The rise of non-canonical subjects and semantic alignments in Hindi

Annie Montaut

Annie MONTAUT   INALCO/CNRS-SeDyL

The rise of non-canonical subjects and semantic alignments in Hindi

Abstract

An inquiry into the emergence of non-canonical subjects in Hindi may be a first step in evaluating the importance of semantic alignments in the language. The modern data display a cline of subject properties depending on the case and semantic role (section 1). A brief account of the ancient data in Sanskrit (section 2) shows that the most innovative development was that of experiencer subjects, which emerged in early NIA along with inefficient and involuntary actors (section 3). This raise was related to global systemic changes in the language, particularly the development of new case markers (section 4). The last section examines the main factors responsible for this shift, particularly re-analysis, usage and speaker's viewpoint, lexical renewal and contact.

1. What is a non-canonical 'subject'? Its properties in Modern Standard Hindi

It is well known that Hindi as well as other Indian languages display a variety of non-nominative subjects (from Verma & al. 1991 to Bhaskararao & Subbarao 2004), particularly the so-called dative subjects of experiential predicates, whether the predicate involves one or two participants.[1] In no other Indo-Aryan language the variety of such 'non-canonical' subjects is such as in Modern Standard Hindi (hence MSH), or the closely related Urdu and Panjabi. Do these non-canonical subjects, defined in this section on the basis of the criteria developed by Aikhenvald & al. (2001), correspond to semantic alignments as defined in Wichmann (2008)? Out of the five non-nominative basic alignments, only two have been thoroughly analysed, namely the ergative and dative ones, whereas the locative, the genitive and the instrumental ones are most of the time mentioned rather than integrated into the core argument structures of Hindi/Urdu. Ergativity depends on aspect (perfective) and is usually considered a syntactic alignment, but the other non-nominative alignments depend on lexical semantics, including those constrained by verbs which require a complement clause. Experiencers tend to require the dative, possessors the genitive (non contingent possession) or the locative (contingent possession), and inefficient or unable agents, the instrumental. Only five simple verbs require the dative for their first argument, but hundreds of complex predicates do so, and the predicates requiring a genitive or locative, very weakly transitive (meet, lack, be able, contact, etc.) in the hierarchy established by Tsunoda (1985), are always complex. What degree of subjecthood do these oblique arguments exhibit?

1.1. Canonical Subject properties

Apart from the usual coding properties (nominative: unmarked case, control of verb agreement), a canonical subject in MSH occupies the initial position (in discursively unmarked statements), can become the addressee of imperative, be passivized, it controls reflexivization and coreferential deletion with converbs (1a), complement infinitives (1b) or adverbial participles. The Hindi converb (-*kar*) is a used for verbal coordination, but can also convey various dependent meanings.

> (1) a *laṛkā   pen lekar   apnā pepar likhtā hai*
>   boy.M.SG pen take.CVB REFL   paper write PRS.3M.SG
>   'The boy takes his pen and writes his paper'

---

[1] In MSH simple verbs with dative subjects always involve two arguments, and the extension of the pattern to complex predicates (light verbs) with only one argument, corresponding to a semantic role, is explained in 5.2.

(1) b *laṛkā   apnā pepar likhnā śurū   kartā hai*
       boy.M.SG REFL paper write.INF begin do   PRS.3M.SG.
       'The boy takes his pen and starts writing'

Nominative statements also undergo converb reduction: *pen lekar* is the reduced form with converb of *pen letā hai* 'takes his pen', with finite verb in the present. Apart from these core properties, canonical subjects are usually omitted in coordinate clauses.

1.2. Core properties of non-canonical subjects

Although devoid of coding properties since the verb always agrees with the second, unmarked, argument, non-canonical first arguments display most of the core subject properties: [2] position, control of reflexivization, coreferential deletion. However, the ergative argument alone has the complete array of subject properties: it is the antecedent of the reflexive and the deleted subject of the converb *lekar* in (2a), and it can also undergo converb reduction, as shown by the equivalence between the finite verb *liyā* 'took' (2b) and the converb *lekar* (2a):[3]

(2) a *laṛke   ne   pen lekar   apnī kahānī   jaldī   likhī*
       boy.M.SG ERG   pen take.CVB   REFL story.F.SG quickly write.PFV.F.SG

=   b *laṛke   ne   pen liyā   aur   apnī kahānī   jaldī   likhī*
       boy.M.SG ERG   pen take.PFV.M.SG and   REFL story.F.SG quickly write.PFV.F.SG
       'The boy took his pen and quickly wrote his story'

Dative experiencers, similarly devoid of coding properties, cannot undergo passive and imperative transformations. However, they control reflexivization and coreferential deletion with converbs (3), but do not undergo themselves converb reduction (4), except if the main verb is also a verb with experiencer, e.g. if both predicates have the same argument structure DAT-NOM (5):

(3) *yah sočkar   mujhko apne par khīj   huī*
     this think.CVB 1SG.DAT REFL on   exasperation.F.SG be.PFV.F.SG
     'When thinking of that I got crossed at myself'

(4) a *usko   bhūkh   lagī   aur   (vah) khāne lagā*
       3SG.DAT hunger.F.SG touch.PFV.F.SG and   3M.SG eat   start.PFV.M.SG
       'He felt hungry and (he) started eating'

   *b *usko   /vah   bhūkh lagkar   khāne lagā*
       3SG.DAT 3SG   hunger touch.CVB eat   start.PFV.M.SG
       (intended meaning 'He felt hungry and started eating')

(5) *bhāī   ko   čiṭṭhī milkar   xuśī   huī*
     brother.M.SG DAT letter get.CVB joy.F.SG be.PFV.F.SG
     'Brother was happy getting the letter'

The two locatives ('in', 'at') and the genitive can be used for real possession as in (9) or as arguments of weakly transitive complex predicates or even as complement of body parts, as in (8a) in the periphrastic expression of a mental or physiological state. They always display one subject property less than experiencers. While they control coreferential deletion (of the infinitive's subject in (6) and (7) or of converb's subject in (8a)), they themselves never undergo

---

[2] The second argument may be the nominal component of a complex predicate, as *bhūkh* in (3b). *Cf.* 5.2.

[3] In the following examples, the sign = stands for' equivalent to', the signs * and ? indicate respectively unacceptability and weak acceptability of the statement, ∅ indicates null anaphora, / indicates alternative form. The gloss is simplified for verbs in the present, imperfect or present perfect, where morphological features are distributed on the participle and the tense auxiliary, and often redundant.

converb reduction: (8b) and (9b) are ungrammatical[4].

(6) *usmẽ is tarah kā kām karne kī himmat kabhī nahī̃ huī*
 3S.LOC this way of work do GEN courage ever NEG be.PFV.3F.SG
 'He never had the courage to do this kind of work'

(7) *merā apnī bahan se milne dillī jāne kā irādā thā*
. 1S.GEN REFL sister COM meet Delhi go of intention.M.SG be.PST.M.SG
 'I intended to visit my sister in Delhi'

(8) a. *yah sunkar sītā kī ā̃khẽ bhar aī̃*
 this hear.CVB Sita GEN eye.F.PL fill come.PVF.F.PL
 'When hearing this, Sita was about to cry'

 *b. *Sītā (kī) ā̃khẽ bharkar cillāī...*
 Sita (GEN) eyes fill.CVB shout.PFV.F.SG
 (Intended meaning): 'Sita was about to cry, and shouted…'

(9) a *uske che baččhe haĩ aur uske pās samay kā abhāv rahtā hai*
 3SG.GEN six children are and 3SG LOC time of lack stay PRST.3SG
 'She has six children and (*having six children,) she is always short of time'

 *b *(uske) che baččhe hokar uske pās samay kā abhāv rahtā hai*
 3SG.GEN six children be.CVB 3SG LOC time of lack stay PRST.3SG
 'She has six children and (*having six children,) she is always short of time'

