

**PHRASAL PREDICATES**  
**HOW N COMBINES WITH V IN HINDI/URDU**

Alice Davison, U. of Iowa  
alice-davison@uiowa.edu

April 27, 2005

1. Introduction

Many languages form a large part of their verbal lexicon by combining a verb such as ‘do/make’ with an adjective or complex event noun.<sup>1</sup> The combination of N and V constitutes a syntactic phrase with a single argument structure. Complex predicates of this sort are found in Japanese, Turkish, and Persian, among many others. Since complex predicates are so generally available, they must be the result of very general linguistic processes, which I will propose are part of syntax and the syntax/interpretative interface, or logical form. In this paper I consider two kinds of syntactic derivation, incorporation and argument merger, or formation of a composite argument structure. The goal of this analysis is to show how N and V contribute to the properties of the complex predicate, and to derive the right phrase structure, case and agreement properties.

Complex predicates form a large and productive part of the verbal lexicon of Hindi/Urdu (1)-(2):

- (1) maiN-nee/\*mujhee -[[un-kii     **yaad**]           **kii**]  
I-erg     /\*I-dat     3pl-gen-fs memory-fs do-pf-fs  
‘I remembered, recalled them’
- (2) mujhee/\*maiN-nee [[un-kii     **yaad**]           **aaii**]  
I-dat     /\*I-erg           3pl-gen memory-fs come-pf-fs  
‘I remembered/missed them.’
- (3) [un-kii] **yaad**  
3pl memory-fs  
‘a memory, recollection about them.’<sup>2</sup>
- (4) maiN-nee kyaa     **ki-yaa?**  
I -erg what? do-pf-ms  
‘What did I do?’

The N *yaad* ‘memory’ takes a thematic object, the thing remembered (3), and implies an experiencer, the one who remembers. In combination with a V (4), this external argument is syntactically projected with nominative, ergative (1) or dative (2) case.

While sentences like (1)-(2) are common and unremarkable in Hindi/Urdu, their structure poses both general and language-specific problems for deriving the right syntax, morphology and interpretation. The N component of N-V is the syntactic object of V, with the null nominative object case and verbal agreement. (fs in this case) required when the subject is marked by a postposition, blocking agreement with the subject. The N *yaad* ‘memory’ has a syntactic object of its own which require case licensing (3). The N selects the case for it, in this instance genitive (1)-(3). The V determines the case of the external argument of the sentence,

ergative when the V is verb like *kar-naa* ‘do’ (1), dative with V *aa-naa* ‘come’. The combination forms a syntactic phrase, which can be interrupted by negation and particles, modified by an adjective on N or reordered.

There are also interpretive/semantic problems. N is a complex event nominal (Grimshaw 1990), which like V has an argument structure, as independently occurring words (3)-(4). But the semantic roles in these argument structures may not be the same. The external argument of *yaad* ‘memory’ has primarily an experiencer role (1)-(3), while the external argument of *kar-naa* ‘do/make’ is typically agentive (4) and the subject of *aa-naa* ‘come’ refers to an entity with or without internal motivation which moves to a destination. The argument structures of both N and V have to be reconciled as one argument structure, preserving the right semantic relations for the arguments. The subjects in both (1)-(2) have experiencer interpretation, so that these sentences may be synonymous in specific contexts. They overlap in meaning, allowing also for different interpretations determined by V because V also is responsible for the event structure and verbal aspect of the combination. (For arguments that ergative case is a structural case not linked to semantic agency, see Davison (2004, to appear).)

Many but not all of these problems have been articulated from different descriptive or theoretical viewpoints, from earlier writers such as Bailey (1963), Kachru (1982) Bahl (1974), Hook (1979), to more recent work by Montaut (1991), Verma (1993, Mohanan (1994) Butt (1995), Butt and Ramchand (2003). The goals of this paper are (a) to form a comprehensive picture of the facts about N-V complex predicates in Hindi/Urdu, (b) to propose an account which preserves the productive syntactic nature of complex predicates, showing their relation to related simplex predicates in the lexicon of the language, (c) to account for case and syntactic structure, and (d) to explain how the properties of N and V are combined. I will relate N-V predicates to other classes in the lexicon of Hindi/Urdu, drawing some conclusions about the differences in their derivation and syntactic projection from N-V predicates. I compare two possible theoretical approaches, one based on incorporation, as in Hale and Keyser’s conflation process in lexical syntax (1993), the other based on Higginbotham’s predicate composition (1985, 1999).

## 2. The phrasal syntax of N-V combinations

The N of the N-V combination may be independently modified by an adjective or adverb (5)-(6). Negation or emphatic particles may intervene between N and V (6)-(7):

- (5) us-nee moohan-kii **bahut yaad**            **kii**  
 3s-erg Mohan-gen much memory-fs do-pf-fs  
 ‘He/she remembered Mohan very much.’ Bahl 1974:xxix)
- (6) coor-nee [apnee dimaag-kaa *sahii isteemal*] *nahiiN ki-yaa* ]  
 thief-erg self’s mind-m-gen right use-m not do-pf-ms  
 ‘The thief did not make right use of his mind.’ (Hook 1979:159)
- (7) mujhee **pataa**            *hii/bhii nahiiN thaa*  
 I-dat knowledge-m only/also not was-ms

‘I did not even know.’ Montaut (to appear: 296)

The adj/adv *bahut* ‘much’ in (5) intervenes between the genitive object phrase and N. Negation in (6) and (7) intervenes between N and V, while N has an adjective modifier *sahii* ‘correct, right’ in (6).

