a), analagous to the Case Filter.

---

[3]This a-structure can be conceived of as an attribute-value matrix (Butt 1998).

In Urdu/Hindi the nominative is a default case. As such, there are principles which assign nominative case to subjects and objects, as in (34b,c). In languages in which all subjects have nominative case, (34b) would be obligatory; in languages such as Urdu/Hindi, in which there are non-nominative subjects, the default principles are optional and only apply if nothing else assigns case to the subject.

(34) a. Wellformedness principle: KP: (↑CASE)

b. Default: ((↑SUBJ CASE)=NOM)

c. Default: ((↑OBJ CASE)=NOM)

While the identification of grammatical functions is not necessarily tied to positional information within Mapping Theory, some languages may restrict a position to a particular case-marked grammatical function. Thus, non-thematic grammatical relations and their corresponding case marking may be licensed by structural position (King 1995). We have not found an example of positional case in Urdu/Hindi.[4]

4.14 Semantic Case

We take semantic case to be the most general type of case marking in Urdu/Hindi. The defining characteristics of semantic case are: (i) predictability via the formulation of generalizations across predicates and constructions; (ii) a subjection to syntactic restrictions (such as only appearing on certain grammatical functions). The association of case morphology with grammatical functions can be restricted by the case markers themselves. This is in line with Nordlinger (1998), who proposes the notion of *constructive case* for Australian languages whereby the case morphology provides information as to grammatical relations.

4.15 Quirky Case

Finally, quirky case is used only when there is no regularity to be captured: the case assignment is truly exceptional to the system. For example, consider the Urdu/Hindi transitive verb *la* 'bring' in (35). The subject should be ergative since this is a perfect transitive verb. However, it is nominative; this requirement must be stipulated in the lexical entry, as in (36).

(35) nadya   kıtab   la-yi
     Nadya.F.Sg.Nom book.F.Sg.Nom bring-Perf.F.Sg
     'Nadya brought a book.'

(36) *la* 'bring'   (↑PRED)=< ag[−o] th[−r] >
                    (↑SUBJ CASE) = NOM

Our notion of quirky case is extremely restricted. Quirky case only occurs when *no* generalizations can be made about the choice of case with the predicate in question.

---

[4]A canonical example is the assignment of adnominal genitive in English.

(i)   English Adnominal Genitives (simplified structure)

```
                    NP
          _____|_____
         NP                    N
  (↑ADJUNCT)=↓              ↑=↓
  (↓CASE)=GEN              hat
     Boris's
```
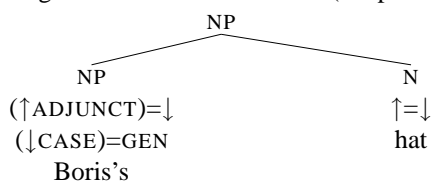
4.16 Sample Analyses

4.16.1 Lexical Entries for Case Markers

The analysis for Urdu case is closely related to Nordlinger's *Constructive Case* idea for Australian languages and incorporate Bashir's (1999) analysis of the ergative (see section 4.1).

(37) a. ne
        ($\uparrow$CASE) = ERG
        (SUBJ$\uparrow$)
          [($\uparrow$SEM-PROP CONTROL) = INT
           $\vee$
          ((SUBJ $\uparrow$) OBJ)
          ((SUBJ $\uparrow$) VFORM) = PERF]

    b. ko
        [($\uparrow$CASE) = ACC
        (OBJ$\uparrow$)
        ($\uparrow$SEM-PROP SPECIFIC) = +
          $\vee$
        ($\uparrow$CASE) = DAT
        (OBJ$_{go}\uparrow$) $\vee$ (SUBJ$_{exp}\uparrow$)
        ($\uparrow$SEM-PROP CONTROL)]

The skeletal f-structures resulting from the entry for *ne* are shown in (38).

(38) a.
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{CASE} & \text{ERG} \end{bmatrix} \\ \text{OBJ} & [\ ] \\ \text{VFORM} & \text{PERF} \end{bmatrix}$$
b.
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{CASE} & \text{ERG} \\ \text{SEM-PROP} & \begin{bmatrix} \text{CONTROL} & \text{INT} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The skeletal f-structures for *ko* are shown in (39).

