



**HAL**  
open science

**Noun-Verb Complex Predicates in Hindi and the rise of  
non-canonical subjects Annie MONTAUT Sedyl  
(INALCO/CNRS/IRD)**

Annie Montaut

► **To cite this version:**

Annie Montaut. Noun-Verb Complex Predicates in Hindi and the rise of non-canonical subjects Annie MONTAUT Sedyl (INALCO/CNRS/IRD). Nash Lea & Samvelian Pollet. Approaches to Complex Predicates., 41, Brill, 2016, Syntax and Semantics, 9789004306981. hal-01313043

**HAL Id: hal-01313043**

**<https://hal.science/hal-01313043>**

Submitted on 27 Nov 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Noun-Verb Complex Predicates in Hindi and the rise of non-canonical subjects**  
**Annie MONTAUT**  
**Sedyl (INALCO/CNRS/IRD)**

**Published in** NASH, LEA & SAMVELIAN, POLLET (EDS.).2016  
*APPROACHES TO COMPLEX PREDICATES*. L: LEYDEN: BRILL (SAS 41), PP. 142-174.  
ISBN 13 9789004306981

ABSTRACT

Verbo-nominal complex predicates, the bulk of the Hindi predicate lexicon, a highly heterogeneous category, regarding compositionality and idiomaticity, can however be sub-classified in three types regarding agreement, internal and external syntax, separability, omissibility, extraction, and acceptable modifiers. However, in spite of the highly heterogeneous nature of the category, they behave similarly in widening the scope of Aktionsart, aspect and voice specifications in comparison to the grammatical markers of TAM in simplex predicates. The most important fact deals with the diachronic evolution of the language: all sub-types of CPs, massively introduced during the renewal of the verbal lexicon by means of borrowing, have contributed to a general shift towards semantic alignments by licensing non-canonical subjects, mainly dative and genitive subject constructions.

**1. Introduction: The place of nominal complex predicate in the language**

The existence of complex predicates formed by the association of a non verbal element with a light verb is considered as one of the pan-indian features which make India a linguistic area: they now exist in all the various families of languages spoken in the subcontinent (Emeneau 1980), whereas they used to be absent or marginal in earlier stages of the language. The same is true for other pan-Indian features such as dative experiential subjects, and the alternation of transitive/intransitive on the same verbal basis. Within the general category of complex predicates, the noun/adjective-verb class is the one less studied compared with the verb verb class (Hook 1974, Butt & Geuder 2003, 2006, Raina & Mukerjee 2005, Mukerjee, Sinha & Mahesh 2009 among others). The first class, sometimes called conjunct verb (Burton-Page 1957, Begum & al. 2011), has most frequently for its first constituent a noun but can also host an adjective, a frozen participial form or even an adverb. In such combination the predicative notion or semantic content is conveyed by the noun (or adjective) and actualized in terms of TAM and voice by the verb. Predicates such as *kām karnā* (work do) “to work”, *pyār karnā* (love do) “to love”, *icchā honā* (desire be) “to desire”, *nazar ānā* (look come) “to see”, *band karnā* (closed do) “to close”, *khālī karnā* (empty do) “to empty”, *xatm karnā* (end do) “to finish”, *cakit honā* (surprised be) “to startle”, *andar ānā* (inside come) “to enter”, belong to this class of complex predicates which I will further mention under the generic label noun verb CP.

These complex predicates represent a massive proportion of the Hindi verbal lexicon (Kellogg’s list, non exhaustive, from the mid 19<sup>th</sup> century, gives 175 frequent CP with only the light verb *karnā* “do”, which almost doubles with the intransitive counterpart of *karnā* “do”, *honā* “be”. Contrary to most European languages, they occur rarely in pairs with simplex predicates, and the predicative notions “to work” or “to wait” cannot for instance be expressed by a simplex predicate: Verma (1993: 200) lists only half a dozen of complex/simplex pairs “which readily come to mind”. Besides, and that too has been noticed recurrently (from Gambhir 1993 to Davison 2005), a simplex predicate when alternating with a complex predicate rarely has the same meaning: *bāt karnā* and *batānā* for instance, two extremely frequent predicates, have different meanings, respectively “to speak” and “to tell, relate”).

Among the verbs used as light verbs (Butt 1995, Mohanan 1995) or verbalizers (Begum & al. 2011), “do” and “be” (*karnā* and *honā*) are by far the most frequent. Others consist in a dozen verbs, mainly motion verbs, either transitive, such as *rakhnā* “to place”, *ḍālnā* “to throw”, *mārnā*

«to strike», *denā* “to give”, *lenā* “to take”, *khānā* “to eat”, or intransitive such as *rahnā* “to stay”, *ānā* “to come”, *lagnā* “to touch”, *paṛnā* “to fall”, *uṭhnā* “to rise”. The verb *khānā* “to eat” is an exception because it usually does not produce agentive CP in meaning although it is transitive, and forms mostly idiomatic CPs. A good half of these verbs are also used as vectors in the verb-verb combination which will not be considered here. As for the first constituent, it represent an open class, action nouns being the bulk of the verbo-nominal class, but concrete nouns are not ruled out, particularly in idiomatic CP, and adjectives are not always clearly identifiable as such<sup>1</sup>. The study will concentrate on verbo-nominal combinations, the most interesting regarding the shift from nominative to dative alignments.

The aim of this paper is to draw a tentative typology of such complex predicates, a category highly heterogeneous regarding agreement and internal syntax, both resulting from the relation between the two components of the CP, as well as regarding the linking device of the CP with its closest argument, and the autonomy of the non-verbal component (section 2). However, in spite of the highly heterogeneous nature of the category, it displays common features regarding the availability of voice and aspect alternations (section 3), as well as with argument structure alternations, particularly the so-called dative subjects alignments, a major typological shift in Indo-Aryan languages (section 4).

## 2. Towards a typology of nominal complex predicates

### 2.1. Identifying the category

Among the most often mentioned criteria for distinguishing a CP from an ordinary N V sequence, one is semantic: the predicative notion is conveyed more by the host noun than by the verb. The other is syntactic: in Noun Verb transitive complex predicates, this Noun cannot appear in the accusative (+ *ko*), a case used for human or specific inanimate objects, but an external object can occur with the accusative marker (2).<sup>2</sup>

(1)	<i>intazār</i>	* <i>ko</i>	<i>karegī</i>
	waiting	ACC	do.FUT.3F.SG
	‘She will wait’		

(2)	<i>vah</i>	<i>Kamlā ko</i>	<i>pasand</i>	<i>kartā-hai</i>
	3SG	Kamla ACC	taste	do-PRS.3MS
	‘He likes Kamla’			

Although a sequence of Noun Verb in which the noun does not behave as a canonical object seems to be a simple and convenient criteria for isolating the class of noun-verb CPs, most of the noun involved in CPs also behave as independent nouns in the language, sometimes with the same verb as the CP’s light verb, and non-agentive CPs (with light verb “be”) obviously take no marked/unmarked accusative, so that this criteria does not help much. Besides, idiomatic

<sup>1</sup> Some of the units used as first constituent of such CP, such as *mālūm* (*honā* “know”), *gāyab* (*honā* “disappear”), *qabūl* (*honā/karnā* “accept”) are not used independently, although dictionaries translate them as adjectives, whatever their original category in Arabic or Persian was. Others, such as *afsos* (*honā* “be sorry”) *pasand* (*honā/karnā* “like”) occur outside the CP combinations, but always in highly constrained contexts, and their « nominal » status is not that of regular nouns (*afsos kī bāt* “sorry, sad thing” (sadness of thing), *āpkī pasand* ‘your liking’. A term like *manzūr*, from Persian (‘accept’), is considered an adjective but *inkār*, from Arabic (‘refuse’), as a noun. These “adjectives” however do not behave as ordinary adjectives which can be noun modifiers in noun phrases.

<sup>2</sup> Glosses correspond to single words, except for certain compound tenses like the indicative present, formed the adjunction of the ‘be’ finite verb (person, gender/number) to the present participle (+gender and number), like the English progressive : since the detailed gloss would only add heavy useless morphological complexities and opacity, I chose to write the hindi compound form with hyphens and simplify the gloss to the lexical meaning plus TAM and agreement only once at the end.

combinations, such as *bāt johnā* or *rāh dekhnā* (path look) “to wait anxiously for/expect”, which also fall under this constraint, are usually considered as phraseology (*muhāvvrā*) rather than complex predicates. It is probably the reason why most of the studies either concentrate on clearly compositional combinations (Sinha & Mahesh 2009, Chakrabarti & al. 2007) or recognize the artificial character of such a definition as above, and the fuzzy limits of the category (Gambhir 1993). Whatever the idiomaticity, there is a form of compositionality in the sense that each of the constituents have a specific input which varies according to the distinct combinations in which they belong (Samvelian & Faghiri under press), and clear compositionality cannot be an operational criteria: the same sequence V N can behave differently in different contexts and still be listed in all the descriptions of the category CP and in CP lexicons such as Indian language wordnet (Chakrabarti & al. 2007).

This is the case for instance of *pyār karnā* “to love”, a very frequent combination in which the degree of autonomy of the noun (crucial for the discrimination of the sequence Noun Verb vs CP) is far from being stable. In the following example (3), the yes/no question has scope on the CP, and the answers makes the noun distinctly autonomous since it is relativized<sup>3</sup>:

(3)	- <i>kyā</i>	<i>tumne</i>	<i>kisī</i>	<i>se</i>	<i>saccā</i>	<i>pyār</i>	<i>kiyā-hai?</i>
	INTER	2.ERG	INDEF	SOC	true	love	do-PRF
	- <i>jaise</i>	<i>pyār</i>	<i>kī</i>	<i>maĩ</i>	<i>bāt</i>	<i>kar-rahī-hũ...</i>	
	such	love	of	1.SG	speech	do-PROG-PRS.1.SG	
	‘- Have you really loved (CP) anybody? - Such a love (N?) I am talking about/ to love in the way I am talking about...’						

Similarly in (4), the answer denies a clause with the CP “to think” (thought be), by using the noun “thought” as a possibly independent noun:<sup>4</sup>

(4)	- <i>merā</i>	<i>khyāl</i>	<i>thā...</i>	- <i>tumhārā</i>	<i>khyāl</i>	<i>galat</i>	<i>thā</i>
	my	thought.M.SG	be.IMPFT.M.SG	2.GEN	thought	wrong	be.IMPFT.M.SG
	‘- I thought (that) (CP)... - You were wrong /your thought was wrong’ (N? CP?)						