Actors, either inadvertent (10a) or, in negative contexts, ineffective (11a) with a meaning close to the incapacitative passive (11b), are not a homogeneous class: the inadvertent agent seems to lack control over reflexivization and converb (10b), whereas the ineffective actor of negative intransitive verbs controls converb reduction (10a), as does the inhibited agent of modal passives (11b):

(10) a *mujhse (merā kimtī) gilās ṭūṭ gayā*
 1SG.INS POSS.1 costly glass broken went
 'I broke the (my precious) glass by inadvertence'

(10) *?b *mujhse safāī karke apnā sab-se kimtī gilās ṭūṭ gayā*
 1SG.INS cleanness do.CVB REFL most costly glass broken went
 'I broke my most precious glass by inadvertence in doing the cleaning'

(11) a *pūrī kośiś karke bhī mujhse darvāzā nahī̃ khul rahā hai*
 full effort do.CVB even 1SG.INS door NEG be.opened PROG PRS.3SG
 'Although I try my best, I can't manage to open the door'

 b *pūrī kośiś karke bhī mujhse yahã̄ baiṭhā nahī̃ gayā*
 full effort do.CVB even 1S.INS here sit NEG PASSIVE.PFV
 'Although I tried my best, I could not bring myself to sit here'

1.3. Non core properties

Subjects, as other NPs, can be omitted in Hindi (zero anaphora), but not all non-canonical subjects are freely omissible. In clause coordination the pronoun corefering with an ergative agent (12) is preferably omitted, in contrast with a possessor, which is never deleted (9a, above, 14), and a dative experiencer, which can optionally be deleted (4a above, 13):

---

[4] The Hindi genitive is always also marked for gender and number: it is an adjective-like "postposition" (*kā, kī, ke* derive from the past participle from verb root *kar*) which agrees with the following NP.

(12) *usne   gilās mez par rakhā   aur (?vah) bāhar gayā / (?usne) kahā*
     3SG.ERG glass table on put.PFV.MS and (3SG) outside went 3SG.ERG said
     'He put the glass on the table and went out / said'

(13) *us vaqt mujhe   tumse   irṣyā   thī   magar (maĩ) nahĩ jāntī thī*
     that time 1SG.DAT 2.from jealousy was but   (1SG.) NEG know PST.F.SG
     'At that time I was jealous of you, but (I) did not know'

(14) *sītā kī   ãkhẽ   bhar aĩ   aur (*∅) vah   bhāg   gaī*
     Sita GEN eyes   fill come.PFV.F.PL and   3SG   run.away go.PFV.F.SG
     'Sita was about to cry and she ran away'

Significantly, shifts in position, particularly the topicalizing fronting of the second argument, do not interfere in the referential properties of ergative agents or nominative subjects, but it does for other obliques. A topicalized unmarked stimulus or theme may disobey the reflexive constraint: 'Chazi', the theme in series (15), requires the reflexive only when the experiencer comes first in the sequence (15b) but not when it itself is topicalized and occupies the initial position (15a). This is never the case with a transitive verb with ergative 'agent': in (16) the close synonym *pānā* 'find/get', with NOM-ACC argument structure, always requires the reflexive *apnā* even with a focalized theme. One may obviously doubt that the dative experiencer of (15a) is at all a subject:

(15) a *hamārī   čāzī   hamko   mil gaī hai*
       1PL.POSS   Chazi 1PL.DAT   find go PRF
       'Our Chazi we got her back / We have found (back) our Chazi' (girl's name)

     b *hamko   apnī čāzī   mil   gaī hai*
       1PL   REFL Chazi find   go PRF
       'We have found (back) our Chazi' (girl's name)

(16) *a *hamārī   čāzī   ko   hamne   pāyā hai*
       POSS.1PL Chazi ACC 1P.ERG   get PRF
       'We have found/got our Chazi'

     b *apnī čāzī   ko   hamne   pāyā hai*
       REFL Chazi ACC 1PL.ERG   get PRF
       'Our Chazi, we found her'

1.4. A semantically consistent class?

To sum up, the syntactic and discursive properties of non-nominative 'subjects' show that only the ergative agent is endowed with full subject behavioural properties, dative exhibits less, and genitives and locatives even less. This cline is distinct from Onishi's (2001b) conclusions for Bengali, where locative (in relation with definiteness) ranks higher than genitive (which encodes experiencers and possessors in Bengali), themselves higher than other obliques. Definiteness does not trigger non-canonical alignment in Hindi, but aspect does, and possessors are case marked distinctly from experiencers, and distinctly whether they refer to contingent or non-contingent possession. But the predicates requiring a non-canonical subject are comparable, including deontic verbs (17), either invariable (*čāhie* 'should') or inflected (*paṛ* 'fall, have to'). Contrary to the above predicates, such deontics require a complement clause, as is the case with expressions such as 'be easy/difficult', 'take time' (18):

(17) *āpko   ārām karnā čāhie / paṛegā*
     2H DAT rest make   should / have.to.FUT.3M.SG
     'You should /will have to take rest.'

(18) *mujhe   āne   mẽ   dikkat   huī,   do   ghaṇṭe   lage*
      1SG.DAT come in   difficulty.F.SG be.PFV.F.SG two hour.M.PL touch.PFV.M.PL
      'I had problems for coming, it took (me) two hours'

Although transitive verbs are usually agentive (2), which means that NOM-ACC alignment can be viewed as a semantic one, (16) shows that the NOM (or ERG) encoding does not systematically require the expected semantic features of agents (volition, control, capacity of initiating a process) and only display conscious awareness (*cf.* 5.3). However, the lack of only one of these features in action processes results in the shift from nominative or ergative to instrumental alignment which encodes participants having no control but acting volitionally ([-control], [+volition]) as in (11a). The argument structure of a predicate now largely, yet not totally, reflects the semantic roles of participants, with the main argument case marked according to its position in the hierarchy of roles. The notion of semantic alignments suggested in Wichmann (2008) then seems more appropriate than that of syntactic alignments. NOM-ACC and ERG-ACC correspond to the higher part of Tsunoda's transitive hierarchy (1985), immediately followed by the DAT-NOM alignment, then GEN/LOC– NOM, and finally the INS-NOM involving inadvertent actors (10). The fact that other INS deficient agents maintain more subject properties (11) shows that a similar case marking may cover several related roles of main arguments, and that the lack of consciousness and volition makes the actor farther from a subject than the simple lack of efficiency and control. This suggests that a correct description of argument structure in Hindi should rely on finer grained distinctions that those usually used. Although case markers, particularly the unmarked case (nominative /unmarked accusative), may be syntactic, they most of the time encode semantic roles, yet a given case marker (dative, or instrumental) may encode quite distinct roles according to the semantic class and morphological structure of the predicate. This suggests that the language is evolving towards a more semantic type of alignments, while retaining to some degree syntactic alignments, an evolution confirmed by a comparison to earlier stages from Sanskrit to premodern Hindi.

2. Continuity from Sanskrit: non-nominative 'agents' (instrumental/genitive)

Can the present system in Hindi be considered a development of the ancestral language Sanskrit, which allowed certain cases to behave as main arguments and display the syntactic properties (mainly control of coreferential deletion) of the first case or *kartā* 'agent'?

2.1. Instrumental subjects
It is well known that the nominal sentence in classical Sanskrit with a verbal adjective (passive past participle) as the predicate and an agent in the instrumental (or genitive particularly for pronoun) is the origin of the modern ergative alignment in Indo-Aryan such as (2). The instrumental agent of a passive participle used as a predicate in (19), adapted from Hock (1991a: 59), controls the coreferential deletion of the subject of converb *gatvā* 'go'. Even with intransitive past participles like 'come' in (20), from the *Mahabharata* (III.199.9) commented in Bloch 1906: 36), a genitive agent controls coreferential deletion of the subject of complement infinitive 'ask':[5]

(19) *tatra gatvā   janai   rājā   dṛṣṭaḥ*
      here go.CVB people.INSTR.M.PL king.NOM.M.SG see.PPP.NOM.M.SG
      lit. 'having gone there (the people/*the king), the king was seen by the people'

---
[5] More examples in Montaut (*forthc.*), which argues that similar 'pre-ergative' structure with verbal adjective could be found with intransitive past verbs and in modal future too. In modern Hindi, agents of finite passive verbs also control coreference as the do in (14), in the same way they do in Sanskrit as shown by Hock (1991a: 59).