(39) a.
$$\begin{bmatrix} \text{OBJ} & \begin{bmatrix} \text{CASE} & \text{ACC} \\ \text{SEM-PROP} & \begin{bmatrix} \text{SPECIFIC} & + \end{bmatrix} \end{bmatrix} \end{bmatrix}$$
b.
$$\begin{bmatrix} \text{OBJ}_{go} & \begin{bmatrix} \text{CASE} & \text{DAT} \\ \text{SEM-PROP} & [\ \text{CONTROL}\ ] \end{bmatrix} \end{bmatrix}$$
c.
$$\begin{bmatrix} \text{SUBJ}_{exp} & \begin{bmatrix} \text{CASE} & \text{DAT} \\ \text{SEM-PROP} & [\ \text{CONTROL}\ ] \end{bmatrix} \end{bmatrix}$$

4.16.2 Ergative/Nominative

Intransitive
Consider (40) which shows an alternation with respect to volitionality.

(40) a. ram   k^{h}ãs-a
    Ram.M.Sg.Nom cough-Perf.M.Sg
    'Ram coughed.'

    b. ram=ne   k^{h}ãs-a
    Ram.M.Sg=Erg cough-Perf.M.Sg
    'Ram coughed (purposefully).'

Simplified lexical entries for the subject and verb are shown in (41) (e.g., ϕ-features are omitted). Mapping Theory associates the $[-o]$ agent with a SUBJ.

(41) ram N kʰãs V -a
  (↑PRED)='Ram' (↑PRED)='cough<ag[−o]>' (↑TNS-ASP TENSE) = PAST
  (↑VFORM) = PERF

The entry for the ergative *ne* given in (37a) allows for the two possibilities in (38). However, only the possibility in (38b) will be well formed for (40b), as (38a) requires the presence of an object. The only well formed f-structure resulting from the combination of the noun *ram* with the ergative *ne* is as shown in (42). It has the desired effect that ergative subjects of intransitives entail volitionality.

(42)
$$\left[\text{SUBJ} \begin{array}{ll} \text{PRED} & '\text{Ram}' \\ \text{CASE} & \text{ERG} \\ \text{SEM-PROP} & \left[\text{CONTROL} \quad \text{INT} \right] \end{array} \right]$$

Thus, the case marker itself assigns ergative case to its head noun. In addition, it forces its head noun to be a subject in the f-structure and to be compatible with the given semantic interpretation. If it is incompatible with the other requirements listed in the lexical entry, the result is an ill-formed structure.

The entry for *kʰãs-a* 'cough' does not specify the ergative in its lexical entry. It is therefore free to occur with a nominative subject as well, as in (40a). This nominative is assigned by default ((34b)).

Transitive

With transitives, the ergative is required when the verbal morphology is perfect.

(43) a. ram   gaṛi   cɑla-ta   (hɛ)
      Ram.M.Sg.Nom car.M.Sg.Nom drive-Impf.M.Sg be.Pres.3.Sg
      'Ram drives a car.'
    b. ram=ne   gaṛi   cɑla-yi   (hɛ)
      Ram.M.Sg=Erg car.M.Sg.Nom drive-Perf.M.F.Sg be.Pres.3.Sg
      'Ram has driven a/the car.'

With perfect transitive verbs, only the second disjunct of the ergative entry in (37) comes into play, i.e., in the f-structure in (38a). This disjunct does not include any information about internal or external control, so the semantics of (43b) are left underspecified. The ergative cannot appear in (43a) because the verbal form is not of the right kind for either of the disjuncts to apply (the first disjunct is understood to be constrained to apply only to perfects and infinitives).

4.16.3 Ergative/Dative

Next consider the ergative/dative alternation in the infinitival construction in (44).

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

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

We analyze this construction as an instance of functional control: the verb 'be' acts as a modal which predicates an event of the subject. The subject of the modal controls the PRO subject of the embedded clause. The entry for this form of the verb 'be' is given in (45). Again, the verb makes no direct specifications as to case. It does, however, require that some notion of CONTROL be involved. Thus, dative and ergative subjects are compatible with this entry, but nothing else.

(45)   hɛ $(\uparrow\text{PRED})='$be$<$ag/go$[-$o$]$,Ev$[-$r$]>'$
          $(\uparrow$SUBJ SEM-PROP CONTROL$)$
          $(\uparrow$SUBJ$)=(\uparrow$XCOMP SUBJ$)$
          $(\uparrow$TNS-ASP TENSE$) = $PRES
          $(\uparrow$XCOMP VFORM$) = $INF

With respect to the disjunction in the entry of the ergative marker in (37), only one of the possibilities is viable: because the VFORM of the embedded XCOMP is INF, and because the lexical entry of the verb requires the feature SEM-PROP CONTROL, only the first disjunct applies (i.e, the f-structure in (38b)).[5] The resulting f-structure analysis is shown in (46).[6]