The N itself behaves like a syntactic direct object. It has gender, person and number agreement features reflected on the verbal complex under just the same circumstances that thematic object DPs require agreement. Object agreement is required when the subject has non-nominative case, and the direct object also is not marked with a postpositional case. Compare the object agreement in (1) above with the subject agreement in (8). The simplex verbs in (9a,b), show the same contrast of subject and object agreement. Verbal agreement is determined by the nominative subjects in (8) and (9a), but when the subject is marked by postpositional case, the verb agrees with a nominative object. The verb agrees with *yaad* ‘memory’ in (1)-(2), and the nominative syntactic direct object in (9b).

- (8) pulis (waalaa) coor -kii      **talaash kar-eegaa**  
police (man) thief -gen-fs search-fs do-fut-3ms  
‘The policeman will search for the thief.’
- (9) a. baccee      billii      **deekh-eeNgee**  
child-mpl cat-fs see-fut-3pl-m  
‘The children will see a cat.’  
b. baccuN-nee billii **deekh-ii**  
child-mpl-erg cat-fs see-pf-fs  
‘The children saw a cat.’

In passive sentences, the constituent singled out as the direct object is the N component of N-V in (10), while it is the thematic object of the simplex verb in (11).

- (10) pulis-kee dwaaraa coor -kii      **talaash kii**      ga-ii  
police-gen means thief-ms-gen-fs search-fs do-pf-fs go-pf-fs  
‘The thief was not searched for by the police.’
- (11) baccuN-see      billii      **deekh-ii**      nahiiN      gaii  
child-mpl-from cat-fs see-pf-fs not go-pf-fs  
‘The children couldn’t bear to see the cat; lit. the cat was not seen by the children.’<sup>3</sup>

The N component of N-V complex predicates has these two properties in common with referential direct object DPs: it triggers agreement when the subject is not nominative, and it is the object promoted to subject in passive sentences.<sup>4</sup> The N component also behaves like an independent N linked with a DP complement by the genitive postposition, which is suffixed to DP, and agrees with N.

- (12) a. raam      -kii      **gaaRii**  
Ram-ms gen-f car-fs  
b. raam-kii      **yaad**  
Ram-gen-f memory-fs

‘Ram’s car’

‘the memory of/about Ram’

I will argue below that complex predicates are syntactically derived, not lexical idioms.

### 3. Other lexical classes of predicates in Hindi/Urdu

In this section I compare N-V predicates like (5)-(7) with other kinds of predicates related to nouns, especially complex event nominals. The Hindi lexicon includes simplex verbs related to the same kind of N found in N-V combinations, and in many cases coexisting with N-V. built on the same N stem, and coexisting with the N-V combination. There are also N-V combinations in which the N and V form a closer unit than in the preceding examples, in some sense intermediate between a phrasal verb and a verb stem. In section 4, I use the discussion of these predicate classes to ask whether there is a common derivation for all types of predicate, by incorporation.

#### 3.1 Denominal verbs

There are many denominal verbs in Hindi/Urdu, such as those in (13)-(15). The formation of denominal verbs is no longer freely productive, but it remained productive after the massive of borrowing of Perso-Arabic vocabulary, such as Persian *daaG*, *talaash*, etc.

(13) N. *daaG* ‘mark of burning’ V. *daaG-naa* ‘burn, brand’ V *daG-naa* ‘be branded’<sup>5</sup>

(14) a. *khooj* ‘search’ N *khooj-naa* ‘to search for’ V  
b. *bahas* ‘argument, dispute’ N. *bahas-naa* ‘to argue’ V  
c. *talaash* ‘search, investigation’ N *talaash-naa* ‘search according to abstract criteria’ V  
d. *bhuul* ‘oversight, omission’ N. *bhuul-naa* ‘forget, overlook, omit’  
e. *pahcaan* ‘recognition, judgement’ N *pahcaan-naa* ‘recognize, perceive’  
f. *baat* ‘what is said, conversation’ N *bataa-naa* ‘tell, explain’

(15) a. *khooj kar-naa* ‘search, discover’  
b. *bahas kar-naa* ‘argue, debate (with)’  
c. *talaash kar-naa* ‘search for a concrete object’  
d. *bhuul kar-naa* ‘to make a mistake’  
e. *pahcaan kar-naa* ‘recognize, identify.’  
f. *baat kar-naa* ‘talk with’

The simplex denominal verbs (14) coexist with the corresponding N-V complex predicates (14), often with a different range of meaning (Gambhir 1993). In particular, the verb *khooj-naa* ‘search’ contrasts with *khooj kar-naa* ‘search, find’. The N-V predicate as an accomplishment meaning not found with the simplex verb (16).<sup>6</sup>

(16) *ganapat singh-nee* [*eek naai bimaarii-kii khooj*] **kii** *hai*  
Ganpat Singh-erg one new illness-gen search(fs)-nom do-pf-fs. is  
‘Ganpat Singh has discovered a new disease.’ (Bahl 1974:222)

The meaning of the complex predicate is not always completely predicted from the meaning of N and the simplex denominal verb. The selection of lexical case properties, however, often relates denominal verbs to complex predicates (17)-(18) via the shared N stems, suggesting that lexical case is associated with N.