'Having gone there, the people saw the king'

(20) *asmākam   abhipretaṃ   bhavantaṃ kiṃcid artham   abhipraṣṭum*
     1PL.GEN   come.PPP.NOM.N.SG respected   some   thing ACC.N.SG ask.INF
     'We have come here in order, O respected, to ask you something'

A similar alignment, also in nominal sentences with a passive modal participle as the predicate (-*tavya* 'to be V.*ed*, V-*able*'), developed for obligation, with the instrumental similarly controlling coreferential deletion, even when omitted. Thus, the adjectival predicate agrees with the patient in (21), from early Prakrit (Ashoka's first Edict), while the omitted instrumental agent is the subject of the converb 'kill':

(21) *na kimci jīvam   arābhitpā prajuhitavyam   na ca samājo   kattavyo*
     NEG some living.N.SG kill.CVB   sacrifice.OVA.N.SG NEGand assembly.M.SG make.OVA.M.SG
     '(One) should not sacrifice by killing a leaving creature nor hold asssembly'

This pattern was later to become a modal future, then a future, with intransitive and transitive verbs in Eastern languages, and is still traceable in modal Marathi statements (Montaut, forthc.).

2.2. Genitive subjects

Sanskrit also displays genitive subjects which control coreference, as shown by Hock (1991a), Deshpande (1991), Dahl (forthc.). The Vedic example (22) is from Hock (1991a: 63):

(22) *svasti vācayitvā   svasti ha eṣām   bhavati*
     svasti recite.CAUS.CVB   svasti PRT 3PL.GEN be.PRS.3SG
     'Having caused 'svasti' to be recited, 'svasti' is theirs= they have 'svasti' '

The reason given for this syntactic prominence is that such a genitive, like the instrumental agent, is a semantic *kartā* 'agent' and therefore can be syntactically promoted to a *kartā*. A *kartā* is not necessarily endowed with agentivity but it is the first case role because it represents the entity most prevalent in the action. Depending on the other roles and their contribution to the action, it may be a simple possessor, providing there is no higher NP in the hierarchy of action-participation.[6] For this reason, possessors of body parts in physiological (or psychological) transient states may also display subject properties as in (23), from *Vetikā, 15.37.38*, in Hock 1991a: 64), which can be considered as a direct source for Hindi (14):

(23) *susnātaṃ puruṣaṃ dṛṣtvā   klidyante   yonayaḥ   strīṇāṃ*
     well.bathed man.ACC see.CVB   get.wet.3PL vaginae.NOM.PL woman.GEN.PL
     '(Women) having seen a well-bathed man, women's vaginae get wet'

Other non subject (or 'agent': *kartā*) cases, like locative and accusative may also be promoted if they refer to the main contributor to the action. Yet no dative, the expected source for modern experiencers, is found in Sanskrit to be endowed with subject properties.

2.3. Dative: a beneficiary and a target, not an experiencer subject

It has been argued on the basis of (24), from Hook (1976), that dative subjects occur in Sanskrit:

(24) *śrutvā   idam upākhyānam ... anyan na rocate   [tasmai]*
     hear.CVB   this story   other NEG please.PRS.3SG[3.SG.DAT]
     'Having heard this story, another one doesn't please him/he does like another (story)'

---

[6] Hock's argument: non agent possessors tend to be co-referred by the pronoun and not the reflexive (1991a: 65).

But Hock (1991a, 1991b), Cardona (1991) and Deshpande (1991) have shown that such control of coreference is not systematically displayed by the dative beneficiary: statements like (24), from classical Sanskrit (*Mahabharata* 1.2.2. 36, quoted from Hook by Hock 1991a), are treated as exceptions like a few other cases of sloppy reference illustrated by Hock.

However, the verbs *rucati/rocate* has, according to Deshpande (1991), evolved by the later stages of Sanskrit in such a way that it may be considered as a possible source for the future development of dative subjects. *Rucati* 'shine' was initially a one place predicate, which allowed the 'observer' to surface in the dative (beneficiary) but not, because of its very weak participation in the action, to be promoted to the main syntactic rank. When the meaning shifted to 'appeal, interest', the dative came to refer to a person as a target still too dependant on the stimulus to be an autonomous participant and allow promotion to the main syntactic rank. The semantic prevalence of the experiencer later emerged out of the meaning 'appeal, interest' and the verb semantic shifted to 'like', while the experiencer was encoded either as a genitive (*mama ruci* 'my liking') or a nominative with the causative base (*aham rocayami* 'I like').[7] The dative remains a target, or a beneficiary, with weaker intensity of participation and lower autonomy, a reason why, like the ablative (point of departure), it can never rank first syntactically (Deshpande 1991: 154-7).

Autonomy, the essential feature of agents in Panini's grammar, is then required for the promotion of a NP to the status of syntactic subject. Such emphasis on the independent contribution in performing the 'action' is related to a theory of language for which action is *par excellence* the sacrificial rite, therefore the action clause serves as a model for all predicative relations. The prominence of experiencer is not acquired before late Sanskrit, and similarly, nominal sentences where an adjective like 'easy, difficult, available' is the predicate, do not seem to promote their 'unnecessary' argument to *kartā*. Significantly, later religious texts dealing with heterodox faiths such as Buddhism or Jainism, no longer focused on sacrificial action, display clear occurrences of experiencers controlling coreference, such as (25) from *Uttarajjhaya* 17, echoing of Hindi (18):

(25) *labdhunavi   āryattanam   ahīna-pancendiyayā   dullahā*
    get.CVB.even   aryanness.ACC   intact-five.sense.NOM.M.PL   hard.to.get.NOM.M.PL
    'Even having obtained Aryanness, it is hard to keep one's five senses intact'

3. The rise of experiencer 'subjects' in early NIA

3.1. The problems in constructing the data

Example (25) illustrates the frustrating situation for a proper analysis when the term under scrutiny is an omitted NP in elliptic clauses, a situation which prevails in what is called 'medieval Hindi', or early New Indo-Aryan, from 13th to 17th century. Besides, and this is part of addressing a non-standardized language, there is considerable variation in the morphological coding of NPs, with postpositions only starting being used in replacement of the old case endings. Most of these case endings got eroded, except for a –*i/ī/ĭ/he/hi* locative case, used for instance to encode agents of past transitive verbs, in continuity of (19). In (26) from Kabir (14th c.), one of the most reliable testimonies of what will be later called Hindi, copula is omitted in the first clause (Maya magician), and in the second both subject and object of -*i* ending converb *mãgi* 'ask' are omitted.

---

[7] An experiencer more involved in the process, with verbs such as *spṛhayati* 'wants, longs for', is promoted to the main syntactic rank with its theme in the accusative while the instrumental of the theme indicates a lesser degree of involvement: *Devadattah puṣpāṇi spṛhayati* [Devadatta.NOM.M.SG flower.ACC.N.PL longs for.PRES.3SG], from the *Sidhantakaumudi* E23, commented by (Deshpande (1991: 153).

Hence we can read the main verb *milai* as transitive, with omitted nominative experiencer:[8]

(26) *māyā   mohinī,   m̃ãgi   milai   na   hāthi* (16.9)
    illusion charmer ask.CVB get.PRS.3SG NEG hand.LOC
    'Maya is a magician, asking, it is not found at hand (even when asking you do not get it)'

But when both arguments are expressed with a clearly analysable case marker, a rare occurrence, the argument structure appears to be OBL-NOM: the animate experiencer is in the first position, either in the oblique case, *tāhi* 'to them' in (27), or in the postpositional dative *premī̃ kaũ* in (28). The verb always agrees with the other participant.