And in (5), with a disjunctive coordination “or”, the two elements made parallel are the CP and the noun:

(5)	<i>tum</i>	<i>pyār</i>	<i>cāhtī-ho</i>	<i>yā</i>	<i>pyār</i>	<i>karnā</i>	<i>cāhtī-ho?</i>
	2	love	want-PRS.2F	or	love	make.INF	want-PRS.2F
	‘Do you want love (N) or do you want to love (CP)’						

The last example provides a full range of varying degrees in the autonomy of the noun, from canonical CP in (6a) and (6b) with two occurrences with the verb elided, to canonical Noun which is not used as an argument (6c), but yet answers the questioning on the initial CP “to complain”, till the final occurrence as a specified noun (6d):

(6)	a.	<i>to</i>	<i>hamē</i>	<i>śikāyat</i>	<i>na</i>	<i>hotī</i>
		<i>then</i>	1PL.DAT	complaint	NEG	be.COND
		‘- Then we would not complain’ (CP)				

<sup>3</sup> For the adjectival modifier which modifies the CP and not only the noun, cf. *infra* 2.3.4.

<sup>4</sup> Both translations “you were wrong” and “your thought was wrong” reflect this ambiguity. In the case of “think”, a simple verb would have been possible, and the same meaning achieved by the similar use of the adjectival/adverbial modifier “wrong” (*maĩne socā... tumne galat socā*). In the context (A wishes to justify his position (“I thought that”), the first occurrence differs in no way from the standard use of the CP *khyāl honā* ‘think’. Examples (3) to (7) are from contemporary theatre (Vaid, 2005 and 2009).

(6)	b.	- <i>inhē</i>	<i>ham</i>	<i>se</i>	<i>śikāyat</i>	<i>kyō?</i>	- <i>śikāyat</i>	<i>kyō</i>	<i>nahī?</i>
		3.PL.DAT	1.PL	from	complaint	why	complaint	why	NEG
	‘- Why should they complain about us? - Why should they not?’								

(6)	c.	<i>kyōki</i>	<i>śikāyat</i>	<i>se</i>	<i>śikāyat</i>	<i>kā</i>	<i>mūl</i>	<i>kāraṇ</i>	<i>dūr</i>	<i>nahī</i>
		because	complaint	from	complaint	of	basic	reason	far	NEG
	‘- Because the basic reason of complaint (for complaining) is not far from a complaint /from complaining’									

(6)	d.	<i>kis</i>	<i>śikāyat</i>	<i>kā?</i>
		which	complaint	of
	– (The reason) of which complaint?’ (N)			

Such examples show, again, a great deal of variability regarding the degree of independence of the noun: the real contexts of production allow all kinds of a priori unexpected behaviours, which show that in its really productive usages, CPs can be as close to simple verbs as to a mere sequence of noun verb, and that for a same noun, in not particularly marked discursive sequences.

The constituent response test (WH questions) has been shown to isolate the CP and not the verb as a constituent (Begum & al. 2011)<sup>5</sup>. “He yawned” (yawn took) cannot be an answer to the question “what did he take?” since the question should bear on the whole constituent (“what did he do?”). That seems to be the best test, providing it takes voice alternations into account (“what did he do/was he?”), as suggested by the following dialogue: the predicative notion that is questioned, then answered or corrected in (7), is obviously not the verbal constituent, but neither is it the CP (“doubt be”) as a whole. It is the noun in association with the paradigm of possible light verbs for voice derivation (allowing transitive and intransitive constructions, with respectively agentive and medio passive meanings, cf. sections 3 and 4).

(7)	<i>tumhē</i>	<i>śak</i>	<i>hai?</i>	- <i>śak</i>	<i>karnā</i>	<i>merā</i>	<i>dharm</i>	<i>hai</i>
	2.DAT	doubt	is?	doubt	do.INF	my	religion	is
	‘Do you doubt/have you doubts? (INTR) – Doubting (TR) is my religion’							

## 2. 2 The various classes of complex predicates

### 2.2.1 State of the art

Kellogg’s categorization (1875: 272-79), the first attempt to classify the proliferating data of NV complex predicates, obtains 15 distinct types, by using both the argument structure (case of external argument) and the verb as discriminating features. Verma (1993:) recognizes 3 types according to case marking: type 1 is a “purely analytical sequence” (*kām karnā* “to work”) behaving as a sequence object-main verb; type 2 is “syntactically complex”, since the verb case marks the nominal constituent but the noun assigns theta roles and selects arguments (*kī madad karnā* “to help”); and type 3 is an “amalgamation” of the noun and verb and behaves as a simple verb (*praṇām karnā* “to greet, to salute”). Davison (2005) and Montaut (2004, 2012) also recognize three types (on account of agreement, case marking and licensing of external argument). Ahmed & Butt’s (2011) is the only study to my knowledge which proposes a semantic classification entirely based on the selectional properties of the noun conveying the predicate meaning. Since argument

<sup>5</sup> Similarly the coordination test excludes the coordination of a noun to the noun within the CP, but not of a verb, providing it is compatible with the noun: ...*pasand nahī āi aur kabhī nahī hogī*, [taste NEG come.PFV and never be.FUT] “do not like, and will never like”.

structure is largely constrained by semantic roles, their study also involves case marking and argument structure: the 3 types identified are the following (authors phrasing): (1) nouns allowing the full range of constructions (light verbs *kar* “do”, *hai* “be” in the present with experiencer, *hu-* “be” in the perfective with an eventive meaning), such as “memory”, (2) those excluding the experiencer, such as “construction”, and (3) the “non eventive” ones which exclude *hu-* (base of ‘be’ verb for anterior tense), such as “waiting”<sup>6</sup>.

## 2.2 The data

Although regarding agreement, there are only two types, the one ruling out internal agreement, or “close association” type in the terminology of Davison (2005), which corresponds to the “coalescent” in Lazard 1991 and Montaut 2004) where the noun never controls agreement, and the one requiring internal agreement, where the verb agrees with the nominal constituent of the CP, the latter is not a homogeneous category. In transitive sentences with nominative alignments (non perfective), the light verb always agrees with the subject, like a simplex predicate, and the difference between both types shows only with intransitives light verbs, in experiential and ergative alignments, as well as in the passive, similar to ergative in this respect.

### 2.2.1. External agreement

The light verb agrees with the external argument when the statement is experiential (8), ergative (9) or passive in example (10), where *anubhav karnā* “feel” behaves as the simplex verb *samajhna* “understand”.

(8)	<i>mujhe</i>	<i>unkī</i>	<i>filmē</i>	<i>bahut</i>	<i>pasand</i>	<i>āī</i>
	1SG.DAT	his	film.F.PL	much	taste	come.PFV.F.PL
	‘I liked his films very much’					

(9)	<i>Rānā</i>	<i>ne</i>	<i>śukravār</i>	<i>ko</i>	<i>kāgres</i>	<i>kī</i>	<i>sadasyatā</i>	<i>grahan</i>	<i>kar</i>	<i>lī</i>
	Rana	ERG	Friday K	DAT	Congress	of	membership.F.SG	grasp	do	take.PFV.F.SG
	‘Rana received the Congress membership on Friday’ (Press: <i>Jagaran</i> June 2015)									

(10)	<i>ye</i>	<i>bātē</i>	<i>samjhī</i>	<i>nahī</i>	<i>anubhav</i>	<i>kī</i>	<i>jā</i>	<i>saktī haī</i>
	these	thing.F.PL	understand.F	NEG	experience.M.SG	do	PASS	can PRS.F.PL
	‘These things [cannot be] understood, [they] can be experienced’ (Yadav, novel 1980)							

Mohanani (1994, 1995) then Dayal (2011) suggested the hypothesis of «semantic» or “pseudo-incorporation” to account for such facts, a label all the more ambiguous since verbs usually incorporate their object and here subject are also “incorporated” in the case intransitive light verbs. Davison (2005) has a more classical analysis of this type of CP as a merging of features<sup>7</sup>.

As a corollary the CP may, if transitive, licence an external object which can be optionally marked as specific or human as seen in example (2).

<sup>6</sup> Or *bardāst* “tolerate”, which however, at least in Hindi, freely licenses the *-hu* construction: *Pāk ko bardāst nahī huā, hatyā kar pahūcā* [Pak DAT tolerate NEG be.PFV murder do.CV arrive.PFV] ‘Pak could not stand [it], he ended up committing murder’; *garmī bardāst nahī huī, ‘murde’ uṭh khare hue* [heat.F.SG tolerate NEG be.PFV.3F.SG dea d.M.PL rise stand be.PFV.3M.PL] ‘[They] could not bear the heat, the ‘dead’ rose up’

<sup>7</sup> Predicates such as *śurū karnā/honā* “to begin”, *pasand karnā/honā* “to chose, to like/to please”, *svikār/manzūr karnā/honā* “to agree”, *yād karnā* “to remember”, *istemāl karnā* “to use”, *anubhav/mahsūs karnā* “to feel”, *khatm karnā* “to finish”, *grahan karnā* “to grasp, take”, *mālūm honā* and *patā honā* “to know”, like all participial CP such as *band karnā* “to close”, *bardāst karnā* “to bear, to tolerate” belong to this class.

### 2.2.2 Internal agreement

In the first subtype, agreement in the relevant syntactic contexts (as above) is controlled by the noun of the CP and there is no vacant place for an external argument in transitive clauses (10), since it is obligatorily case marked in the genitive (the genitive postposition has adjectival features since it agrees with the head noun in gender and number). Transitive predicates show agreement with the noun of the CP in the past with ergative agent (11a) as well as in the passive and middle (11b):

(11)	a	<i>atākvādī</i>	<i>ne</i>	<i>das</i>	<i>ādmīyō</i>	<i>kī</i>	<i>hatyā</i>	<i>kī</i>
		terrorist.M.SG	ERG	ten	man.M.PL	GEN	murder. FS	do.PFV.F.SG
		‘The terrorist killed ten people’						

(11)	b	<i>das</i>	<i>ādmīyō</i>	<i>kī</i>	<i>hatyā</i>	<i>kī-gaī /</i>	<i>huī</i>
		ten	man.M.PL	GEN	murder. FS	do-PASS.F.SG	be.PFV.F.SG
		‘Ten people were killed’					

Although not case marked as an argument (Hindi has no simplex verb requiring genitive objects), this external argument retains the discursive properties of arguments.