- (17) a. *baccee billii-see Dar-tee haiN*  
 child-mpl cat-from fear-impf-mpl are  
 ‘The children are afraid of the cat.’  
 b. *bacoonN-koo billii-see Dar lag-taa hai*  
 child-mpl-dat cat-from fear-ms strike-impf is  
 ‘The children are afraid of the cat.’
- (18) a. *coor apnii karnii-par pachtaa-yaa*  
 thief self’s misdeed-on .repent-pf  
 ‘The thief repented his misdeeds.’  
 b. *coor-koo apnii karnii-par pachtaavaa hai*  
 thief-dat self’s misdeed-on repentance is  
 ‘The thief repented his misdeeds.’

### 3.2 Diatheses, verb alternations

Complex predicates in Hindi/Urdu fill in paradigmatic gaps with respect to denominal verbs. There is a large number of related verbs of the kind in (19). Related Ns include *khood* ‘hollow’, *phair* ‘turn’ and *kaat* ‘cutting’. The verbs in each set share a basic root, but are distinguished by stem alternations and suffixes, as well as by the number of arguments projected syntactically. The transitive stem (19b) takes both subject and object arguments. A related detransitive verb (19a) has a phonologically reduced stem, and projects the thematic object as syntactic subject. The same reduced stem combines with the causative affixes *-aa* and *-vaa*, increasing the valence of the verb by the addition of a causer.

- |      |                 |                  |                         |                            |
|------|-----------------|------------------|-------------------------|----------------------------|
| (19) | a. detransitive | b. transitive    | c. causative 1          | d. causative 2             |
|      | <i>khul-naa</i> | <i>khol-naa</i>  | <i>khul-aa-naa</i>      | <i>khul-vaa-naa</i>        |
|      | ‘to be open’    | ‘to open’        | ‘to cause to be open’   | ‘to get opened by another’ |
|      | <i>phir-naa</i> | <i>phair-naa</i> | <i>phir-aa-naa</i>      | <i>phir-vaa-naa</i>        |
|      | ‘to turn’       | ‘to turn’        | ‘to cause to be turned’ | ‘to cause to be turned’    |
|      | <i>kaT-naa</i>  | <i>kaat-naa</i>  | <i>kaT-aa-naa</i>       | <i>kaT-vaa-naa</i>         |
|      | ‘to be cut’     | ‘to cut’         | ‘to cause to be cut’    | ‘to cause to be cut’       |

As Gambhir (1993) noted, some denominal verbs do not participate in this paradigm (20), and instead complex predicates are formed with the same valence properties as the missing derivatives.

- (20) a. intransitive                      b. transitive                      c. causative

<b>talaash hoo-naa</b>	talaash kar-naa	<b>talaash kar-vaa-naa</b>
'be searched'	'search for'	'have someone search for'
* <i>talash-naa</i>	talaash-naa	* <i>talash-vaa-naab.</i>
'be searched'	'search for'	'have someone search'
(Gambhir 1993:89)		

There is a category of complex predicates in which the N and V, while remaining separate words, form a closer association than the N and V which I have called phrasal (1), (2), (6). The difference from the truly phrasal N-V combinations is that the thematic object is syntactically the direct object. It can trigger verb agreement (20), (21), (23):

- (20) bhaiyaa-nee [**apnii taaqat**]                      **isteemaal kii**  
brother-erg. self's strength-nom-fs      use-ms      do-pf-fs.  
'Brother used his strength.' (Hook (1979:158)
- (21) coor -nee [apnaa dimaag] (\*-kaa)      (\**sahii*) **isteemal** nahiiN **ki-yaa**  
thief-erg self's mind(m)      (\*gen)      right      use(m)      not      do-pf-ms  
'The thief did not make right use of his mind.' (Hook 1979:159)
- (22) us-nee [moohan -koo] bahut **yaad**                      **ki-yaa**  
3s-erg Mohan-**dat**      much memory-f      do-pf-ms  
'He/she remembered Mohan very much.' (Bahl 1974:xxix)
- (23) siitaa-koo [apnee maaN-baap]                      **yaad**                      **aa-ee**  
Sita-dat      self's mother-father-mpl      memory-fs      come-pf-mpl  
'Sita remembered/missed her parents.' Saxena 1985)

The thematic object can have nominative case (20), (21), (23), or dative direct object case (22). Unlike N in examples about such as (2), N is not linked by genitive case or locative case to its thematic object. Rather, the V itself determines the direct object case. This option is possible only for certain Ns, and for adjectives like *mazbuur* 'forced' which assign no case (see section 5).

Both denominal verbs and N-V complex predicates are transparently related to N. But the facts of the language suggest that denominal verbs N-V are not optional variants of one another. A simplex verb and a complex predicate related to the same N coexist in the language, but often with different nuances of meaning or belonging to different aspectual classes, such as *khooj-naa* 'search' and *khooj kar-naa* 'search, discover'(14)-(15). Many denominal verbs participate in transitivity alternations (19), but not all do, and the gaps in the paradigm are filled by N-V predicates (20), suggesting again that V and N-V have semantic commonalities but are not automatically available variants of one another. Finally, N-V combinations are for the most part phrasal, linked to a genitive or locative-marked thematic object, the choice determined by N. But a subset of these N-V complex predicates also form a more closely associated combination distinct from a simplex verb.