(27) *kām   krodh   triṣṇā   tajai   tāhi   milai   bhagvān* (264)
    desire anger cupidity renounce.PRS.3SG 3SG.OBL find.PRS.3SG Lord
    'The one who gives up desire, anger and greed, will meet <u>God</u>' (*lit.* 'him is found God')

(28) *premī̃ kaũ premī̃ milai,   tab sab biṣ   amrt   hoi* (43.12)
    Lover DAT lover meet/find.PRS.3SG then all poison nectar be.PRS.3SG
    'When the lover finds the lover, all poison becomes nectar'

The verb *mil* accounts for a consistent part of the data supposedly involving a non-nominative subject, a situation resembling Old Islandic verbs 'seem' and 'like/please' (Barddal 2003), and it is clearly not by chance that a reciprocal verb with its two arguments ranking equal in the hierarchy of roles, is the most frequent with the construction under analysis. We can assume that *mil* is already a non-nominative subject verb, by far the most recurrent verb so constructed in Kabir.[9]

3.2. A semantically consistent class

3.2.1. Classes of predicates
The four classes of predicates identified across languages by Onishi (2001a) as requiring non-canonical subjects are attested in the earlier stages of early NIA: physiological and psychological states (29), perception and liking verbs (30), verbs of lacking and possession (31), verbs of happening expressing uncontrolled events (32), which are probably an extension to intransitive verbs of the modal incapacitative passive (33):

(29) *anaratā   sukh sovaṇã̃,   ratai   nīnd   na   āī* (29.4)
    non.lover.NOM.PL happily sleep.PRS.M.PL lover.OBL sleep.F.SG NEG come.PFV.F.SG
    'Those who do not love sleep peacefully, the one who loves cannot find sleep'

(30) *kāmī̃   amī   na   bhavai* (20.19)
    lascivious nectar NEG please.PRS.3SG
    'Lascivious men do not like nectar'

(31) *kartā   kere bahut guṇ,   auguṇ   koī nāhī̃* (56.3)
    Creator GEN many qualities, bad.qualities some NEG

---

[8] The root *mil-*, "encounter", noun and verb, is of Sanskrit origin, and was initially a nominative verb (Callevert's dictionary for Raidas lists it as transitive). Bracketed numbers after the example stand for the number of the stanza.
[9] However its close synonym *pānā* "obtain, find", always remained transitive, and agrees with the nominative subject (*sukh pāũ̃* [happiness get.PRS.1SG] "I find pleasure"). *Cf. infra.*

'The Creator has many qualities, no defects at all'

(32) *sãĩ   sũn sab hota hai,   bande thaĩ kuch   nāhĩ* (38.12)
      Lord INS   all be   PRS.3 slave INS something   NEG
      'The Lord can [do] everything, the slave is not able of [doing] anything'

(33) *mau   paĩ   lakhyā na jāī*   (38-6)
      1SG   INS   perceive NEG PASSIVE
      'I am totally unable of perceiving [it]'

Particularly frequent, yet not as much as the omnipresent *mil* (a verb of possession 'get', cognition 'find/discover' or perception 'meet'), is the construction involving a dative first argument and a main verb expressing time duration (days went, time passed) with an embedded non finite clause (present participle) expressing an action (35). The verb *lag* 'touch, be connected', today one of the four simple verbs requiring dative subject in MSH in the meaning 'seem', can also be used for expressing time duration of a process whose subject is in the dative (34):

(34) *sãĩ   kau siyat   mās   das lāge* (3.4)
      Swami/Lord DAT sewing   months ten be.connected.PFV.3PL
      'It took ten months for the Lord to sow' (*lit*. the Lord took ten months sewing)

(35) *kahat   sunat   sab din gae* (55.6)
      saying hearing   all days go.PFV.M.PL
      'Speaking and hearing, all days went (you lost your time)',

The exact nature of the semantic role of the dative noun in (34) is difficult to specify, and even more abstract than the dative of "have difficulties": it is only the locus of the process 'time pass in doing X'.

3.2.2. Evidence for subjecthood in ancient Hindi
Nominative subjects regularly control coreferential deletion in early Hindi/Urdu (Kabir): in (36) the nominative omitted subject of the main verb *calyā* 'went' controls coreferential deletion of the subject of the converb *dekhi* 'see' as well as the subject of the present reduplicated participle *hansat* 'laughing':

(36) *dokh   parāye dekhi   kari,   calyã   hãsat   hãsat* (54.1)
      defect other see.CVB   do.CVB   go.PFV.M.PL laughing laughing
      'Looking at other's defects they went by, laughing'

Contrary to nominatives, experiencers do not allow passive and imperative transformation, but, even when omitted, they control coreferential deletion with adverbial present participles such as *siyat* 'sewing' in (34), *kahat sunat* 'saying hearing' in (35) and with the converb *mili* 'having found' in (37):

(37*) *satgur   mili   parcā bhayā* (54.2)
      true.guru find.CVB   vision be.PFV
      'Having found the true guide, I had the vision'

They also control reflexivization: in (38), which comes right after (36) in Kabir's text, the reflexive is coreferent with the omitted experiencer of *āvai* 'comes' (to knowledge):

(38) *apnāi cyanti   na āvaī jinkī   ādi   na ant* (54.1)
      REFL consciousness NEG come.3SG REL.GEN beginning NEG end
      'They have no consciousness of their own (defects), which have no beginning nor end'

Similarly the (omitted) genitive possessor of body parts controls coreference when it occurs in periphrastic locutions used for expressing physical states, such as 'tongue aching' in (39) with the omitted possessor of the 'tongue' as the syntactic subject (*cf.* MSH (8), Sk (23) *supra.*

(39) *jibhṛiyã   chālā   pariā   nām pukāri pukāri* (16.2)
      tongue.F.OBL swell.M.PL come.PFV.M.PL name call.CVB call.CVB
      'I got swells in my tongue by calling [His] name'

Further evolution of the language shows the development of such structures as (34) and (39), extremely frequent some two centuries later in the Hindi *Ramayana* of Tulsidas (16th c.), from which are the two following examples:[10] (40) shows that the body part possessor (omitted 'of me') controls coreferencial deletion with converb, (41) shows that the omitted experiencer controls coreference with present-participle used adverbially:

(40) *tohi   dekhi   sital bhai   čhātī* (35)
      2SG.ACC see.CVB cold be.PFV breast
      'Seeing you my heart got cool'

(41) *sunat   bacan   ujā   ati   krodha* (135.3)
      hearing speech rise.PFV much anger
      'As soon as he heard these words, [Narada] was filled with rage'

The reflexive anaphora is also controlled by the experiencer: in (42), *nij*, the Eastern variant of the reflexive, is coreferential with the dative *mohi* 'to me', the experiencer of the verb 'seem' (*lāg*). In the first clause of the same verse in Tulsidas's *Ramayana*, the omitted dative subject of the adjectival predicate 'reachable' corefers with the deleted subject of the present participle 'giving' *det*:

(42) *tumhahi det   ati sugam   gosãi,* (133)
      2.DAT   giving very well.reachable God
      agam   lāg   mohi   nij kripnāī
      un.reachable seems 1SG.DAT REFL narrowness
      'It is easy for you to grant [it], God, due to my wretchedness it seems to me so hard'

3.3. The main shifts in early NIA: towards a new case marking?

The argument structure is clearly identifiable in OIA and MSH, both in their own way languages with a clear case morphology (flectional and postpositional respectively). Early NIA, in contrast, displays considerable opacity in case morphology (*cf.* 3.1) and the argument structure often remains unclear. In Kabir, non-agents of two place predicates are more frequent in the oblique or locative, a case which has then become extremely syncretic, to such an extent that we find it for animate objects, experiencers, locatives, ablatives, and transitive agents in pre-ergative sentences. Besides, the same postposition may be used for two or three distinct roles, and two distinct postpositions may alternate for the same role (*sūn, thaĩ, paĩ* in (32-33) for instrumental).

---

[10] Numbers after the examples refer to the stanza number in all standard editions.