In the second subtype, the noun of the CP controls agreement in the same conditions as above (intransitive light verb, passive and ergative of transitive light verb), but the external argument retains a standard argument status (sociative, locative, less often accusative; *pranām k.* “to salute”, *mardan karnā* “to rub, massage”), similar to the case marking of simplex predicates:<sup>8</sup>

(12)	a	<i>maī</i>	<i>us</i>	<i>ādmī</i>	<i>se</i>	<i>nafrat</i>	<i>/pyār</i>	<i>kartā-hū</i>
		1SG	DEM	man	SOC	hatred.F.SG	love.M.SG	do-PRS.1SG
		‘I hate/ love this man’						

(12)	b	<i>mujhe</i>	<i>us</i>	<i>ādmī</i>	<i>se</i>	<i>nafrat</i>	<i>/pyār</i>	<i>hai</i>
		1SG.DAT	DEM	man	SOC	hatred.F.SG	love.M.SG	be PRS.3SG
		‘I hate/ love this man’						

(13)	<i>maīne</i>	<i>tum</i>	<i>par</i>	<i>viśvās</i>	<i>kiyā</i>
	1SG.ERG	2	on	trust	do.PFV.3M.SG
	‘I trust you’				

(14)	<i>maīne</i>	<i>ḍāktar</i>	<i>se</i>	<i>bāt</i>	<i>kī</i>
	1SG.ERG	doctor	SOC	speech.F.SG	do.PFV.3F.SG
	‘I spoke to the doctor’				

A few NV combinations allow two constructions (with external agreement and with genitive case marking) such as *yād* “memory”, *patā* “know”, *anubhav* “experience”, *istemāl* “use”, *qatl* “murder”, *tabāh* “destruction” but they can be listed as exceptions.

### 2.3 Autonomy of the two constituents

<sup>8</sup> The external argument, analysed as adjunct in other theoretical frames, is similarly a sociative in equivalent statements with simplex predicates such as :

<i>maī</i>	<i>ḍāktar se</i>	<i>bolā</i>	<i>maīne ḍāktar</i>	<i>se</i>	<i>kahā</i>
1SG[M]	doctor INS	speak.PFV.M.SG	1SG.ERG doctor	SOC	say. PFV.M.SG

‘I spoke to the doctor / I told the doctor’

Regarding agreement, Hindi CP is a two-fold category. Regarding case marking, it is also a twofold category, but both criteria do not coincide. Type 1 and second subtype 2 are similar regarding the external argument, dissimilar regarding agreement, according to which both subtypes of type 2 are similar (internal agreement). Does the autonomy of the noun argues in favour of a threefold or twofold classification? For brevity and clarity below, let us tag the three classes as NV (external agreement), *kā* N-V (genitive external argument, internal agreement), and N-V (standard external argument, internal agreement).

### 2.3.1 Separability of the two constituents

The existing literature on Hindi/Urdu (Butt 1995, Begum & al 2011) as well as on Persian (Samvelian 2012, Karimi-Doostan 2014) shows that negative, interrogative and discourse particles freely occur between the noun (or adjective) and the verb. That is expected for CP with internal agreement as in example (6), but it also systematically occurs with CPs displaying external agreement, supposedly with both constituents in “close association”.

(15)	<i>kūrā</i>	<i>tumhē</i>	<i>nazar</i>	<b><i>nahī̃</i></b>	<i>āyā?</i>	<i>-nahī̃</i>	<i>lekin</i>	<i>mahsūs</i>	<i>to</i>	<i>huā</i>	<i>hogā?</i>
	garbage	2.DAT	look	NEG	came	no	but	feel	TOP	be.PFV	PRSP
	‘- Didn’t you see the garbage? — No..– But you must at least have felt (something)?’										

(16)	<i>tumne</i>	<i>śādī</i>	<i>karnā</i>	<i>māzūr</i>	<b><i>kyō</i></b>	<b><i>nahī̃</i></b>	<i>kiyā?</i>
	2.ERG	marriage	do.INF	agree	why	NEG	do.PFV.M.SG
	‘Why didn’t you agree to marry?’						

Intervening heavier elements, such as interrogative constituents or focus constituents, which are preverbal in Hindi, however show that the three types do not behave exactly the same way. The internal agreement class is far more constrained than the others, and its lexicon contains a far more limited number of entries.

A focused constituent freely occurs between noun and verb in the N-V class in statements as natural as those with the focused constituent immediately before the CP and for some speakers more natural (17), whereas in the NV class with external agreement class the pre CP position is far more natural (18b) and the preverbal position highly marked or disliked by some speakers (18a) :

(17)	a	<i>mausam</i>	<i>par</i>	<i>bharosā</i>	<i>kaun</i>	<i>kar</i>	<i>saktā-hai</i>
		weather	on	trust	who	do	can PRS.3M.SG
	b	<i>mausam</i>	<i>par</i>	<i>kaun</i>	<i>bharosā</i>	<i>kar</i>	<i>saktā-hai</i>
		weather	on	who	trust	do	can PRS.3M.SG
		‘Who can trust the weather?’					

(18)	a?	<i>yah</i>	<i>sastī</i>	<i>sārī</i>	<i>pasand</i>	<i>kaunsī</i>	<i>laṛkī</i>	<i>karegī?</i>
		DEM	cheap	sari	taste	which	girl	do.FUT.3F.SG
	b	<i>yah</i>	<i>sastī</i>	<i>sārī</i>	<i>kaunsī</i>	<i>laṛkī</i>	<i>pasand</i>	<i>karegī?</i>
		DEM	cheap	sari	which	girl	taste	do.FUT.3F.SG
		‘Which idiot girl would chose/like this cheap sari?’						

The *kā* N-V class allows both positions, although the pre CP position in intransitive predicates is more marked (19b), although less than in the NV class with external agreement:

(19)	a	<i>rām</i>	<i>kī</i>	<i>madad</i>	<i>kaunsā</i>	<i>bevkūf</i>	<i>laṛkā</i>	<i>karegā?</i>
------	---	------------	-----------	--------------	---------------	---------------	--------------	----------------



		Rām	GEN	help	which	stupid	boy	do.FUT.3M.SG
		<i>rām</i>	<i>kī</i>	<i>kaunsā</i>	<i>bevkūf</i>	<i>laṛkā</i>	<i>madad</i>	<i>karegā?</i>
	b	Rām	GEN	which	stupid	boy	help	do.FUT.3M.SG
		‘Which stupid boy would help Ram?’						

Contrasts with CP allowing two alternative constructions (external agreement NV and *kā* N-V with internal agreement) are particularly revealing: focus constituents in between Noun and Verb are far more natural with the *kā* construction (20a) than in the external agreement constructions (20b), disapproved by some speakers:

(20)	a	<i>is</i>	<i>zamīn</i>	<i>kā</i>	<i>istemāl</i>	<i>kāfī</i>	<i>samay</i>	<i>se</i>	<i>ho-rahā-thā</i>
		DEM	ground.F.SG	GEN	use.M.SG	some	time	from	be PROG.IMPF.M.SG
		‘This ground was used since quite some time’							

(20)	b	?	<i>yah</i>	<i>zamīn</i>	<i>kāfī</i>	<i>samay</i>	<i>se</i>	<i>istemāl</i>	<i>ho-rahī-thī</i>
			DEM	ground.F.SG	some	time	from	use.M.SG	be PROG.IMPF.M.SG
			‘This ground was used since quite some time’						

The N-V is the less constrained regarding separability, but the difference between the two other types is more a question of markedness and frequency than of grammaticality.

### 2.3.2. Omissibility in parallel constructions and question/answer sequences

Examples of verb omission such as (6b) occurs mainly in the N-V class (*śikāyat*), but ellipsis of noun is extremely frequent with the class *kā* N-V class (21) as well as N-V and NV classes (22):

(21)	<i>merā</i>	<i>intazār</i>	<i>karoge?</i>	<i>-zarūr</i>	<i>karūṅgā</i>
	1SG.GEN	waiting	do.FUT.2	surely	do.FUT.1M.SG
		‘Will you wait for me? - Of course I will’			

(22)	<i>tumhē</i>	<i>mahsūs</i>	<i>nahī</i>	<i>huā?</i>	<i>mujhe</i>	<i>hotā hai</i>	<i>to</i>	<i>tumhē</i>	<i>kyō</i>	<i>nahī</i>	<i>huā?</i>
	1SG.DAT	feeling	NEG	was	1SG.DAT	be PRS	so	2.DAT	why	NEG	was
		‘You did not feel it? I do, so how (come) you didn’t?’ (Contemporary theatre: Vaid)									

### 2.3.4. Specifiers and modifiers

Adverbial and adjectival modifiers are interpreted as bearing on the CP as a whole and not on its first component only, in all types. But whereas *bahut* freely modifies all classes, the adjectives *baṛā* ‘big’, *sahī* ‘true’, *pūrā* ‘full’, occur only with classes *kā* N-V and N-V (23-24). They are ruled out with the NV type (25), an example from Hook (1974), although one of my informants accepted it.

(23)	<i>hamē</i>	<i>baṛī</i>	<i>khuśī</i>	<i>huī /</i>	<i>baṛā</i>	<i>mazā</i>	<i>āyā</i>
	1PL.DAT	great	happiness.F.SG	be.PFV.F.SG	great	pleasure.M.SG	come.PFV.M.SG
		‘We were very happy / enjoyed a lot’					

(24)	<i>māṭ</i>	<i>āne</i>	<i>kī</i>	<i>pūrī</i>	<i>kośīś</i>	<i>karūṅgā</i>
	1SG	come.INF	GEN	full.F.SG	effort.F.SG	do.FUT.1SG
		‘I will try hard/ do my best to come’				

(25)	a	<i>cor</i>	<i>ne</i>	<i>apnā</i>	<i>dimāg</i>		* <i>sahī</i>	<i>istemāl</i>	<i>kiyā</i>	
		thief	ERG	REFL	brain		true	use	do. PFV	
	b	<i>cor</i>	<i>ne</i>	<i>apne</i>	<i>dimāg</i>	<i>kā</i>	<i>sahī</i>	<i>istemāl</i>	<i>kiyā</i>	
		thief	ERG	REFL	brain	GEN	true	use	do. PFV	
		‘The thief made a good use of / used quite well his brain’ (from Hook 1974)								

Quantifiers varying as adjectives are similarly ruled out in the NV class and rejected by all informants whereas the adverbial type (*bahut*) is allowed (26). The indefinite used for high degree with negated verbs (“not even”) are similarly ruled out in the NV class (27a), whereas they freely occur in the alternating *kā* N V construction in (27a):

(26)	<i>in</i>	<i>dinō</i>	<i>bacce</i>	<i>bahut /</i>	* <i>baṛī /</i>	* <i>zor-kī</i>	<i>yād</i>	<i>āte haī</i>
	these	days	child.M.PL	much	great	pang-of	memory	come PRS.3M.PL
	‘These days [I] remember the children a lot’ (I miss the kids)							

(27)	a	<i>mujhe</i>	<i>iskā</i>	<i>koī</i>	<i>patā</i>	<i>bhī</i>	<i>nahī</i>	<i>thā</i>
		1S.DAT	this.GEN	INDEF	knowledge	even	NEG	was
		‘I did not even know that’						

(27)	b	<i>mujhe</i>	<i>yah</i>	* <i>koī</i>	<i>patā</i>	<i>bhī</i>	<i>nahī</i>	<i>thā</i>
		1S.DAT	this	INDEF	knowledge	even	NEG	was

To sum up, constraints on modifiers, extraposition, omissibility, and separability apply very differently to the class NV and the two others, whereas differences are far less clear between these last two types, except for separability.