#### 4. Incorporation as model of derivation and interpretation

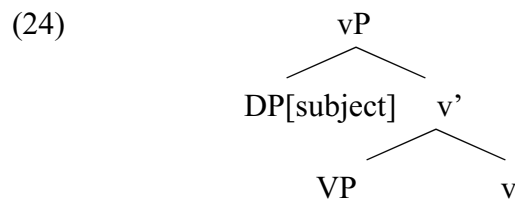
In this section I consider proposals for incorporating nominal information into verbs, whether

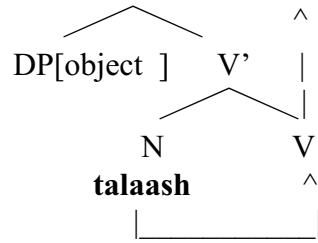
null, overt or present as causative affixes, so as to explain how N can be related in various ways to V. I will first consider the derivation of denominal verbs and the transitivity alternation by means of Hale and Keyser's conflation, or movement of N into a verbal head in lexical syntax (1993). I then explore the extension of conflation to analytic or overt phrasal causatives, asking if this kind of derivation can account for the difference between the clearly phrasal complex predicates (1)(2) and the more closely associated N-V counterparts (20)-(23). I next consider whether the lexical syntax of the conflation account captures the full range of event structures in the verbs discussed here, particularly those with experiencer subjects, and particularly when the experiencer interpretation is dictated by N and not V. Finally I consider semantic incorporation, which preserves phrasal structure but combines information in Logical Form; I compare the interpretation of incorporation of indefinite NP objects with the incorporation of complex event nominals with their own full argument structure. This discussion serves as a background to the alternative I will propose in the section 5, based on Higginbotham's predicate composition by argument identification (1985, 1999).

#### 4.1 Denominal verbs and transitivity alternations

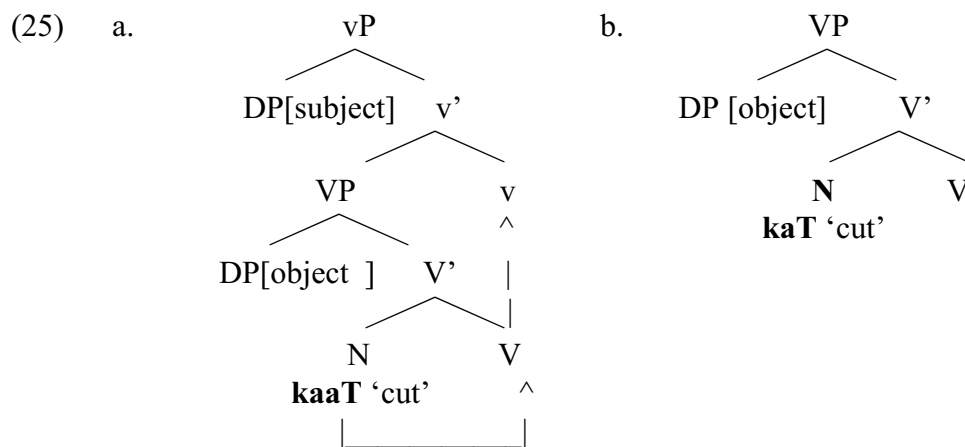
Hale and Keyser (1993) draw comparisons between the overt phrasal syntax which licenses objects and the object relations implicit in derived verbs, especially verbs of change, such as *clear*. This deadjectival verb is derived by their process of conflation by incorporating the Adj *clear*. In l(exical)-syntax, so that the verb *clear* is equivalent to 'V-cause the screen V-become clear', with the Adj *clear* incorporated into the abstract null V of causation and the inchoative V-become. The theta roles are directly read off the l-syntax: the subject of V-cause has the agent interpretation, while 'the screen' gets the role of subject of change or theme. Intransitive activity verbs which are 'unergative', like *dance* are derived in l-syntax by incorporating the N *dance* which is the l-syntactic object of V 'do'.

We must extrapolate a little from Hale and Keyser's account of English to derive the Hindi transitive denominal verbs such as those in (14) (17a) and (18a). An activity verb like *talaash-naa* 'search for' would have the l-syntactic structure (24):





The N *talaash* ‘search’ undergoes head movement to the null V and v in I-syntax. The N takes on the category V, and the syntactic DPs which get interpreted as semantic arguments of the verb. The same structure is the basis for verbs of change and causation, such as *kaaT-naa* ‘to cut’, and *kaT-naa* ‘to be cut’, related to N *kaat* ‘cutting’:



The syntactic object (25a) is the subject of V the verb of change (25b), while the syntactic subject is the specifier of v or the overt causative affix (25a). The various verb stems associate with V and v. The association is not free however, as the full form of the stem *kaaT-* is associated only with the transitive structure. The reduced form *kaT* associates in the intransitive and causative versions. Assuming there is some way to guarantee that the right form of the stem is associated with the right projection in I-syntax, the conflation approach gives a reasonably good results.<sup>7</sup>

#### 4.2. Can incorporation be extended to ‘analytic’ complex predicates?

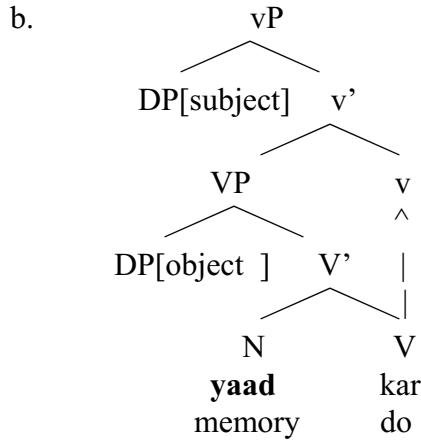
Here I consider whether complex predicates can be given a similar account, substituting lexical verbs for the null V and v. The presence of lexical verbs blocks head movement in I-syntax. A proposal of this sort is made for Persian complex predicates by Folli et al (to appear).