Postpositions are, as expected, more frequent in the later texts (Tulsidas). Ellipsis also complicates matters, particularly with predicates such as 'difficult to obtain', 'easy to reach', which appear at that period, with an animate argument controlling corefence (34, 42), but most of the time non expressed (35), in continuity with late Middle Indian (25). Such adjectival predicates, requiring non-nominative main arguments, may be considered as non-canonically marked, following Haspelmath's conception (2001: 65-6).[11] Following Onishi (2001b), other complex predicates with stative light verbs are to be considered non-canonically marked.

Whatever the predicate, the main innovation in terms or roles is the promotion of the dative experiencer to the main syntactical rank. The more striking innovation, however, is the limitation of the transitive argument structure, which tends to be restricted to predicates involving agent and patient roles. Exceptions are those still observable in MSH (transitive verbs like 'know', 'see', 'obtain'), but in MSH such verbs all have a counterpart with non-canonical subject, which was not yet the case in early NIA or at least not systematically (*cf.* 5.3). The class of demoted agents in particular (incapacitative intransitive and passives) is part of this general shift from purely syntactic transitivity to a semantically conditioned transitivity (requiring agent and patient roles).

The verb *mil* is a good illustration of the shift from nominative to dative. It was originally a nominative verb meaning 'meet, merge' in Sanskrit, with a second argument in various cases (accusative, genitive, instrumental) but not dative. In the early NIA texts, its argument structure is mainly DAT – NOM, with the main argument (the one who gets) in the oblique or dative (27, 28). But we also find in Kabir statements like (43), with the old argument structure, and (44), which treats the nominative as the most crucial for topic continuity, contrary to the new DAT/Obl-NOM argument structure where the DAT/Obl is the discursive pivot:

(43)   *tan   mati maĩ mili   gayā*   (12/14)
       body.M.SG earth LOC mix/meet.CVB go.PFV.M.SG
       'The body got mixed with/into the earth'

(44)   *Kabīr sāĩ̃   milhĩge   pūchihĩge   kusalāt*   (56.1)
       Kabir Lord[HON.] meet.FUT.M.PL ask.FUT.M.PL   state
       'Kabir, [you] will meet the Lord, [He] will ask about your condition'

Yet, the verb *mil* has most of the time the argument structure DAT/Obl – NOM, where DAT is the experiencer, first term in the sequence, and controller of coreferential deletion.

   Another, less frequent verb, which underwent the same shift from NOM/theme – DAT/beneficiary to DAT or OBL / experiencer – NOM/theme, is *bhānā* 'please' (30). *Bhānā* comes from the Sanskrit root BHĀ 'shine', but in the 14th c. is clearly re-interpreted as meaning 'please' and re-categorized as a non-nominative predicate. This means that the ex-beneficiary (X shine for Y), the 'unnecessary' argument, has become a subject-experiencer. This evolution strongly echoes that of verbs *rocate/rocayati/rucayati* as described by Cardona (1991) and Deshpande (1991). The meaning and argument structure of *bhānā* in MSH are the same as in Kabir, but the verb has been marginalized by the various predicates with light verbs which will proliferate in the next centuries (*cf. infra*). Interestingly, *pānā*, from the NOM-ACC Sk verb *prāpayati* 'obtain, find' did not change its argument structure because it already had its human experiencer in the first rank (nominative).

4. Systemic changes and the development of new case markers

---

[11] Haspelmath considers the equivalents of 'have' predicates, the reflexive, middle or anticausative verbs as non-canonically marked, as well as adjectival or adverbial predicates.

If we compare this state of affairs with the picture displayed by the modern language, which tremendously enriched its case system (Kulikov 2009: 455-6), early NIA appears as a transitional system. Typologically, this transitional phase did not simply prepare the shift from nominative to ergative alignment, but also to dative and other oblique case alignments, with a marginalization of the transitive NOM-ACC alignment. Such a shift suggests that the argument structure tends to encode semantic roles rather than grammatical relations. Yet we cannot view early NIA as a role dominated language since the specificity of the numerous non agentive roles is not precisely encoded, with the syncretic oblique form prevailing on postpositions, which only start developing. Since their origin is usually poorly treated and may generate wrong generalisations, I will go in some detail below.[12]

4.1. Origin of postpositions as case markers

In the 14-15th century, postpositions are both non systematic and not fixed in a given meaning. What is striking when observing the emergence of the new case markers in the various dialects of Hindi[13] is both the diversity of forms grammaticalized as case markers and the low differentiation of the semantic content of such forms.

The postposition *kau* occurring in (28), (34), with the alternate forms *kaũ, kũ̃, kū* (in MSH *ko* marking DAT/ACC/EXP), is from the Sanskrit noun *kakṣa* 'side, place' (a noun maintained in Hindi with the meaning 'room, classroom'). This clearly locative origin is still visible in the alternate forms *kakh, kākh, kakhã̃* which occur in the Pahari languages (North). It is not related with the *kā/kī* marking the genitive, which agrees in gender and number with the possessed NP like an adjective, because it is derived from the past participle of the verb *kar* 'do' (*kʃa, kia, ka*), with such variants as *ro/rā/rī* in Western speeches or *ker/ kera* in the East (*cf*. ex. 31). Other frequent markers for dative in the Pahari languages are *saṇī (haṇī*, with the regular aspiration of sibilants), from the Sk noun *sanga* 'society, company', then 'with', or *tai/tāĩ* which comes from *tāvati*, a locative form of the resumptive indefinite *tāvat* (*tāvahĩ, tāmhĩ *taaĩ, *tannĩ, tāĩ*) 'so long, so far, up to, till', and *khuṇī,* from the Sk noun 'ear' (*karṇa* in the locative: *karṇe*) 'at the hear': 'aside, close to'. It is clear that the markers for dative (used for experiencer, beneficiary and accusative) have originally no allative meaning but are simply locators with almost no semantic content except closeness. Similarly *lā, laĩ, lai*, rare in Hindi dialects (but frequent in other IA languages) comes from a verb meaning 'touch', 'be connected'.

The postpositions today considered as ablative/instrumental also share this vague relational meaning without concrete directionality. Hindi *se* (with variants such as *sī, si*) still bears traces of this original bleak meaning since it occurs for ablative, cause, instrument, symmetric relation ('marry', 'be in contact', 'quarrel') and quasi beneficiary ('speak to', 'say to'). Although some scholars trace its origin to *santo/sunto* from Sk AS 'be, existing' (Hoernle 1880: § 376, discussed in Tiwari 1955: 351), or *sanga* (Kellogg), it is now admitted that is comes from *sama* 'equal, even' (Tiwari 1955, Chatak 1966) in the instrumental form *sam.ena* (*saẽ > saĩ > sẽ*). How can a meaning 'equal' evolve into a meaning 'by/from/since'? Part of the answer goes back to Sanskrit, where *samena* already adverbialized in the meaning of 'equally, in parallel'. What may appear as a very bizarre semantic shift amounts in fact to the selection of some of the meanings already proliferating in Sanskrit and now used for relating, again in a spatial vague way, the predication to some participant.

Te/tī, also used for the ablative in alternation with *se/sī*, is derived from a semantically even more empty basis, *hunti < hontai < hontako*, a present participle of the verb 'be' (Sk *bhavati*), meaning

---

[12] Bubenik for instance derives the Awadhi *ker* from Sk *kṛta* "done, made", but Hindi *kā, ke* from Skt *kārya* "to be done" (2006: 122f) without quoting sources, and against Tiwari 1955 or Chatterji 1926.

[13] Hindi is said to accommodate 331 various dialects, among which a dozen are really distinct languages.