The three classes of CP described above, whatever their internal structure, constitute an original contribution to the renewal of a pan-Indian feature, the existence of cognate transitive/intransitive pairs of verbs. The following sections are devoted to the role played by such predicates in the coherence of the Indian linguistic area, first in systematizing the correlation between transitive and intransitive predicates, then in the emergence of new argument structures and new (neither nominative nor ergative) sentence alignments.

### 3. Complex Predicates in the typological characterization of the Hindi predicate lexicon

Cognate verbs, or morphologically related verbs, are a well-known feature characterizing the Indian area as a linguistic area (Emeneau 1980). Old Indo-aryan had a triple voice alternation (active, middle, passive) correlated with a two-fold paradigm of personal endings for each tense. The alternation between active, middle and passive, and the so-called “causative derivation” required a phonological alternation in the verbal base, traditionally described as a shift in vowel degree. The zero degree is correlated with the middle (or passive) basis, the normal degree with the active base and the augmented degree with the causative. For instance, the verbal root KR “do” has three correlated bases: *kr* (zero), *kar* (median or “qualified”), *kār* (augmented). But consonant alternations can also occur, particularly with retroflex consonants. Although not all Hindi intransitive verbs are inherited from the Sanskrit base with low degree and not all transitive from the one with middle or high degree, the general logic is maintained in simplex indigenous verbs, the non agentive intransitive verb having a different vowel than the agentive transitive. The system of vowel/consonant alternation is no longer transparent nor productive and verbs are usually perceived as two lexically distinct units, although correlated, according to the terminology used to describe this pan-Indian feature. It is easy for instance to identify the augmented degree of the lengthened *ā* in *chāpnā* “to print” compared with *chapnā* “to be printed”, or in *nikālnā* “to put out,

extract” compared with *nikalnā* “to go out, ooze”, in *mārṇā* “to kill” vs *marnā* “to die”, or in *pīṭnā* “to beat” vs *piṭnā* “to be beaten”, but less easy in *dekhnā* “to see, look”, compared with *dikhnā* “to be visible”, or in *roknā* “to stop” (tr.) compared with *rukṇā* “to stop” (intr.), not to speak of pairs with different consonants (*biknā* “to sell/get sold” vs *becnā* “to sell” (tr.), *ṭūṭnā* “to break/get broken” (intr.) vs *torṇā* “to break” (tr.)). In a number of cases, the transitive base is formed by the addition of the *-ā* suffix, like *cal* “to walk/go”, *calānā* “to make walk, drive”. As opposed to these irregularities, further causativity (make do) is systematically obtained by the *-vā* suffixation, a recycling of the Sanskrit causative derivative (*-apa/apaya*)<sup>9</sup>. Besides, some roots may have up to four forms (*dikhnā* “to be visible”, *dekhnā* “to see/look”, *dikhānā* “to show”, *dikhvānā* “make show”, some only one (*jācnā* “to check, verify”), many lack the intransitive counterpart (*bhejnā* “to send”, *khānā* “to eat”, *sunṇā* “to hear/listen”), the intransitive and double causative (*khojnā* “to look for, search”, or the transitive one (*jānā* “to go”). And a few are labile verbs, with both agentive and non agentive meanings (*bharnā* “to fill”, tr. and intr., *baṛhānā* “to augment”, tr. and intr.).

### 3.1. The complex predicate as a means to simplify and systematize the traditional correlation

In front of this complicated and now unpredictable system, which is no longer productive, the complex predicates offer an easy way for producing the paired forms by simply alternating the light verb. The most common alternation is between *kar* “do” and *ho* “be”: *intazām (prabandh) honā* “to be organized”, *intazām karnā* “to organize”; *tabāh (naṣṭ) honā* “to be destroyed”, *tabāh karnā* “to destroy”, *hatyā (khūn) honā* “to be killed”, *hatyā karnā* “to kill”, *jāc honā* “to be checked”, *jāc karnā* “to check”, *kharā honā* “to stand”, *kharā karnā* “to make stand”, *ārambh (śūrū) honā* “to begin” (intr.), *ārambh karnā* “to begin” (tr.), etc.<sup>10</sup>.

Hence a considerable proportion of equipollent predicates in the Hindi lexicon, that is, predicates which can be either transitive or intransitive according to the verbal affix or auxiliary added on the basis (Haspelmath 1993)<sup>11</sup>. Thus the CP morphology not only strengthens the indigenous tendency of deriving two cognate transitive/intransitive verbs, it also provides for the “empty” slots existing in the defective morphology of simple verbs: for instance *khojnā* “to search for, inquire” and *dhūṇḍhnā* “to search” which have only one form as well as the Persian synonym *talāsnā*, are suppleted by the complex predicates *khoj+* light verb with two forms (tr. *khoj karnā*, intr. *khoj honā*) and *talaś karnā/honā*. Similarly, *batānā* “to tell” and *bolnā* “to speak” have only one form whereas *bāt karnā* (tr.) pairs with *bāt honā* (intr.). Other verbs which have only the active transitive form get a suppletive non agentive meaning by means of a CP, like *sunṇā* “to listen” (*sunāī denā* “be audible”), *jānnā* “to know” (*patā (mālūm) honā* “to be known”)<sup>12</sup>. Other light verbs than “be” and “do” are also available, particularly for idiomatic predicates such as *khānā* “to eat” (intr.) vs *denā* “to give” (tr.), with the same voice alternation, “eat” involving a patient, “give” an agent<sup>13</sup>.

<sup>9</sup> The vowel flexion is called a “degree increase” in the Sanskrit grammatical terminology (*guṇa* “quality” and *vridhhi* “augmentation”). On the controversial question of the sense of the “derivation”, arguments for basic transitivity (meaning of the bare noun sometimes active like *rok* “a stop”) are fewer than those in favour of basic intransitivity (heavier form of the transitive verb, intransitive form used for further derivations such as the “double causative” (*rukṇānā* “to make stop”, existence of nouns also from the intransitive base), cf. Montaut (2011).

<sup>10</sup> In the list, synonyms from a different origin, Persian, Arabic or Sanskrit, are given in brackets.

<sup>11</sup> The third class proposed in Haspelmath & Comrie’s typology, the labile verbs, is almost non-existent in Hindi and concerns two or three verbs. The recognition of an equipollent class in Hindi ith CPs, for some of the 40 basic notions in the language sample the WATP (Pardeshi 2014), suggests an analysis of the noun-verb CP where the light verb behaves as a verbal affix or an auxiliary.

<sup>12</sup> CP with only one voice (*būkh honā* “be hungry”) do exist but they are far less numerous than simplex predicates with only one form, as noted by Gambhir (1993).

<sup>13</sup> In idiomatic expressions involving undergoing of a generally negative process, the verb « eat » (*khānā*) corresponds to the passive, and its active agentive counterpart uses verb « give » (*denā*): *dhokā khānā* (fraud/betrayal eat) « to be deceived, fooled », vs *dhokā denā* (fraud/betrayal give) « to fool, deceive sbd »; *khoyī khānā* (whip eat) « to be whipped », vs *khoyī denā* (whip give) « to whip sbd ».

This re-creates in the language a clear opposition between semantically active and middle or passive verbal bases. Whereas formerly the distinction was morphologically complex and the sense of the derivation unclear (although more arguments favour the primary intransitivity of simplex basis), the Noun Verb predicate allows for a clear and easy system, hence its unbroken productivity since about four centuries. It contributed to even more relegate the class of labile verbs to an insignificant role, making the general system more pregnant and in this way typologically more relevant. If Hindi could be considered as a language with primary intransitivity (Pardeshi 2014) on the basis of its simplex verbs, on the basis of the whole predicate lexicon, it is equipollent. This means that the much controverted direction of the derivation (intransitive derivation as in most west-European languages or transitive derivation as in Caucasian languages, cf. Comrie & Haspelmath 1993) emerges as a weakly relevant typological feature. The question is not whether the language privileges a passive or middle orientation with primary intransitivity or an active orientation with primary transitivity: both are treated on equal terms, because of the recent dominance of equipollent predicates in the lexicon.

The importance of complex predicates in Hindi in particular and other Indo-Aryan languages in general is indeed quite recent – their occurrence in Sanskrit is as marginal as it is in English for instance. They began to proliferate in Hindi during the Moghol empire (16-18<sup>th</sup> c.) when northern India was ruled by a dynasty from turko-mongol origin and Persian became the language of culture and administration. After the three earlier centuries of military raids, partial occupation, and Sufi presence in the North-West from 12<sup>th</sup> to 15<sup>th</sup> c., during which contacts with the population were limited, the establishment of the Empire resulted in a widespread language interaction, and, consequently important borrowings. As noted by Gambhir (1993: 83), “both communities needed to develop linguistic channels to communicate with each other”, an interaction attested “by the large number of borrowings in the areas of administration, war and law”. As verbs are less easy to borrow than other parts of speech (Moravcsik 1975), they are most of the time borrowed as nouns or with the addition of a derivational morphology – although more recent and extensive surveys like Wolgemuth (2009) also find verb stems borrowed with no modifications. In Hindi, Persian and Arabic verbs were systematically borrowed from the language of the new rulers as nouns or participles used with light verbs. For instance the participles or invariable verbal constituents *qatl* “kill”, *qabūl* “accept”, *mālūm* “know”, *manzūr* “agree” from Arabic, *band* “close”, *bardašt* “tolerate”, from Persian, are till now only used with light verbs, whereas nouns such as *intazām* “organisation” or *intazār* “waiting”, from Persian are also used as independent nouns. Among the reasons given for this recategorization of foreign verbs and the success of the new construction over the indigenous simplex verb is the fact that “there was no complicated phonotactic changes in the borrowed element for deriving various verb forms”, whereas the indigenous “causative derivation” involves “complex phonotactic rules and irregularities” (Gambir (1993: 83).