(46)
$$
\begin{bmatrix}
\text{PRED} & '\text{be}<\underline{\ \ },\underline{\ \ }>' \\
\text{TNS-ASP} & \begin{bmatrix} \text{TENSE} & \text{PRES} \end{bmatrix} \\
\text{SUBJ} & \begin{bmatrix} \text{PRED} & '\text{Nadya}' \\ \text{CASE} & \text{ERG} \\ \text{SEM-PROP} & \begin{bmatrix} \text{CONTROL} & \text{INT} \end{bmatrix} \end{bmatrix} \\
\text{XCOMP} & \begin{bmatrix} \text{PRED} & '\text{go}<\underline{\ \ }>' \\ \text{SUBJ} & [\ ] \\ \text{VFORM} & \text{INF} \end{bmatrix}
\end{bmatrix}
$$

Now consider the dative version in (44b). The entry for the dative case in (37b) allows for the three cases in (39). The disjunctions for the marking of object ((OBJ$\uparrow$)) and indirect object ((OBJ$_{go}\uparrow$)) will not result in a well formed analysis for (44) as the argument structure of *hɛ* can only be linked to a subject and a clausal complement by the mapping theory. The only viable option is a dative subject ((SUBJ$\uparrow$)) with the feature SEM-PROP CONTROL. This feature in the f-structure in (47) tells the semantic component that there is something to be said about the control the subject had over the action. Context-dependently this can give rise to a modal interpretation such as □ 'must' (external control), as in (43b), but it need not.

---

[5]The locative oblique 'zoo', not shown in the f-structure, functions like the ergative and dative case markers in that it is associated with information specifying that it must be an OBL$_{loc}$, etc.

[6]Event arguments link to an XCOMP or COMP; see Butt (1995) for discussion.

(47)

$$
\begin{bmatrix}
\text{PRED} & '\text{be}< \rule{0.6em}{0.4pt}, \rule{0.6em}{0.4pt} >' \\
\text{TNS-ASP} & \begin{bmatrix} \text{TENSE} & \text{PRES} \end{bmatrix} \\
\text{SUBJ} & \begin{bmatrix} \text{PRED} & '\text{Nadya}' \\ \text{CASE} & \text{DAT} \\ \text{SEM-PROP} & \begin{bmatrix} \text{CONTROL} \end{bmatrix} \end{bmatrix} \\
\text{XCOMP} & \begin{bmatrix} \text{PRED} & '\text{go}< \rule{0.6em}{0.4pt} >' \\ \text{SUBJ} & [\ \ ] \\ \text{VFORM} & \text{INF} \end{bmatrix}
\end{bmatrix}
$$
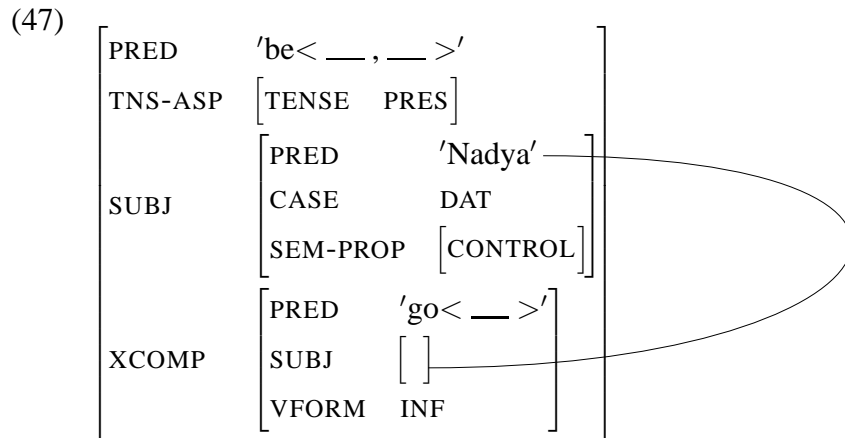
References

Allen, W.S. 1951. A Study in the Analysis of Hindi Sentence-Structure. *Acta Linguistica Hafniensia* 6:68–86.

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

Böhtlingk, Otto. 1839–1840. *Pāṇini's Grammatik*. Delhi: Motilal Banarsidass. Republished in 1998.

Bayer, Josef. 1999. Final complementizers in hybrid languages. Journal of Linguistics 35. 233-271.

Bashir, Elena. 1999. The Urdu and Hindi Ergative Postposition *ne*: Its Changing Role in the Grammar. In *The Yearbook of South Asian Languages and Linguistics*, ed. Rajendra Singh. 11–36. New Delhi: Sage Publications.

Beames, John. 1872–79. *A Comparative Grammar of the Modern Aryan Languages of India*. Delhi: Munshiram Manoharlal. Republished 1966.