'being, staying'. This is not the only example of participles of the existential verb being grammaticized into case markers since the very common *thaï* is also considered to be derived from the past participle of STHĀ 'stand, stay' (cf. the copula, and tense auxiliary for past, in MSH *thā/the/thī*).[14] This explains that one can find case markers which can be used for apparently opposite concrete relations like agent, object, beneficiary and instrument as is the case in Bangru, a northern dialect close to Panjabi and Hindi (from Tiwari 1955: 177):

(45)   *rupay   tī   us-tī   le   lo*
        money ACC 3SG-ABL take take.IMPER
        'Take the money from him'

The same ambivalence occurs with the *nae, nai, ne/ni* form in Western languages, also derived from the locative of *karṇa* 'ear': Panjabi is known for having the *na* morph for both ergative and accusative/dative, although in two distinct reflexes (*ne/nṹ*), whereas Rajasthani has *nai/ne*. One Southern variety of Bangaru uses the same form *nae* for ERG, DAT/ACC, but also INS.

A study of the other ergative markers, *le/la*, from *lag* 'touch, be connected', would similarly show that the form is semantically rather empty and has many uses, according to the languages which use it, from DAT/ACC to INS/ERG (more details in Montaut forthcoming).

4.2. Typological correlations

What is striking in the history of IA case development is that practically all cases etymologically simply indicate that the predication is related to a given participant, without specifying the semantics of the relation. Keeping this in mind, it no longer appears as a contradiction that destination and source or cause may use identical or similar markers (even in the same language, although exceptionally, as in Bangru). This suggests that they were primarily used not as specific case markers, but as simple (vague) locators of the predication. In this way, in early NIA, emerging 'case markers' did not behave very differently from the remaining old case endings, phonologically so much eroded that they were reduced to the syncretic marking of all kinds of obliques.

Later on, case markers got stabilized in standardized languages (Hindi, Marathi, Panjabi) and acquired a distinct semantic meaning, sometimes syncretic (like the so-called Hindi 'instrumental'), although they all stem from a semantically similar origin (localisation). Standard languages today display a definite set of markers, different for each language. But when we look at the morphological material used for case marking across various related languages and dialects, we find practically all of the markers used in practically all functions (yet rarely in the same language), with different extensions according to the various languages: for instance *na/nai* in Rajasthani extended to DAT/ACC and ERG, whereas in Garhwali, *na/le* extended to ERG, INS and cause roles, but not DAT/ACC. The proliferation of forms is associated with the proliferation of meanings of the etymon, making merging and analogy easy.[15] The great number (up to ten) of markers in a given dialect for a given case, some unrelated, is understandable since they all derive from vaguely localizing words.

This suggests a transitional phase in the evolution of the IA language when all non direct arguments (SO, mostly mapping agent and patient) were simply peripherized in relation to the predication. Additional marking appeared to compensate for the degradation of case endings, initially as a continuation of the Apabhramsha stage (early second millenium) with a very weak semantic load, suggesting that the predicative nucleus was only loosely related to other participants, whatever their role, as simple locators for the predication. What later emerged as the most important strategy for presenting information was the encoding of both

---

[14] *Thaï* and *ti* occurr in Kabir but are also very common in Rajasthani (*nakṣatra thaï* "from the constellation").

[15] Hence the sometimes competing claims for two different sources, as is the case for *tai/taï* and *te/tī*.

the semantic roles and the information flow (order of constituents). So, late MIA was not only, as developed in Bubenik (2006), a transition from case to adposition and configurational syntax as in Romance languages. It was also a transition to different preferred alignments: from a nominative language, essentially dominated by grammatical relations and subject oriented, the language shifted to a partly ergative language (aspectual semantics) and largely role and flow dominated language, where the category of subject has little relevance. Whereas subject oriented languages are 'cumulative' languages, in Kibrik's terms (1997), with the subject NP cumulating functions of various orders (topic, semantic, referential), languages which encode separately the various dimensions of communication such as semantic roles and information flow show a weak relevance of the category of subject.

5. Factors involved in the emergence and stabilization of non-nominative 'subjects'

With this background of a language deprived of morphological endings and still without a clear set of postpositions to replace them, where the predicate was simply related to its non-direct arguments in a vaguely localistic way, with no strict association between roles and formal markers, how comes that roles emerged as prevalent over syntax and came to be associated with specific postpositions, each language now devoting a given set of forms to specific functions? In the 14th c., the selection of postposition is not fixed, but a non agent of a two place predicate tends not to remain in the nominative. In addition, an agent of a transitive perfect verb is marked (which is now the case only in Western IA). How can we account for these structurally crucial changes?

5.1. Reanalysis and speaker's choice

Reanalysis has been recently articulated with a speaker-oriented and usage base grammar (Andersen 2008: 32-33, Barddal 2009). It is obviously at play in the emergence of the ergative alignment (and modal futures up to the 15th century: section 2). The initial motivation for preferring the nominal sentence to tensed verbs in the perfect was the speaker's view point mapping the result as more important than the source, as also happened in Romance languages (Cennamo 2008). When this originally stylistic or 'expressive' construction started generalizing then grammaticizing as the only expression of past, including simple anterior events (preterit), the viewpoint shifted from the result to the source, which resulted in the fronting of the agent and in Eastern languages to the shift towards nominative alignments. Similarly when modal future shifted to simple future, the goal aimed at was perceived as less important than the source, which got encoded in the nominative (Montaut forthcoming).

As for the dative/experiencer shift, when an initially transitive or nominative verb like *milnā* (or *bhānā*, or Sk root *RUC*) becomes intransitive with an experiencer subject, e.g. when a formerly marginal argument structure becomes the norm, we may hypothesis that this shift is connected with the more general shift towards the conditioning of valency by semantics: the human participant affected by a stimulus (animate or inanimate) acquires prevalence in the argument structure over the inanimate stimulus, yet not full agent properties because it is not an agent, a scenario also observed in European languages (Haspelmath 2001: 79). The raising of experiencer subjects (- volition, - control) may also be related to the cultural function of the notion *mil-* itself in the new devotional religion. This new form of mystical religion, flourishing between the 13th and 16th century in North India, centred on the surrender of personal will and fusion with the deity, came to replace the old religion of ritual sacrifice which relied on sacrificial action. Merging with the deity by surrender of personal will, getting united, became cardinal cultural notions, and the spiritual experience derived from this

union became the central aim (as opposed to the performance of ritual sacrifice previously). The fact that *mil* was by far the most frequent predicate in the preachers' poems, which came to constitute the bulk of folk culture, may have played a role in triggering the change in argument structure for other non agentive predicates, an extension which became systematic with the renewal of the predicative lexicon later. Since this devotional trend started around the 6th century in South Indian, only an enquiry on the (expectedly earlier) raise of dative subjects in the South Indian languages could confirm this hypothesis.

5.2. Lexical renewal and attraction

The major difference between early NIA and modern NIA, apart from the enrichment of case markers and TAM paradigms, is lexical: in MSH simple verbs are few compared with the profusion of complex predicates formed with a noun and a light verb. In modern NIA, a huge majority of the so-called 'dative' predicates are noun-verb complex predicates, hundreds of them in Hindi for instance, compared with the four usual simple verbs requiring the dative construction. As noted by Barddal (2009: 142)

> 'a rapid change in vocabulary may speed up the development, precisely because the bulk of new verbs will be attracted by the high frequency construction, thereby lowering the proportional type frequency of the low type frequency constructions, increasing the chances of others becoming extinct'.

That is exactly what happened in Hindi at the end of the 16th century during the heydays of Mughal empire, with a massive renewal of the lexicon starting during the reign of Akbar and his cultural politics of translation (Persian/Sanskrit/NIA). From then on, borrowed verbs have been lexicalized in the form of complex predicates with 'be' (or other statives, depending on Aktionsart or aspect) as the light verb for non agentive meanings, and 'do' for agentive meanings. Already in Kabir such an alternation appears between the simple verb (*so(vā)* 'sleep' + NOM) and the complex predicate (*nīnd ā(v)* 'sleep come' + DAT). But this kind of alternation with semantic correlates only generalized when a massive input of Persian and Arabic (later English, Sanskrit) verb roots came to regenerate the Hindi predicative lexicon. This resulted in a considerable class of equipollent predicates. *Irṣyati* 'be jealous' is re-borrowed from Sanskrit as *irṣyā* ('jealousy') *honā* ('be'), with experiencer subject (13). *Jānnā* 'know' as a simple stative verb is nominative whereas the borrowings (from Arabic or Sanskrit) *mālūm/patā honā* has the experiential argument structure and the open meaning of 'know' and 'get informed, aware' according to the tense/aspect of the light verb, now far more usual than the simple verb. Even without borrowing new verbs, some verbal bases, like *samajhnā* 'to understand', are now in competition with the complex predicate with noun (X *kī samajh mẽ ānā* : *lit.* 'come in understanding of X').