Till to-day, foreign verbs are always recategorized in order to be borrowed, mainly from Sanskrit (particularly in the technical neology) and English, and the recategorization as noun is the more frequent: for the predicate “to try” the complex predicates borrowed from Sanskrit (*prayatn karnā* effort do) and Persian (*kośiś karnā* effort do), which supplanted the indigenous simplex verb long ago, are now paralleled by the English borrowing *trāī karnā*, which displays internal agreement, either as a N V predicate or a *kā* N V predicate, with feminine agreement by analogy with *kośiś*, a feminine noun:

(28)	<i>mañne</i>	<i>kām</i>	<i>pūrā</i>	<i>karne</i>	<i>kī</i>	<i>kośiś</i> /	<i>kī</i>	<i>trāī</i>	<i>kī</i>
	1SG.ERG	work	complete	do.INF	GEN.F.SG	effort	GEN.F.SG	try	do.PFV.FS
	‘I tried to complete the work’								

### 3.2 An enriching extension of the transitive/intransitive correlation: aspect and Aktionsart

It has been noted in several studies, particularly Davison (2005) and Montaut (2004, 2011), that the light verb resembles auxiliaries by conveying tense and aspect, as well as voice alternation. The

alternation of *karnā* “do” (agentive) and *honā* “be” intransitive (non agentive) light verbs obviously evokes voice alternations (cf. section 3.3), but the choice of different light verbs also modifies aspect and Aktionsart, when several are available for the same predicative notion. It may simply parallel the grammatical system of aspectual auxiliaries which can combine with any simplex verb: the use of *rahnā* “to stay” as an intransitive light verb or *rakhnā* “to place, keep” as a transitive light verb produces the same aspectual meaning as the auxiliary *rahnā* with a participial verbal basis (durative: “keep doing/being”). Changing *denā* “to give” into *rakhnā* “to put” as the unmarked light verb for *dhyān* “attention” (pay attention) makes it a durative with the meaning “keep in mind” (plus direct object). But the availability of *lagānā* “to place, stick” in a similar function makes the subject’s intentionality more pregnant than the object (prepositional), with the meaning “concentrate/maintain or strengthen one’s mindfulness (on)”. *Dhyān lagānā* is also compatible with the durative auxiliary *rahnā*, as in example (29), a program of the US army significantly coined “mind fitness based on mindfulness”:

(29)	<i>dhyān</i>	<i>lagā</i>	<i>rahī</i>	<i>Amerīkā</i>	<i>senā</i>
	attention	place/stick	DUR.PFV.F.SG	America	army.F.SG
	‘American army kept concentrating’ (webdunya 23 June 2015)				

Similarly, the predicate *ḍar honā* “to be afraid” can combine also with the verb *lagnā* (“be placed/touch”), which makes it inchoative as does the inceptive auxiliary *lagnā* for simple verbs, and *ḍar lagānā* can additionally combine with the inceptive auxiliary (30):

(30)	<i>ḍar</i>	<i>lagne</i>	<i>lagā</i>	<i>apne</i>	<i>hī</i>	<i>pyāre</i>	<i>gharaūde</i>	<i>mē</i>
	fear	be.placed/touch	INC.PFV	REFL	FOC	dear	house	in
	‘(We) started becoming afraid in our own beloved houses’ (Amarujala 23 June 2015)							

Whether such subtle differences better accounted for in terms of aspect or Aktionsart, is a question for further reflexion, particularly in the case the light verb belongs to the class of vector verbs in VV constructions, a semi-auxiliarized class of motion verbs conveying perfectivity and « attitudinal meanings » (Abbi & Gopalkrishnan 1992). For instance, the complex predicate *dikhāi denā* (visibility give) « be visible, appear », or *nazar ānā* (view come) with the same meaning, can both combine also with light verb *paṛnā* « to fall ». The verb *paṛnā* as a vector verb is used with intransitive verbs to convey both perfectivity and suddenness. As a light verb, not only does it convey suddenness, but it can itself combine with a vector verb like « go » (31a) and with the progressive marker (31b), in contrast with the vector verb which cannot:

(31)	a	<i>acānak</i>	<i>vah</i>	<i>phisal</i>	<i>gayā</i>	<i>aur</i>	<i>unkī</i>	<i>brest</i>	<i>dikhāi</i>	<i>paṛ</i>	<i>gaī</i>
		suddenly	3.SG	slip	go.PFT	and	her	breast	visible	fall	go.PFT
		‘It (her shawl) suddenly fell and her breast became visible’ (Bollywoodtakda)									

(31)	b	<i>jo</i>	<i>dikhāi</i>	<i>paṛ-rahā-hai</i>	<i>vah</i>	<i>māyā</i>	<i>hai</i>
		REL	visible	fall-PROG-PST.3SG	3.SG	illusion	be.PST.3SG
		‘What is immediately appearing is illusion’					

If *lagnā* is the most usual light verb for conveying inchoation with host nouns, such as « hunger », « thirst », « fear », etc., *ānā* « to come » also occurs with others, such as « memory », « thought » in lexicalized combinations (some notions require *ānā* and other *lagnā*, with practically no interchangeability). The verb *ānā* also conveys a vague notion of internalization, and inchoation itself may be over-marked by the additional aspect marker *lagnā* for inception:

(32)	<i>mujhe</i>	<i>ve.hī</i>	<i>yād</i>	<i>āne</i>	<i>lage-hai</i>	<i>jinhē</i>	<i>bhūlne</i>	<i>mē</i>
	1.SG.DAT	3M.PL.FOC	memory	come	INC-PFT	REL.ACC	forget	in

	'I started remembering those whom I had been forgetting
--	---

<i>zamāne</i>	<i>lage-hai</i>
periods	INC-PFT
for years' (famous ghazal song)	

This means, as in (30), that the substitution of an inchoative light verb to the neutral *honā* “be” conveys something more than inception. Yet when contrasted with the neutral CP, *yād honā* (state) or *ḍar honā* (state), *yād ānā* or *ḍar lagnā* conveys the notion of entry in the state: aspectual distinctions, or for the matter Aktionsart distinctions, which both can be reconciled in the same logic (Sasse 2001, 2006)<sup>14</sup>, are richer and more flexible with complex predicates since they can additionally use the regular paradigm of aspect markers.

### 3.3 Valency correlates

The opposition between transitive and intransitive predicates structures the verb lexicon whereas simplex or complex predicates (in a more systematic way as for complex predicates). The intransitive paradigm behaves as a substitute for passive and middle, which makes the morphological passive quite marginal in Hindi and always associated with the emphasis on the agent volitionality. What is in Romance and Russian languages expressed by the reflexive voice is expressed in Hindi by intransitive verbs, whether it relates to medio-passive, middle meanings or to spontaneous processes, pertaining to what Kemmer (1993) encapsulates with the middle category. The enrichment of the predicative lexicon by means of complex predicates not only contributed to centre stage the already existing feature of cognate verbal pairs for transitive/intransitive, it also allowed sentences to background all the arguments, since the noun in the class of CP with internal agreement, by far more numerous than the one with external agreement, can behave as a quasi argument.

As expected, the transitive agentive predicate (33a) has one argument more, the agent, than the intransitive (33b), a similar valency alternation as for simplex verbs:

(33)	a	<i>dakuō</i>	<i>ne</i>	<i>do</i>	<i>masūm</i>	<i>baccō</i>	<i>ko</i>	<i>qatl</i>	<i>kiyā</i>
		bandit.M.PL	ERG	two	innocent	child.M.PL	ACC	murder	do.PFV.M.SG
		‘The bandits murdered two innocent children’							

(33)	b	<i>do</i>	<i>masūm</i>	<i>bacce</i>	<i>qatl</i>	<i>hue</i>
		two	innocent	child.M.PL	murder	be.PFV.M.PL
		‘Two innocent children were murdered’				

With the *kā* N-V type the agent is similarly backgrounded (omitted) and the patient is demoted since it appears in the genitive case, which never happens in the case of simplex verbs (cf 4):

(34)	a	<i>atākvādiyō</i>	<i>ne</i>	<i>das</i>	<i>ādmīyō</i>	<i>kī</i>	<i>hatyā</i>	<i>kī</i>
		terrorist.M.PL	ERG	ten	man.M.PL	GEN	murder. FS	do.PFV.F.SG
		‘The terrorists killed ten people’						

(34)	b	<i>das</i>	<i>ādmīyō</i>	<i>kī</i>	<i>hatyā</i>	<i>huī</i>
		ten	man.M.PL	GEN	murder. FS	do.PFV.F.SG

<sup>14</sup> Aktionsart is a category used mainly in Germanic and French grammatical traditions for modes of action such as iteration, continuation or other distinctions which belong to lexical aspect in many languages (Sasse 2006). Although cognitively not distinct from aspect (Sasse 1991), it can help describing verbal categories in languages with a rich paradigm of aspectual markers such as Hindi, some of which pertaining to lexical aspect.

		‘Ten people were killed’
--	--	--------------------------

Furthermore, the patient too may be omitted: *qatl huā* “it was murdered”. Sentences such as (35) are very frequent in the language, omitting all the participants, even in the case of a complex predicate with external agreement such as *manzūr honā* “accept” in (35b):

(35)	a	<i>jāc /</i>	<i>talās</i>	<i>ho-rahī-thī</i>
		inspection. F.SG	search. F.SG	nbe-PROG-IMPFT.F.SG
		‘It was being inspected / searched/discussed’ (one was/they were inquiring, searching)		

(35)	b	<i>bāt</i>	<i>huī,</i>	<i>māzūr</i>	<i>huā</i>
		speech.F.SG	be.PFV.FS	accepted	be.PFV.M.SG
		‘[I] spoke with [them], they accepted’ (lit. ‘It was spoken, was accepted)			

If (35a) requires a context for recovering the missing arguments, (35b) does not, and the frequency of such minimal sentences, similar to the impersonal passive which occurs in German (*es wird getanzt* “it is danced”) or the impersonal French “on”, is by far higher than the impersonal passive in Hindi devoid of agent and patient, which is highly restricted and usually requires the recoverability of the main argument.