Generalizing this proposal to Hindi/Urdu confronts problems in object case marking, and subject interpretation. There are differences of case marking between Persian and Hindi/Urdu. Complex predicates in Persian mark the thematic object for case in the same way as simplex verbs, with null nominative or the differential object marker *-ra*. Complex predicates in Hindi/Urdu have three different options for marking the thematic object of complex predicates. First, the thematic object may be linked by genitive case to N (agreeing with N in number-gender

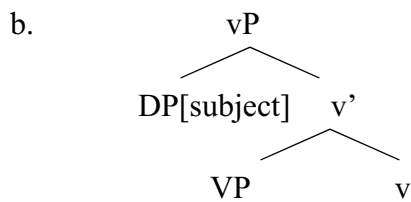
features) (1)(2). Second, some Ns select locative lexical case in combination with V (17b) (18b). These options are the most common ones. Finally, the N-V forms a close combination giving the thematic object the same case as a simplex verb, either null nominative case or the differential object marker *-koo* (20-23). This third option is the least common in Hindi, although it seems to be the most the preferred marking in Persian.

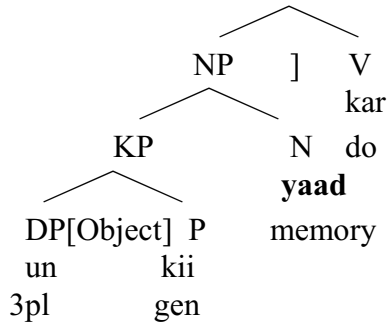
This third possibility might be captured using a structure similar to (25a) for a sentence like (22), repeated here as (26a):

- (26) a. us-nee moohan -koo bahut **yaad** **ki-yaa**  
 3s-erg Mohan-dat much memory-f do-pf-ms  
 ‘He/she remembered Mohan very much.’ (Bahl 1974:xxix)



- (27) a. maiN-nee [[un-kii **yaad** **kii**]  
 I-erg 3pl-gen-fs memory-fs do-pf-fs  
 ‘I remembered, recalled them’





The more usual expression of complex predicates involves a NP object containing the predicate N and its thematic object, which gets a case determined by N, genitive or locative (27a). The syntactic structure required for such sentences would seem to be (27b), which represents the thematic object as a constituent of NP, rather than a specifier of VP. It is hard to see how (27b) could be derived from (26b) without stipulation of object lowering or V' raising, and subsequent restructuring of the thematic object DP as a constituent of NP.

A further problem is the semantic role assigned to the subject, which does not follow from these structures. Neither (26b) nor (27b) explains how the N determines the semantic role of experiencer for the syntactic subject, as the subject is not a syntactic argument of N. Both structures predict that only V assigns a theta role to the subject (as is the case for Persian, according to Folli et al. (to appear)). In section 5, I will propose an alternative account which directly generates (27b) and the less usual (26a), and allows the N to determine the semantic roles of the subject.

#### 4.3 The semantic role of the subject

While many N-V complex predicates like *talaash kar-naa* 'search' have agentive subjects, not all do. Complex predicates such as *yaad kar-naa* 'remember, recall, miss', *afsoos kar-naa* 'regret' and *mahsuus kar-naa* 'experience, feel' have non-agentive experiencer interpretations for the subject:

- (28) ab diwaalii paas aa rahii hai, is-liyee baccee maaN-koo **bahut yaad kar-tee haiN**  
 now Diwali near come prog is so children mother-dat much memory do-impf are  
 'Now Diwali is approaching, and for this reason the children miss their mother very much.'  
 (Bahl 1974: 73)

- (29) yee gumraah aurateen peeshtar nahiiN [ too sharaab-kaa nashaa utar -nee- kee  
 these wayward women earlier not then liquor-gen intoxication come down-inf-gen  
 baad][ apnii haalat] -par **afsoos kar-tii haiN**.  
 after self's condition-on regret do-impf are  
 'These wayward women, not at first, but after coming down from the intoxication of

- liquor, feel regret about their condition.' (Bahl 1974:98)
- (30) kanyaa-nee **mahsuus ki-yaa** [ki woo jiivan-meeN pahlii baar kisii puruS-kee  
 Kanyaa-erg feeling do-pf that 3s life-in first time some man-gen  
 aakarSaNpaash-meeN puurii taur-par bandh ga-ii hai.  
 attraction-snare-in full state-in tied go-pf  
 'Kanya felt [that for the first time in her life that she was completely snared in the trap of  
 attraction to some man].' (Bahl 1974: 253).