An argument in favour of extension, through attraction towards a generalizing argument structure with dative experiencer, is that *bhūlnā* 'forget', and *samajhnā* 'understand', both traditionally transitive, now tend to be used with dative subject:

(46) *mujhe   apne bačpan   kī   bātẽ   bhūl   gaĩ*
      1SG.DAT   REFL childhood GEN thing.F.PL forget go.PFV.F.PL
      'Now I have forgotten (things of) my childhood'

5.3. Differentiation between alternative argument structures: avoidance of synonymy?

Milnā and *pānā* 'get' are a good example of quasi synonymic constructions, a synonymy which occurred when *milnā* shifted its original meaning 'meet', 'mix', to the meaning 'find, get'. The transitive *pānā* and intransitive *milnā* constructions still co-exist in MSH for the meaning 'find', but *pānā* is the only possible verb to mean 'obtain', whereas *milnā* is the only

possible verb to mean 'come across, meet'. This suggests that a subtle semantic difference has come to oppose previously equivalent meanings for two distinct argument structures. This is very clear when intentionality is involved, which is the case for the two constructions of *milnā* in MSH: with NOM-INS (comitative), the verb means '(intentionally) meet', whereas with DAT-NOM it means 'find', 'get'. The many complex predicates which allow both light verbs *kar* 'do' (NOM) and *ho* 'be' (DAT) are now usually distinguished by the feature +/- volition and control (*pasand ho* + DAT: 'like', vs *pasand kar* + NOM 'chose'). But the distinctive feature triggering NOM may only be that the experiencer takes into account and intellectually acknowledges the process he undergoes, the absence of this feature (bare perception), triggering DAT. It is more than simple awareness – the bare fact of voicing jealousy, or anger, or hunger, implies some kind of awareness – but less than 'conscious choice', a feature referred to by Onishi (2001b: 127) as more "freedom" in initiating or expressing the emotion. The experiencer in (12), repeated below, is for instance jealous but not intellectually conscious of it. If it were the case, the statement would involve a nominative subject with light verb 'do', which would not be compatible with the second clause (47):

> us vaqt   mujhe tumse   irṣyā   thī   magar (mujhe) iskā bodh   nahĩ thā
> that time I.DAT 2.from jealousy was   but   (I.DAT) its   awareness NEG was
> 'At that time i was jealous of you, but was not conscious of it'

> (47)   us vaqt maĩ tumse   irṣyā   kartī thī *magar (mujhe) iskā bodh   nahĩ thā
>   that time I   2.from jealousy did   but   (I.DAT) its   awareness NEG was
>   'At that time I was jealous of you' (*but was not conscious of it)

Similarly the verb 'to feel' occurs in both constructions in the following dialogue, opposing a man (M) and a frustrated woman (W), who feels solitude and frustration. When the man comments her complaints, he uses both constructions: the nominative agentive construction (A) adds to the bare perception in the dative of experiential construction (E) a dimension of reflexive consciousness, intellectual acknowledgement, beyond simple perception:

(48) W- maĩ to   itnī   begānī   mahsūs kartī hũ is ghar mḗ   ki...
         1S T   so.much solitude feel   do PRS.1S thishouse in   that
     M- pahle   nahĩ kartī thĩ ?
        before   NEG do PAST
     W- pahle ? pahle to...M- mahsūs karnā hī   mahsūs nahĩ hotā thā aur
        Before   before TOP   feel   do.INF just feel   NEG be PAST and
        kučh-kučh mahsūs   huā   jab to pahlā mauqā milte   hī
        somewhat feel   be.PFV whenthen first   occasion findingjust
        ghar   se   čali-gaĩ
        homef   rom   leave.PFV
     'W – I have such a feeling (A) of solitude in this house that… M – you did not have it (A) before? W – before ? but before… M - you did not feel (E) that you precisely had this feeling (A) and you started having some feeling (E) of it when you left home at the first opportunity'   (Mohan Rakesh, mid 20ᵗʰ c.)

Significantly, this distinction did not seem to be present before the lexical renewal and we find in early NIA apparently meaningless alternations. A nominative subject with active predicate like *kāmīn lajjā na□ karai* 'the lascivious man does not make shame = is not ashamed') is not markedly conscious compared with the oblique experiencer of (30).

5.4. The role of contact

The so-called 'dative subject' construction is deemed to be an innovation of New Indian

languages, attributed to convergence rather than to diffusion, since this alignment was as marginal in ancient Dravidian as it was in Sanskrit (Murugaiyan 2004): epigraphic Tamil shows rare instances of this syntactic pattern with practically no evidence of the dative being a syntactic subject. But such convergence was not born spontaneously in India. Austro-Asiatic languages (particularly the Munda group) are known to be structured by semantics rather than by syntactic relations. Verbal flexion is dominated by the distinction active/non active (or neutral) and nominal flexion by the distinction animate/non animate (Biligiri 1965, Peterson 2006). At the level of predicates (in languages without verbo-nominal polarity), there are two sets of temporal affixes according to the nature of the predicate: *-te, -og, -e* for respectively the present, past and future of active predicates like 'beat' (*gilte, gilog, gile*), and *–ta, -ti, -na* for the same in non active predicates like sleep (*gitagta, gitagti, gitagna*) (Biligiri 1965: 56 sq). In the nominal class, the suffixes differ according to the animate or inanimate nature of the noun (Santali *dadal-ic* 'the one that strikes', vs *dadal-ak* 'the thing that strikes'). There are certainly many factors at play in the rising of the so-called 'dative subject' construction in NIA, e.g., in the differential treatment of human non agentive agents as contrasted with animate agentive entities, and, generally, in the differential treatment of action statements and transient state statements. But the presence of a substratum of 'active' languages in the area concerned has certainly triggered the systematic use of oblique agent. Besides, the fact that de-ergativation (shift of oblique agents to nominative subjects) which occurred in the East of NIA (from Bihar to Bengal and Assam) is explainable by the high density of Austro-Asiatic speakers in these regions. Eastern Hindi (Bihar) also presents an agreement pattern clearly related to the Munda pattern (verb indexing all participants of the clause, thus displaying 'multiple' agreement (Verma 1991).

Conclusion

According to the specialists, the way Sanskrit language has been described by Indian ancient grammarians reflects the properties of the language in its older stage: Deshpande for instance analyses Panini's relevant sections as allowing the syntactic promotion of various oblique cases to the main case or agent (*kartā*), but notices that such a promotion never occurs for the beneficiary/experiencer, until the latest stages of the language. This suggests that instrumental as well as genitive subjects have always been present in IA (another 'change that never happened' in Barddal's (2003) words), in contrast to experiencer subjects and inadvertent or inefficient agents. The syntactic notion of subject as well as transitivity is significantly absent in this description, and the reason for making agent and not subject the basic function in the sentence is related to the philosophical frame underlying the Indian theory of language. Within a Brahmin culture entirely centred on ritual sacrifice, where the most important act was the ritual action, action was perceived as the model for any linguistic expression (Deshpande 1991, Al-George 1957). Hence the privilege conferred to agentivity and action.[16] In contrast, the Western model, inherited from Aristotle's philosophy of language within a conceptual frame dominated by logic and dominated by the relation between 'subject' and 'predicate', i.e. topic and comment, the nominal predication was selected as a privileged model for linguistic expression, because it was the best illustration of the logical judgement (Al-George 1957: 46-7). Whereas ritual speculations oriented grammatical doctrines in India, logical speculations oriented grammatical doctrines in Europe. Obviously none is universal, and when it comes to the description of concrete language, the subject as a grammatical category is no more central in ancient Greek than in Sanskrit, and the agent no less. Despite its huge fortune in Western language theory and language descriptions, the notion of subject

---

[16] Expressions like *puruṣa* "man" or *jīva* "living being" are treated as nouns of agency in Panini III.4.43.

is clearly not a universal, and part of the problems raised by the split of subject properties on several NPs only emphasize its weak relevance for certain languages, particularly topic or role oriented languages (Kibrik 1997).