#### 4 The role of Complex Predicates in remodelling alignments

Apart from the argument loss described above, valency change can affect, not only the number but the semantic role of arguments, as well as their case marking.

##### 4.1 The dative experiencer: semantic alignments

A considerable proportion of the predicative notions in complex predicates do not gain an argument when combining with the active transitive light verb – or do not lose an argument when combining with the stative light verb: all the nouns which are not clearly nouns of action, and consequently do not allow a clear opposition Agent Patient, like “search”, “like”, “doubt”, “feel”, “complain”, maintain the same number of external arguments with both light verbs. The stative light verb simply licenses a dative experiencer in place of the agent licensed by the active light verb:

(36)	a	<i>mujhe</i>	<i>yah</i>	<i>film</i>	<i>bahut</i>	<i>pasand</i>	<i>āī</i>
		1SG.DAT	this	film	much	taste	come.PFV.F.SG
		‘I liked this film very much’					

(36)	b	<i>maīne</i>	<i>yah</i>	<i>film</i>	<i>bahut</i>	<i>pasand</i>	<i>kī</i>
		1SG.ERG	this	film	much	taste	do.PFV.F.SG
		‘I liked this film very much’					

Example (7) above, with *śak* “doubt”, exhibit a similar alternation in argument structure, extremely widespread in modern Hindi. The agentive version, *śak karnā* “to doubt”, may be interpreted as more deliberate than the experiential counterpart *śak honā* “to have doubt”, and the feature conscious activity sometimes accounts for the preference of an active light verb: in specific contexts, *pasand karnā* can mean “to choose” (besides “like”) and *yād karnā* “to convoque/call” (besides “remember”), which is never a possible extension of meaning for the inagentive counterpart<sup>15</sup>. Similarly, *talās karnā*, “to research”, a predicate whose noun refers neither to a

<sup>15</sup> Similarly *pyār* “love” or *viśvās* “trust” convey the contrast between a deliberate process in the agentive construction (in the passive of transitive *kar* “do” (a) and a non deliberate process in the middle construction with *honā* (b):

prototypical action nor to a prototypical experiential state, involves a more active participation of the subject, with the agentive light verb, Agent and Patient (37a) than the experiential construction (37b) where the subject is represented as only affected by the process, in a kind of neglectful involvement, with the intransitive light verb, Experiencer and Theme in<sup>16</sup>:

(37)	a	<i>pulis</i>	<i>corō</i>	<i>kī</i>	<i>talās</i>	<i>kar-rahī-hai</i>
		police.F.SG	thief.M.PL	GEN	search	do-PROG-PRS.3F.SG
		‘The police is searching for the robbers’				

(37)	b	<i>mujhe</i>	<i>tum</i>	<i>donō</i>	<i>kī</i>	<i>talās</i>	<i>thī</i>
		1.SG.DAT	2	two	GEN	search	be.IMPFT.F.SG
		‘I was in search of you both’ (film Shole 1976)					

Whereas the alternation between transitive-active and intransitive-passive alternation, which involves the loss of one argument, occurs also with simple verbs, the experiential pattern occurs practically only with complex predicates in modern Hindi: only half a dozen simple verbs allow dative experiencers, and only three of them are usual, *milnā* ‘to get/find’, *lagnā* ‘to seem’ and *ānā* ‘to come’ when it has the specific meaning of ‘to know’ (a language, a technique)<sup>17</sup>. The massive presence of the so-called ‘dative subjects’, one of the characteristic features of the South Asian as a ‘linguistic area’ (Emeneau 1980), is to be correlated with the massive presence of CPs in the predicate lexicon. ‘Dative subjects’ were marginal in ancient languages such as Sanskrit (Hock 1991) or old Tamil, and in Indo-Aryan they are relatively recent, just starting to appear in the 14<sup>th</sup> c., not yet systematic before the 16<sup>th</sup>- century (Montaut 2013), that is to say, their proliferation in the language coincides with the proliferation of Complex Predicates as a consequence of verb borrowings from Persian.

The consequence of a massive borrowing in the change of argument structure has been 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’. When contact with Persian speakers occurred, the rising construction was the dative-subject for experiencers, mainly with the verb *mil*, massively used in the context of popular devotion for expressing the fusional meeting with the deity and experience of true revelation. The lexical renewal of the predicate lexicon caused attraction to the construction on the rise, hence its extension to more and more experiential predicates. Just before the massive borrowing in the form of CPs, the experiencer dative subject was almost restricted to the verb *milnā* ‘meet, get fused’: this verb makes the bulk of the experiential constructions in Kabir, a mystic poet from the 14<sup>th</sup> century, with only very few isolated occurrences of other verbs (*bhānā* ‘please, like’). Since this use pattern became a representative idiom of the new devotional mystic – ‘meet’ the deity to get fused with it or the absolute principle, and receive or experience the revelation of true knowledge –, was becoming a central cultural axiom, its linguistic expression with the verb *milnā* attracted the newly borrowed predicates when they were experiential predicates (Montaut 2013).

---

a. *viśvās /pyār kiyā nahījātā,*                      b. *ho jātā hai*  
trustlove do NEG PASS.PRS                      be go PRS

‘Trust/love is not done, it becomes/happens = trust/love cannot be commanded, one does not chose to love/trust’

<sup>16</sup> Besides the passive-like construction with no agent and no experiencer:

. *corō kī talās ho rahī hai*  
thief.M.PL GEN search be PROG PRS.3F.SG  
‘The robbers are being searched’

<sup>17</sup> Two of them also occur as light verbs, namely *lagnā* (*ḍar lagnā* ‘be(come) afraid’) and *ānā* (*pasand ānā* ‘please’) but with different meanings and behaviour.



Significantly the concurrence between simplex, when still available, and complex predicates favours till now the complex predicate: the verb *dikhnā* “to appear” is to a large extent replaced by the CP *dikhāī denā* or *nazar ānā* (view come), and similarly *bhānā* “to please” is now almost out of use and replaced by the CP *pasand honā*. The CP *vicār* or *khyāl* (“thought”) *ānā* “to think of” has now far more currency than the simplex verb *sūjhnā* in the same meaning. Moreover, certain simplex verbs, which do not allow the dative subject construction, are recategorized as complex predicates which can license dative experiencers: *ḍarnā* “to fear” for instance has now less currency than the CP *ḍar honā/lagnā*. Far less frequent, and still non-standard, is the remodelling of a nominative simplex verb into an experiential simplex verb as is the case for *samajhnā* “to understand”, or *bhūlnā* “to forget” in Panjabised Hindi. This clearly indicates that the Complex Predicate is felt as the best tool for shifting a nominative argument structure to a dative experiential one.

If the experiential construction with a dative subject is now considered as one of the major pan-Indian features, it should be emphasized that, at least as far as the Indo-Aryan languages are concerned, its extension is intimately correlated with the extension of CPs. This new alignment is all the more interesting since it is a semantic alignment, as opposed to the alignments prevailing till then (nominative and ergative alignments). As noted in the majority of relevant studies, this argument structure is constrained by the lexical semantics of the predicate and not by grammatical parameters, and this way of mapping the linguistic material according to semantic and not syntactic constraints, is typical of semantically aligned languages (Donohue & Whichmann 2008) or, as they were previously called, “active languages”. The fact that a good proportion of the alignments in a language are semantically constrained does not mean that the language as a whole is semantically aligned, but a growing proportion of such alignments is typologically relevant regarding alignment, a typological shift which cannot, for the matter, be dissociated from the lexical renewal in the form of complex predicates.

#### 4.2 The genitive “possessor”

Complex predicates can also licence a genitive subject without losing an argument, a type of subject marking which is unique in finite clauses except possessive sentences with verb “be”<sup>18</sup>.

- (38) *merā man (ghūmne ke.lie/ko) nahī̃ hai, tumhārā jī kartā hai ki*  
 1SG.GEN spirit walk for/to NEG be.PRS.3SG 2SG.GEN spirit do.PRS.3SG that  
 ‘I don’t feel like taking a walk, you wish that...’

This type of subject demotion is in conformity with a general tendency of the language which consists in distancing certain subjects, particularly when subjective states are at stake: *merā man/jī kartā hai* (my spirit/soul does) “I wish”, *uske dil mē* “in his heart = he”, *uskī ākhē bharne lagī* (her eyes fill started) “she started crying” etc. In this way, the genitive marking of objects and of intransitive subjects, recategorized as mere expansions of the action noun of the predicate, represents a demotion of the main argument. This converges with the obliteration of the first or second person in idiomatic expressions instead of the direct expression of person for reasons of etiquette rules (“your slave”= “I”, “X’s Majesty, X’s Honour, X’s Grace” = “you”).

With Complex predicates, genitive subjects in two participants clauses generally occur with very weakly dynamic notions such as “ability”, “habit”, “responsibility”, “lack”, “experience”, which correspond to the category of lower transitivity in Tsunoda’s (1981) scale of transitivity. Some of these CPs do not have an agentive counterpart with the light verb “do”, some have it but marginally like *irādā* “intention” (39), and some others present an alternation dative/genitive, such as *icchā* “desire”, both with the stative light verb (40):

<sup>18</sup> Example: *mere do bhāī haī* (of.me two brothers are) “I have two brothers”. In non finite clauses with infinitives or participles, when the subject is distinct from the main subject, it is however expressed in the genitive: *bhāī ke āne par, maī* (brother GEN come.INF on 1SG) “when my brother arrived, I..”

(39)	a	<i>merā</i>	<i>parsō</i>	<i>jāne</i>	<i>kā</i>	<i>irādā</i>	<i>hai</i>
		1SG.GEN	day.after.tomorrow	go.INF	GEN	intention	be.PRS.3SG

(39)	b	<i>maĩ</i>	<i>parsō</i>	<i>jāne</i>	<i>kā</i>	<i>irādā</i>	<i>kar rahā hũ</i>
		1SG	day.after.tomorrow	go.INF	GEN	intention	do.PRS.1SG
		'I intend to go day after tomorrow'					

(40)	<i>merī</i>	<i>mujhe</i>	<i>cāy</i>	<i>pīne</i>	<i>kī</i>	<i>icchā</i>	<i>nahĩ</i>	<i>hai</i>
	1SG.GEN	1SG.DAT	tea	drink.INF	GEN	desire	NEG	be.PRS.3SG
		'I don't feel like drinking tea'						

Example (40) shows the affinity of the genitive subject construction with the experiential construction in the dative, the latter emphasizing more the subject's affectedness and transience of the state, whereas the genitive pattern emphasizes more the stative nature of the process, represented as an inherent property of the subject (rarely used in the eventive). Both involve predicative notions which are semantically very low in transitivity, and for some the alternation is constrained by the lexical choice without any difference in meaning: for instance, *abhāv honā* (lack be) "to lack" requires a genitive subject whereas *kamī honā* (lack be) requires a dative subject.