Since the event N is the syntactic object of V, and is not raised to vP, it is not associated syntactically with the subject of vP in the syntactic structures (26a)-(27a). This representation predicts that the subject of N-V should have the subject semantic role assigned by V, which should be agent if the V is a transitive activity or accomplishment verb like *kar-naa* 'do, make'. But in (28)-(30) the syntactic subjects of these sentences do not refer to causative agents or events which bring about a change of state.

One of the advantages of the Hale-Keyser representation of the lexical syntax of verbs is that the syntactic projections are at the same time a representation of the event structure of the verb, at least for verbs of causation and change. There is an instigating entity (or event) which causes a change, which implicates a state (*clear*), or a relation (*shelve/on the shelf*). Activities like *laugh* implicate an entity which is produced by the event. Depending on the event complexity of the verb, the event structure can be represented schematically as (31):

- (31) a.  $n > (e \rightarrow r \text{ (relation)/ } s \text{ (state)})$  transitive verbs of causation and change  
 b.  $(e \rightarrow s)$  inchoative verbs, change of state, unaccusatives  
 c.  $e \rightarrow n$  activity verbs, unergative  
 (Hale and Keyser 1993: 69ff)

The subevents correspond to the verbal phrases within the I-syntactic structure underlying a verb. The semantic roles of the verb's arguments follow from the VP/event with which they are associated, making it unnecessary to stipulate a traditional ordered array of theta roles.

Verbs of perception, mental or psychological state would seem not to conform to this otherwise perspicuous description.<sup>8</sup> The direct objects of such verbs are not 'subjects of change', undergoing a transition to a resulting state. Rather they could be interpreted as the locus of a feeling or perception. The instigating event or cause of the feeling or perception can be present in the context without being projected as a sentence constituent. In (28), for example, the approach of the Diwali festival arouses in the children the memory of their mother. The stimulus can be expressed as an optional adverbial clause rather than a syntactic subject or locative source, with both simplex and complex predicates (32)-(33):

- (32) [apnee saamnee doo banduuk-dhaarii DaakuuooN-koo *deekh-kar*] woo **Dar-aa thaa**  
 self's front two gun-carrying robber-dat see-prt 3s fear-pf was  
 'Seeing two gun-carrying robbers in front of him, he became afraid.' Nespital 1997:624
- (33) [kaaminii-kee pahlee din-kii baat *sun-nee-kee baad*] maiN apnee-meeN ajiib  
 Kamini-gen first day gen matter hear-inf-gen after I self-in strange  
 hiintaa -kaa **anubhav kar-nee lag-aa thaa**

deficiency-gen experience do-inf begin-pf was  
 ‘After hearing what Kamini said the first day, I began to feel a strange inferiority.’  
 Bahl 1974: 248

The same surface syntax can encode different event structures, either the causative/agentive schema for (20)-(21). The same event structure for psychological predicates can be expressed syntactically in two ways, one with a nominative or ergative subject (28)-(30), or with a dative experiencer subject (34). In both structures the syntactic subject refers to the experiencer of a mental or physical state, which may come about as a process.

(34) *siitaa-koo apnee maaN-baap yaad aa-ee*  
 Sita-dat self’s mother-father-mpl memory-fs come-pf-mpl  
 ‘Sita remembered/missed her parents.’ (Saxena 1985)

The causative structure (26b) which has an overt V *kar-naa* ‘do’ accounts for only a subset of complex predicates such as (20), but it fails to derive the right case, constituent structure and subject semantic role in many other instances. One of the problems with (26b)-(27b) is that N and not the V provides the experiencer interpretation of the syntactic subject. When verbs like *kar-naa* ‘do, make’ and *aa-naa* ‘come’ are used in isolation, it is not the case that their subjects are interpreted as experiencers, but this interpretation is associated with N-V combinations. But the V does add to the aspectual interpretation of N-V. The sentence (35) contrasts the two expressions for ‘remember’.

(35) *aur us-see mujhee bahut-sii baateeN yaad aa gayii thiiN jinheeN maiN*  
 and 3s-from I-dat many-rather matter-pl memory come go-pf were which-dat I  
***yaad nahiiN kar-tii***  
 memory not do-impf  
 ‘From that I remember many things which I do not want to recall/think about.’  
 Bahl 1974: 73

The dative subject version *yaad aa-naa* ‘memory come’ can only refer to involuntary memory, but some agency as well as perception may be conveyed by the nominative subject version *yaad kar-naa* ‘do memory’.

#### 4.4 The semantics of incorporation and the interpretation of complex predicates

Accounts of complex predicates have proposed many ways for the argument structure of N to be combined with V. For example, Grimshaw and Mester’s process of argument transfer (1988) assumes that the ‘light’ verb is denuded of theta roles, and the arguments of N are transferred into V. Their version of argument transfer requires a number of stipulations about the lexical representation of N and V. But denominal verbs of the kind discussed above are formed by syntactic incorporation of N into a null verb head, and so necessarily the argument structure of the resulting verb is determined by N. The result is a syntactic unit which assigns case to its object.<sup>9</sup>

Phrasal combinations have been analyzed as the result of semantic incorporation, which leaves the phrasal surface syntax intact, but combines N and V at LF. These combinations are typically made up of an indefinite object and an activity verbs (Mohan 1995, Dayal 2003, Chung and Ladusaw 2003) The N is interpreted not as a referential expression but as an intransitive predicate which restricts the meaning of the verb in more or less the way an adverb does. Dayal (2003) argues persuasively for an analysis of this kind for Hindi/Urdu, refining the proposal made by Mohan (1995). This indefinite reading is associated with (36) with a nominative singular object, which can have a plural meaning absent when the case of the object is dative:

- (36) . anu baccaa /baccee-koo **samhaal rahii hai**  
 Anu child-nom/ child-dat look-after prog is  
 ‘Anu is looking after children (one or more)/ the child.’ (Dayal 2003)

The subject’s acts of ‘looking after’ are restricted to children as the direct object.<sup>10</sup>

Suppose that N-V predicates were accounted for in the same way. The event noun *yaad* ‘memory’ with its transitive argument structure ‘x remembering y’ restricts the object of *kar-naa* ‘do’. The result can be paraphrased more or less as ‘z’s acts of doing are restricted to x remembering y’. If this meaning is possible at all, it seems to be a heavily literal meaning with separate subjects of ‘do’ and ‘memory’, quite far from the actual meaning of the complex predicate, which has a single argument structure.<sup>11</sup>

So I will conclude that Hindi/Urdu has both syntactic and semantic incorporation, but for very specific classes of predicate. Denominal verbs and the intransitive-transitive-causative diathesis are derived by syntactic incorporation in l-syntax, while semantic incorporation involves indefinite objects without complex argument structure. Even if complex predicates were to be derived by some sort of semantic incorporation, it is not enough in itself to say that N is linked with V at LF; what is needed is an explanation of how the properties of N including its argument structure are linked to V, yielding a single argument-licensing unit.

## 5. Derivation of complex predicates

In this section I will propose an alternative to incorporation for deriving N-V combinations with the right surface syntax, case marking and semantic roles for the subject and thematic object. The focus here is on syntactic derivation and constraints on derivation. The derivation I propose for N-V predicate formation adopts the bottom to top successive phrase formation of the Minimalist Program (Chomsky 1995 and subsequent work). Syntactic combinations are formed by MERGE, linking heads and phrases and projecting the head syntactically. The arguments of the head correspond to successively merged phrases licensed by a process of Theta Discharge<sup>12</sup>. An alternative process to Discharge is Theta Merger, combining the argument structures of two syntactic elements in combination.

### 5.1 Case, agreement and semantic roles in Hindi/Urdu complex predicates

The goal is to explain and differentiate the three different syntactic outcomes for N-V, summed

up in (37)-(39). The case of the thematic object classifies the three types:

- (37) Type I- genitive thematic object, N-V agreement  
ganapat singh-nee [eek naii bimaarii -*kii khooj*] **kii** hai  
Ganpat Singh-erg one new illness- -gen search(fs)-nom do-pf-fs. is  
'Ganpat Singh has discovered a new disease.' (Bahl 1974:222)
- (38) Type II - locative thematic object, N-V agreement  
yee gumraah aurateen . . . [ apnii haalat -par] **afsoos** **kar-tii haiN**  
these wayward women self's condition-on regret do-impf are  
'These wayward women . . . feel regret about their condition.' (Bahl 1974:98)
- (39) Type III - dative or nominative thematic object, N does not agree with V  
a. siitaa-koo apnee maaN-baap **yaad** **aa-ee**  
Sita-dat self's mother-father-mpl memory-fs come-pf-mpl  
'Sita remembered/missed her parents.' (Saxena 1985)  
b. us-nee moohan -koo bahut **yaad** **ki-yaa**  
3s-erg Mohan-dat much memory-f do-pf-ms  
'He/she remembered Mohan very much.' (Bahl 1974:xxix)

The N is morphologically and syntactically the object of V in Types I and II, but not in Type III. N assigns object case in Types I and II, but not Type III. The N assigns semantic roles to both the subject and thematic object, while V is responsible for the aspectual class of the predicate. V may contribute semantic roles in all cases. The V projection assigns the case of the thematic object.

An account of complex predicates in Hindi/Urdu has to allow for many different kinds of argument structures, including those with experiencer predicates (37)- (39), as well as those with agentive subjects (20)-(21). It must derive correctly the respective contributions of N and V; N provides the semantic roles of the arguments, including the syntactic subject, and in most instances, the case-marking of the thematic object, while V accounts for the case of the subject, the event structure and aspectual class of the verb, and in some instances, the case of the thematic object.

## 5.2. Theta-related processes

To simplify the exposition and to focus on the crucial features of the syntactic derivation, I will assume a lexical representation of the argument structure of verbs and event nouns as in (40). It identifies an ordered list of arguments, and contains a Davidsonian event argument (Higginbotham 1985, Speas 1990):

- (40) <1, 2, 3, e>

The arguments in this lexical structure have semantic roles which are linked to specific phrases by the process of *Theta discharge* process. I assume this association is a consequence of MERGE and PROJECT (Chomsky 1995). Discharged theta roles are distinguished by 2\* indicating that the position of arguments so marked is *saturated* and does not license further

combination. By *Theta projectio*, the argument structure is projected upward to the newly formed phrase.