Beg, Mirza Khalil A. 1988. *Urdu Grammar: History and Structure*. New Delhi: Bahri Publications.

Bhatt, Rajesh. 2003. Long Distance Agreement in Hindi-Urdu. *Natural Language and Linguistic Theory*.

Bloch, Jules. 1920. *La formation de la langue marathe.* Paris.

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

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Cambridge, Massachusetts: Blackwell.

Butt, Miriam. 1993. Object Specificity and Agreement in Hindi/Urdu. In *Papers from the 29th Regional Meeting of the Chicgo Linguistic Society*, 80–103.

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

Butt, Miriam and Tracy H. King. 1996. Structural Topic and Focus without Movement. In *On-line Proceedings of the First LFG Conference*, Rank Xerox, Grenoble.

    http://www-csli.stanford.edu/publications/

Butt, Miriam. 1998. Constraining Argument Merger Through Aspect. In E. Hinrichs, A. Kathol and T. Nakazawa (eds.) *Complex Predicates in Nonderivational Syntax*. Academic Press.

Butt, Miriam and Tracy Holloway King. 2003. Case Systems: Beyond Structural Distinctions. In *New Perspectives on Case Theories*, ed. Ellen Brandner and Heike Zinsmeister. Stanford, California: CSLI Publications.

Butt, Miriam and Tracy Holloway King. 2005. The Status of Case. In *Clause Structure in South Asian Languages*, ed. Veneeta Dayal and Anoop Mahajan. Dordrecht: Kluwer Academic Publishers.

Butt, Miriam, Tracy Holloway King, Maria-Eugenia Niño, and Frederique Segond. 1999. *A Grammar Writer's Cookbook*. Stanford: CSLI Publications.

Chatterji, Suniti Kumar. 1926. *The Origin and Development of the Bengali Literature, Volume* II. Calcutta: D. Mehra, Rupa & Co (1975 edition).

Davison, Alice. 1999. Ergativity: Functional and Formal Issues. In *Functionalism and Formalism in Linguistics, Volume I: General Papers*, ed. Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan, and Kathleen Wheatley. 177–208. Amsterdam: John Benjamins.

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

Evans, Nicholas. 1985. *A Grammar of Kayardild: With Historical-Comparative Notes on Tangkic*. Berlin: Mouton de Gruyter.

Franks, Steven. 2000. Clitics at the interface. In *Clitic Phenomena in European Languages*, ed. Marcel den Dikken and F. Benkema. Amsterdam: John Benjamins. To Appear.

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

Kellogg, S.H. 1893. *Grammar of the Hindi Language*. Delhi: Munshiram Manoharlal Publishers Pvt. Ltd. Second Edition, reprinted 1990.

Masica, Colin. 1991. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.

Mohanan, Tara. 1988. Causatives in Malayalam. Ms., Stanford University.

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

Payne, John. 1995. Inflecting Prepositions in Indic and Kashmiri. In *Double Case*, ed. Frans Plank. 283–298. London: Academic Press.

Pischel, Richard. 1900. *A Grammar of the Prākrit Languages*. Delhi: Motilal Banarsidass. Republished in 1999, translated by Subhadra Jhā.

McGregor, R.S. 1968. *The Language of Indrajit of Orchā*. Cambridge: Cambridge University Press.

Saksena, Anuradha. 1980. The Affected Agent. *Language* 56(4):813–826.

Saksena, Anuradha. 1982. Case Marking Semantics. *Lingua* 56:335–343.

Saksena, Anuradha. 1982. *Topics in the Analysis of Causatives with an Account of Hindi Paradigms*. Los Angeles: University of California Press.

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

Sharma, Devyani 1999. Discourse clitics and constructive morphology in Hindi. In M. Butt and T.H. King (eds.) *Nominals: Inside and Out*. Stanford, California: CSLI Publications.

Speijer, J. S. 1886. *Sanskrit Syntax*. Delhi: Motilal Banarsidas. Republished 1973.

Tuite, Kevin, Asif Agha, and Randolph Graczyk. 1985. Agentivity, Transitivity, and the Question of Active Typology. In *Papers from the Parasession on Causatives and Agentivity at the 21st Regional Meeting of the Chicago Linguistic Society*, ed. William Eilfort, Paul Kroeber, and Karen Peterson, 252–270.

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

Woolner, Alfred. 1917. *Introduction to Prakrit*. Delhi: Motilal Banarsidass. Republished in 1996.

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

Zaenen, Annie, Joan Maling, and H. Thráinsson. 1985. Case and Gramamtical Functions: The Icelandic Passive. *Natural Language and Linguistic Theory* 3:441–483.