But there are also action processes, higher on the transitivity scale yet not in the highest position, which can alternate an agentive nominative construction and a genitive one, while retaining both their arguments, a shift which has not been studied yet. This alternation is particularly frequent with the CPs formed on nouns involving contact, relation, such as *śādī* "marriage", *sampark* "contact", *bāt* "speech", with the second participant in the sociative case and the first one either in the nominative (41a) and ergative if perfective (42a) or in the genitive (41b, 42b):

(41)	a	<i>maĩ</i>	<i>tumse</i>	<i>sampark</i>	<i>zarūr</i>	<i>karūgā</i>
		1SG	2.SOC	contact	surely	do.FUT.1SG
		'I will definitely contact you'				

(41)	b	<i>un</i>	<i>logō</i>	<i>se</i>	<i>sampark</i>	<i>nahĩ</i>	<i>hai</i>
		1SG	people	SOC	contact	NEG	be.PRS.3SG
		'I have no contact with these people'					

(42)	a	<i>maĩne</i>	<i>dāktar</i>	<i>se</i>	<i>bāt</i>	<i>kī</i>
		1SG.ERG	doctor	SOC	speech.F.SG	do.PFV.F.SG
		'I spoke to the doctor'				

(42)	b	<i>merī</i>	<i>dāktar</i>	<i>se</i>	<i>bāt</i>	<i>huĩ</i>
		1SG.GEN	doctor	SOC	speech.F.SG	be.PFV.F.SG
		'I spoke to the doctor'				

One cannot say that the genitive main argument of *bāt* is changed from agent to experiencer, or to a non agentive actor of the process. The genitive case marking is less semantic than discursive and fits a general tendency of the language – to avoid the foregrounding of the agent, particularly the first person, in conformity with rules of etiquette and social behavior, which contributed to shape the polite language during the Mogol empire (cf. *supra*)<sup>19</sup>.

<sup>19</sup> It is for instance deemed more polite to use a periphrastic with genitive argument instead of the nominative expression perceived as ruder:

Apart from these two types of genitive alignments which both maintain two external participants, another sentence pattern also involves a genitive subject: when a *kā* N V predicate shifts from the agentive to the non agentive passive like construction, by shifting verb “do” to verb “be”, the first participant appears in the genitive. This is the case in examples (11b), (21a) or (34) above: “the land was used” (land of use was), “men were murdered” (murder of men was). In such cases, the object in the agentive sentence is case marked in the same way as the subject of the non agentive sentence: *ādmīyō kī* “of men” is the object of the transitive statement with “murder do” and the subject of the intransitive statement with “murder be”.

The identical case marking or construction for subject and object is a well known feature of semantically aligned languages, and represents what Durie (1988) called the “preferred argument structure in active languages”, with examples from Acehnese such as (45) with *gopnyam* “he” (cliticized 3 *geu/geuh*) appears in the same enclitic position (Undergoer) as subject of inagentive intransitive and object of agentive transitive verb, and *lôn* “I” in the same proclitic position (Actor) as subject of active intransitive and subject of agentive transitive verbs:

(43)	a	<i>lôn</i>	<i>teungöh-lôn-jak</i>
		I	MIDDLE-1-go
		‘I am walking’	

(43)	b	<i>gopnyan</i>	<i>galak-geu</i>	<i>that</i>
		he	happy-3	very
		‘He is very happy’		

(43)	c	<i>gopnyan</i>	<i>na-lôn-timbak-geuh</i>
		he	be-1-shoot-3
		‘I shot him’	

The fact that in Hindi both subjects and objects can now be marked in the genitive is perhaps only a morphological consequence of the lexical nature of the relevant CPs (with patient related to the action noun by the genitive marking): according to Gambhir (1993), this class of Complex Predicates is the most productive to-day, and the shifts observable since Kellogg established the first systematic lists of NV constructions in 1875 are mainly towards the *kā* N type, which may also attract predicates borrowed from English (cf. supra *kī trāy karnā* “to try”, example (30)). Whatever the causes which brought such similarities (S= A for action processes, P for non action processes), the fact is that now most CPs licence the availability of case-marking Patients object of transitive verbs and Patient subjects of non active verbs in the genitive, a case which Nichols (2008) calls to closer attention for identifying the so-called active languages or alignments. She questions the general assumption that semantically aligned languages are essentially found in Austronesian and Amerindian languages, but not in Europe nor in South Asia, on the basis of potentially misleading expectations (that, for instance, the main argument of “non-active” statements be strictly coded like an object – accusative—). She mentions that genitive marking, which exists in Basque, Georgian and Kartvelian languages, should also be taken into account. The genitive marking for certain subjects and objects, along with the fact that the marked accusative is identical in Hindi with the dative, making “dative subjects” and specific objects identically case-marked, may not be sufficient to speak of a semantically aligned language, but the role of CPs in licensing non-canonical subjects and non-canonical alignments cannot be questioned.

#### 4.3. Semantic roles and subjecthood

---

<i>unkā dehānt huā</i>	<i>vah mar gayā</i>	
3P.GEN body-end was	3SG die went	‘He died’

As a correlate to these typological changes in argument structure and alignment, semantic roles came to gain more and more importance, and new distinctions to appear. Even when the specific non-canonical marking already existed in the language with simplex verbs, the CP category extends and systematizes this distinction, in the same way as it systematizes the opposition of correlated pairs of transitive /intransitive. Inadvertent actors are for instance case marked in the instrumental, a construction very commonly used for expressing an involuntary acts, therefore irresponsibility: it can convey the innocence (44b) of a subject accused of having committed a crime (44a):

(44)	<i>tumne</i>	<i>khūn</i>	<i>kiyā</i>
	2.ERG	blood	do.PFV
	'You have killed...'		

But the really new distinction which appeared in semantic roles concerns experiencers. Not only the systematic distinction between nominative agents and dative experiencers coincides with the entry of CPs in the lexicon, but within the very class of experiencers a new distinction arose. For psychological verbs of feeling, the alternation of light verbs (“do” vs “be”) is devoid of the semantic distinction mentioned above, since deliberateness or conscious choice is doubtful with predicates like “hate” or “love” (cf. example (12)) and ruled out with predicates meaning “to feel” or “be jealous” as in (45):

(45)	a	<i>tab</i>	<i>mujhe</i>	<i>tumse</i>	<i>irṣyā</i>	<i>thī</i>	<i>magar</i>	<i>mujhe</i>	<i>iskā</i>	<i>bodh</i>	<i>nahī</i>	<i>thā</i>
		then	1SG.DAT	2SOC	envy	was	but	1SG.DAT	its	awareness	NEG	was
		'At that time I was jealous of you, but I was not conscious of it'										

(45)	b *	<i>tab</i>	<i>maĩ</i>	<i>tumse</i>	<i>irṣyā</i>	<i>kartī thī</i>	<i>magar</i>	<i>mujhe</i>	<i>iskā</i>	<i>bodh</i>	<i>nahī</i>	<i>thā</i>
		then	1SG	2SOC	envy	did	but	1SG.DAT	its	awareness	NEG	was

The difference here is not in the choice of being jealous but of consciously assuming such a state in (45b) with “do” verb and ergative subject, which makes the statement unacceptable with a prolongation such as “but I was not aware”. With such a prolongation, the only possible statement is the experiential one with dative subject and “be” verb (45a). 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 on her complaints, he uses both constructions: the nominative agentive construction (A) adds to the bare perception in the dative of experiential construction (Exp) a dimension of reflexive consciousness, intellectual acknowledgement, beyond simple perception:

(46)		<i>maĩ</i>	<i>to</i>	<i>itnī</i>	<i>begānī</i>	<i>mahsūs</i>	<i>kartī-hū</i>	<i>is</i>	<i>ghar</i>	<i>me</i>	<i>ki...</i>
		1SG	TOP	such	solitude	feeling	do-PRS.1SG	this	house	in	that
		Woman – I feel (A) so much solitude in this house that...									

		<i>-pahle</i>	<i>nahī</i>	<i>kartī-thī ?</i>	<i>- pahle ?</i>	<i>pahle</i>	<i>to...</i>	<i>mahasūs</i>	<i>karnā</i>	<i>hī</i>
		before	NEG	do-IMPF.F	before	before	TOP	feeling	do.INF	just
		Man – You did (A) not before? Woman – before? But before... Man - The fact of feeling (A)								

		<i>mahasūs</i>	<i>nahī</i>	<i>hotā-thā</i>	<i>aur</i>	<i>jab</i>	<i>kuch-kuch</i>	<i>mahasūs</i>	<i>huā</i>	<i>to</i>
		feeling	NEG	be-IMPF.F	and	when	somewhat	feeling	be.PFV	then
		you were not aware of (Exp) and when you began more or less feeling (Exp) it								

		<i>pahlā</i>	<i>mauqā</i>	<i>milte</i>	<i>hī</i>	<i>ghar</i>	<i>se</i>	<i>calī-gāĩ</i>
--	--	--------------	--------------	--------------	-----------	-------------	-----------	-----------------

	first	occasion	finding	just	home	from	leave.PFV.F.SG
	you left home at the first opportunity' (Mohan Rakesh, Theater, 1959)						

The interesting sequence in the man last answer (“you did not feel the fact that you felt so”) clearly shows the difference between a state of unaware feeling (“you did not feel” with experiential construction) and awareness of it (“the fact that you felt so”, that is to say, “the fact that you realized your solitude” in the agentive construction).

The feature conscious awareness is then a sufficient condition to case mark an experiencer as an agent in (45) and (46). Symmetrically, the lack of this feature is a sufficient condition to case mark an agent as an instrument with action predicates and light verb “be”, as seen in example (46).