The speculative frame which came to dominate Indian culture during the time when the new languages acquired their modern characteristics has not given birth to a new language theory and grammatical descriptive tools. But it was probably instrumental in the emergence of 'experiencer subjects', in emotional and cognitive predicates like *milan/milnā* ('find/meet' the truth or spiritual leader, 'get' the spiritual revelation, 'get united' with God), referring to the crucial experience in the new devotional mystic. Other factors were probably more decisive, such as lexical renewal and contact, in what we can consider as a shift towards semantic alignments. There is a fairly high frequency in the language of alignments constrained by lexical semantics, whose first argument is encoded in the same way as direct and indirect objects (NOM/DAT). (Arkadiev 2008) allowed perfective aspect to rank first among the language specific properties which can trigger semantic alignments in Loma. The Hindi ergative, with its marker born out of a locative and used for DAT/ACC in related languages, may also be viewed as part of a general language shift towards semantic alignments.

Aikhenvald, Alexandra, Dixon, R.M.W., Onishi, Mayasuki (eds.). 2001. *Non-Canonical marking of Subjects and Objects*. [TSL 46]. Amsterdam: Benjamins.

Al-George, Sergiu. 1957. Le Sujet grammatical chez Panini. *Studia et Acta Orientalia* 1: 39-47.

Andersen, Henning. 2008. Grammaticalization in a speaker-oriented Theory of Change. In *Grammatical Change and Linguistic Theory*. Eythorsson (ed.), 11-44.

Arkadiev, Peter M. 2008. Thematic roles, event structure and argument encoding in semantically aligned languages. In *The Typology of Semantic Alignments*. Donohue, Mark & Winchmann, Soeren (eds.), 101-117, Amsterdam: Benjamins.

Barddal, Johanna. 2003. The change that never happened: the Story of Oblique Subjects. *Journal of Linguistics 39*: 439-472.

Barddal, Johanna. 2009. The Development of Case in Germanic. In *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*. Barddal Johanna and Chelliah L. Shobhana (eds.), 123-59. Amsterdam: Benjamins.

Bhaskararao, P. & Subbarao, K.V. 2004. *Non-nominative Subjects*. Amsterdam: Benjamins.

Biligiri, Hemige Shriniwasarangaghar. 1965. *Kharia*. Poona: Deccan College.

Bloch, Jules. 1906. *La Phrase nominale en sanscrit*. Paris : Champion.

Bubenik, Vit. 2006. Cases and Postpositions in Indo-Aryan. In *From Case to Adposition. The developmernt of configurational syntax in Indo-European Languages*. Hewson, John and Bubenik, Vit, 102-130. Amsterdam: Benjamins.

Cardona, George. 1991. A Note on 'Dative Agents in Sanskrit. In *Experiencer Subjects*, Verma M.K & Mohanan K.P. (ed.), 141-145.

Cennamo, Michela. 2008. The rise and development of analytic perfects in Italo-Romance. In Eythorsson (ed.), 115-142.

Chatterji, S.K. 1926 [1986]. *The Evolution of Bengali Language*. Delhi: Rupa.

Chatak, G. 1966. *Mādhyapahārī kā bhāsāśāstrīya adhyāyan*. Delhi: Radhakrishna Pr.

Dahl, Eystein. Forthcoming. Morphosyntactic Variation and the Differenciation of Semantic Roles: Evidence from Vedic Sanskrit.

Deshpande, Madhav. 1991. Some Features of the *sampradāna kārika* in Panini. *In Experiencer Subjects in South Asian Languages*. Verma, Mahindra K. and Mohanan K. P. (eds.), 147-59.

Donohue, Mark & Soeren Wichmann. 2008. *The Typology of Semantic Alignment*. Oxford: OUP.

Eythorsson, Thorhallur (ed.). 2008. *Grammatical Change and Linguistic Theory*. Amsterdam: Benjamins.

Haspelmath, Martin. 2001. Non canonical marking in European Languages. In *Non-Canonical marking of Subjects and Objects*. Aikhenvald & al., 53-83.

Hock, Hans Heinrich 1991b. Oblique Subjects in Sanskrit. In *Experiencer Subjects in South Asian Languages*. Verma, Mahindra K. and Mohanan K.P. (eds.), 119-140.

Hock, Hans Heinrich. 1991a. Genitive Possessor. *Studies in Sanskrit Syntax*. Delhi: Motilal Banarsidass: 55-69.

Hoernle, August, Friedrich, Rudolf. 1880. *A comparative grammar of the Gaudian languages: with special reference to the Eastern Hindi, accompanied by a language-map and a table of alphabets*. London.

Hook, Peter Edwin. 1976. *As☐tādhyāyī 3.4.2.1* and the role of semantics in Paninian linguistics. *Papers from the 12th Regional Meeting of the Chicago Linguistic Society*: 302-12.

Kibrik, Alexander E. 1997. Beyond subject and object: Toward a comprehensive relational typology. *Linguistic Typology* 1: 279-346.

Kulikov, Leonid. 2009. The Evolution of Case Systems. In *The Oxford Handbook of Case*. Malchukov, Andrei and Spencer, Andrew (eds.), 438-457. Oxford: OUP.

Montaut, Annie. 2004b. *Hindi Grammar*. Munchen: Lincom Europa.

Montaut, Annie. 2004a. Oblique main Arguments in Hindi. In *Non-nominative Subjects*. Bhaskararao & Subbarao (eds.), 33-56.

Montaut, Annie. 2011. Basic Intransitivity: A Typologically Relevant Feature of Indo-Aryan. In *Indo-Aryan Linguistics*. O.N. Koul (ed.), 29-46. Mysore: CIIL,.

Montaut, Annie. Forthcoming. Why the ergative case in modal (in)transitive clauses? Aspect, modality, ergative and locative in Indo-Aryan.

Murugaiyan, Appassamy. 2004. Note sur les prédications expérientielles en tamoul classique. *Bulletin de la Société de Linguistique de Paris* 99: 363-382.

Onishi, Masayuki. 2001a. Non-Canonically Marked Subjects and Objects: Parameters and Properties. In *Non-Canonical Marking of Subjects and Objects*. Aikhenvald & al. (eds.), 1-44.

Onishi, Masayuki. 2001b. Non-canonically Marked S/A in Bengali. In *Non-Canonical Marking of Subjects and Objects*. Aikhenvald & al. (eds.), 113-1147.

Peterson, John. 2006. A *Grammar of Kharia*. Amsterdam: Brill.

Tiwari, Uday Narayan. [1995]1955. *Hindī bhāşā kā udgatan aur uskī vibhinn boliyā̃*. Allahabad: Bharti Press.

Tsunoda, Tasaku.1985. Remarks on Transitivity. *Journal of Linguistics 21*: 385-96.

Verma, Mahindra K. and Mohanan K. P (eds.). 1991. *Experiencer Subjects in South Asian Languages*. Stanford: The Stanford Linguistic Association.

Verma, Mahindra K. 1991. Exploring the Parameters of Agreement: the Case of Magahi. *Language Science 13-2:* 125-43.

Wichmann, Soeren. 2008. The study of Semantic Alignment: Retrospect and state of the art. In *The Typology of Semantic alignment*. Donohue, Mark & Wichmann, Soren (eds.), *The Typology of Semantic Alignment*, Oxford, OUP, 3-23.

Abbreviations (other than in the standard list of Benjamins)
OVA obligative verbal adjective; PPP: past passive participle, H: honorific
The gloss PFV was used in conformity with the usual literature on Hindi for forms which rather correspond to a preterit (anterior event), perfectivity being marked by different devices