Conscious awareness seems to have become a distinctive feature only recently and to be related to the predicate renewal mentioned in section 3.1, since older texts from 14-16<sup>th</sup> century ignore it: for instance “be envious” or “be luxurious” occur with both light verbs in exactly the same sentences and contexts in Kabir (14<sup>th</sup> c.) and even two centuries later in Tulsidas or Mira Bai’s poems no systematic distinction in meaning is observable (Montaut 2013). The emergence of the feature conscious awareness as a distinctive feature has probably been motivated by the need to avoid synonymy (Barddal 2009) in assigning different meanings to the two correlated verbal bases and constructions.

## Conclusion

A device mainly designed for borrowing, initially with the massive integration of Persian predicates in the 17-18<sup>th</sup> centuries, and still systematic with English or Sanskrit borrowed verbs once recategorized as nouns, the Hindi CPs can be deemed responsible, or at least instrumental, in some of the major typological shifts of the language. The greatest impact comes from the rise and development of the so-called dative subjects, which can be considered as a typical semantic alignment, along with other sentence patterns with non-canonical subjects whose argument structure is constrained by semantic roles rather than syntactic function (Donohue & Wichmann 2008). It later resulted in the emergence of a new distinctive feature in the differentiation of semantic roles and their case marking, namely the feature of conscious awareness. The development of the light verb construction as an operator for middle voice, filling the gaps in the flexional alternate pairs of simplex predicates, also contributed to extend the scope of operation of this feature, since an inadvertent agent is case marked as an instrument of the intransitive middle construction. The semantic of the agent role in this way obeys a finer grain grid than the usual definition by volition and control. An agent (case marked as such, NOM/ERG) in Hindi can even lack volition and control if it has conscious awareness, whereas an experiencer case marked in the dative has not. Finally, the genitive coding of the closest argument, either the object or the subject according to the light verb selected, is, for the more productive class of CPs, one more argument for considering Hindi as extensively displaying semantic alignments.

## References

- Abbi, Anvita & Gopalkrishnan, Devi. 1992. Semantic Typology of Explicator Compound Verbs in South Asia. *Language Sciences* 13-2: 161-180.
- Ahmed, Tafseer and Butt, Miriam. 2011. Discovering Semantic Classes for Urdu N-V predicates. *Proceedings of the International Conference on Computational Semantics (WCS 2011)*, 304-9. Oxford: Oxford University Press.
- Ahmed, Tafseer, Butt, Miriam, Hautli, Annette, Sulger, Sebastian Sulger. 2012. A Reference Dependency Bank for Analyzing Complex Predicates. LREC012: 3145-53 <http://www.lrec-conf.org/proceedings/lrec2012/summaries/474.html> (International Language Resource and Evaluation).

- Aikhenvald, Alexandra, Dixon, R.M.W., Onishi, Mayasuki (eds.). 2001. *Non-Canonical marking of Subjects and Objects*. [TSL 46]. Amsterdam: Benjamins.
- Barddal, Johanna. 2009. The Development of Case in Germanic. In *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case* ed. by Johanna Barddal and Shobhana Chelliah, 123-59. Amsterdam: Benjamins.
- Begum Rafia, Jindal Karan, Jain Ashish, Husain Samar and Sarma, Dipti Mishra. 2011. Identification of Conjunct Verbs in Hindi and its effects on parsing accuracy. *CICLing'11 Proceedings of the 12<sup>th</sup> International Conference on Computational Linguistics and Intelligent Text Processing*: 29-40. Heidelberg: Springer Verlag.
- Bhaskararao, P. & Subbarao, K.V. 2004. *Non-nominative Subjects*. Amsterdam: Benjamins.
- Bhattacharya, Pushpak, Chakrabarti, Debasri, and Sarma Vaijayanthi. 2007. Complex predicates in Indian languages and wordnets. *Language Resources and Evaluation* 40(3-4): 331-355
- Burton-Page, John. 1957. Compound and Conjunct Verb in Hindi. *BSOAS* 19: 469-78.
- Butt, Miriam, & Geuder, Wilhelm. 2003. Light Verbs in Urdu and Grammaticalization. *Trends in Linguistics Studies and Monographs, Vol 143*: 295-350.
- Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu*, Chicago: CSLI
- Chakrabarti, Debasri, Mandalia, Hemang, Priya, Ritwik and Bhattacharya, Pushpak. 2008. Hindi Compound Verbs and their automatic extraction. *Computational Linguistics* 8 (COLING08): 27-30.
- Chakrabarti, Debasri, Sarma, Vaijayanti M., Bhattacharya, Pushpak. 2007. Complex Predicates in Indian languages Wordnet. *Lexical Resources and Evaluation Journal* 40: 3-4.
- Davison, Alice. 2005. Phrasal predicates: How N combines with V in Hindi/Urdu. In *Yearbook of South Asian Languages and Linguistics*, ed. by Tanmoy Bhattacharya (ed.), 83-116. Amsterdam/NewYork: Mouton DeGruyter.
- Dayal, Veneeta. 2011. Hindi Pseudo-Incorporation. *Natural Language and Linguistic Theory*, 29.1, 123-167.
- Donohue, Mark. & Wichmann, Soeren. 2008. *The Typology of Semantic Alignment*. Oxford: OUP.
- Durie, Mark. 1988. Preferred Argument Structure in an Active Language. *Lingua* 74: 1-25.
- Emeneau Murray B. 1980. *Essays on linguistic Area*.
- Gambhir, Vijay. 1993. Complex Verb Phrase: A diachronic and synchronic view. In *Experiencer Subjects in South Asian Languages*, edited by Verma, Mahindra K. and K.P. Mohanan. 77-96. Delhi: Manohar Verma.
- Haspelmath, Martin. 1993. More on the Typology of Inchoative/causative Verb Alternations. In *Causatives and transitivity*, ed. by Bernard Comrie & Polinsky, 84-120. Amsterdam: Benjamins.
- Pardeshi, Prashant. 2014. Transitivity Pairs in Hindi. *The World Atlas of Transitivity Pairs* (2014). Tokyo: National Institute for Japanese Language and Linguistics (Available online at <http://www.ninjal.ac.jp/>)
- Hock, Hans.H. 1991. Oblique Subjects in Sanskrit. In Verma, Mahindra K. and Mohanan K.P. (eds.). *Experiencer Subjects in South Asian Languages*, edited by Verma, Mahindra K. and K.P. Mohanan. 119-140. Delhi: Manohar.
- Hook, Peter E. 1974. *The Compound Verb in Hindi*. Ann Arbor: The Michigan University Center for South and South East Asian Studies.
- Karimi, Simin. 1997. Persian Complex Verbs: Idiomatic or compositional. *Lexicology* 3: 273-318.
- Karimi-Doostan, Gholamhossein. 2011. Separability of Light Verb in Persian. *Sudia Linguistica* 65-1: 79-95.
- Kellogg, Rev. H.S., 1875. *A Grammar of the Hindi Language*. London: Kegan.
- Kemmer, Susanne. 1993. *The Middle voice*. Amsterdam: Benjamins.
- Mohanan, Tara. 1994. *Argument Structure in Hindi*. Chicago: CSLI.
- Mohanan, Tara. 1995. Incorporation. Wordhood and lexicality. *NLLT* 13: 75:134.
- Montaut, Annie. 2011. Basic Intransitivity in Hindi/Urdu. In *Indo-Aryan Linguistics*, ed. by Omkar N. Koul, 29-46. Mysore: Central Institute of Indian Languages.
- Montaut, Annie. 2013. The rise of non-canonical subjects and the development of semantic alignments in Hindi. In *Diachronic Typology of non-canonical Subjects* ed. by Ilja Serzants & Leonid Kulikov, 33-56. Amsterdam: Benjamin.

- Moravcsik, Edith. 1975. Verb Borrowing. *Wiener Linguistics Gazette* 8: 3-10.
- Mukerjee, Amitabha, Soni, Ankit and Achla M. Raina. 2006. Detecting Complex Predicates in Hindi using PoS Projection across parallel corpora. *Proceedings of the Workshop on Multiword Expressions: Identifying and exploring underlying properties*, 28-35. Sidney: Association for Computational Linguistics.
- Nichols, Johanna, 2008. Why are Active Languages rare in Eurasia? In *The Typology of active alignments*, ed. by Donohue, Mark & Soren Wichmann, 121-139. Oxford: OUP
- Nunberg, Geoffrey., Sag, Ivan A. and Wastow, Thomas. 1994. Idioms. *Language* 70: 756-96.
- Onishi, Masayuki. 2001. Non-canonically Marked S/A in Bengali. In *Non-Canonical Marking of Subjects and Objects*, ed. by Aikhenvald & al., 113-147.
- Raina, Achla M. and Mukerjee, Amitabha. 2006. Complex predicates in the generative lexicon. In Bouillon, Pierrette and Kansaki Kyoko (eds.) *Third International Workshop on Generative Approaches to the Lexicon (Proceedings of GL'2005)*, 210-21. Geneva: School of Interpretation and Translation of Geneva.
- Samvelian, Pollet. 2012. *Grammaire des prédicats complexes. Les constructions nom-verbe*. Paris : Hermès-Lavoisier.
- Samvelian, Pollet & Faghiri, Pegah. Under press. Rethinking compositionality in Persian Complex Predicates. *BLS* 39.
- Sasse, Hans-Jürgen. 1991. Aspect and Aktionsart: A Reconciliation. In *Perspectives in Aspect and Aktionsart. (Belgian Journal of Linguistics)*, edited by Veters, Carl, Vandeweghe, Willy and Dominicy, Marc, 31-45. Bruxelles: Editions de l'Université de Bruxelles.
- Sasse, Hans-Jürgen. 2006. Aspect and Aktionsart. In Brown, Keith (ed.), *Encyclopedia of Language and Linguistics*, edited by Keith Brown, 535-38. Leiden: Elsevier.
- Sinha, K. and R. Mahesh. 2009. Mining Complex Predicates in Hindi in using a parallel Hindi/English Corpus. *Proceedings of the Workshop on Multiword Expressions: Identification, Interpretation, Disambiguation and Application*, 40-46. Kanpur: Indian Institute of Technology.
- Tsunoda, T. 1985. Remarks on Transitivity. *Journal of Linguistics* 21: 385-96.
- Verma, Mahindra K. and K.P. Mohanan (eds.) 1993. *Complex Predicates in South Asian Languages*. Delhi: Manohar.
- Wohlgemuth, Jan. 2009. *A Typology of verbal borrowings*. Berlin: Mouton de Gruyter.

#### Abbreviations not in the list

- PRSP presumptive  
 PFV perfective  
 TERM terminative