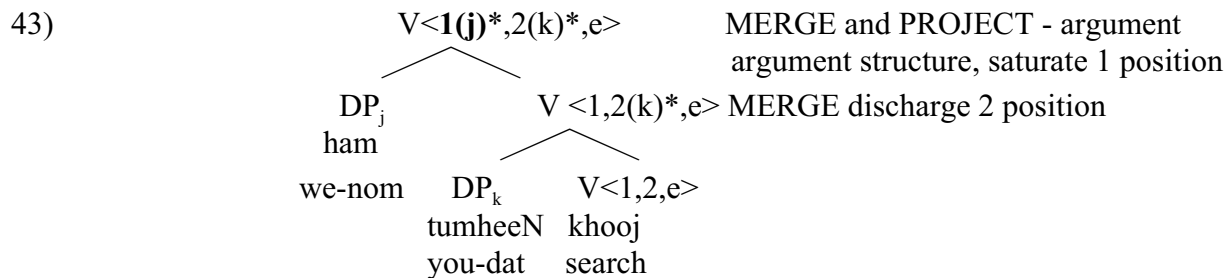
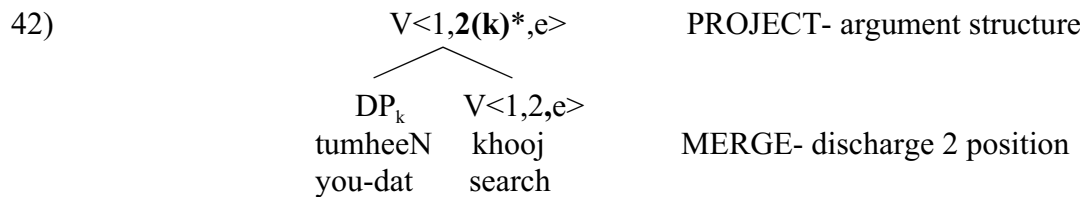
*Theta/argument identification* identifies one argument position in an argument structure with an argument position in another argument structure; *theta or argument merger* is the result of linking or identifying corresponding argument positions in two separate argument structures to form a new composite argument structure.

### 5.3. An example of a well-formed derivation

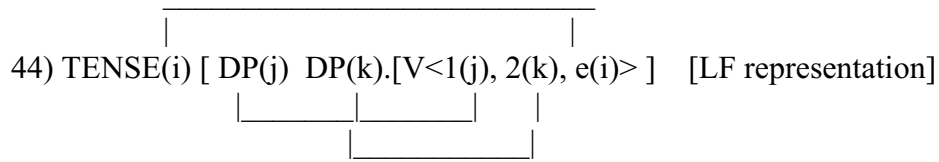
I begin by demonstrating how the operations of MERGE and Theta Discharge apply in a sample derivation of a sentence (41), in which the arguments are saturated constituents, not complex event nominals.<sup>13</sup> A sentence like (41a) is built from the merger of the transitive verb *khooj-naa* ‘search for’ (41b), with its arguments. In addition to internal and external arguments, the argument structure will include the Davidsonian event argument *e*.

- 41) a. ham tumheeN khooj rahee thee  
       we you-dat search prog were  
       ‘We were searching for you.’ Bahri 1992: 138  
       b. *khooj-naa* ‘search (V)’ <1,2,e>

The verb *khooj* ‘search’ merges with the DP *tumheeN* ‘you-dat’, discharging the 2 argument position and the V head PROJECTS upward, with its argument structure. This DP saturates the internal argument position. The result is (42). For concreteness, I use indices to represent the theta-discharge relation between the saturated argument position and the phrase receiving the semantic role. The combination in (42) MERGES with *ham* ‘we-nom’, forming (43):



Finally, the event argument is THETA-BOUND by sentence tense and aspect, anchoring it directly or indirectly within a temporal representation (cf. Smith 1997, Giorgi and Pianesi 1997). The final result of the derivation of the sentence, which goes to the interpretative interface, is the well-formed LF representation in (44):



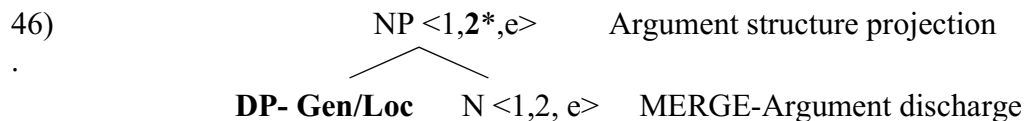
All positions in the argument structure are saturated; all arguments are linked to a position in the argument structure, including the event variable. The event argument is bound by the TENSE.

This sentence derivation and the resulting LF representation conform to Higginbotham's *Saturation Principle*, which can be adopted within the assumptions of the Minimalist Program as an interface condition, relating syntactic structures and LF interpretation.

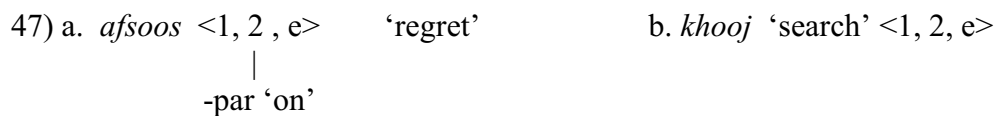
45) *Saturation Principle*: A constituent which has open argument position(s) may not itself discharge a theta role assigned by another head. (Higginbotham 1985: 561-2)

But (45) would be violated if the nominal object is a complex event nominal, with aspectual structure and multiple argument positions to be saturated (Grimshaw 1990).

The derivation of a sentence with a complex predicate of Type I or II works like the formation of a simple sentence in successive stages shown in (46)(48)(49), with one important difference stemming from the Saturation Principle. The N first merges with a DP, which gets the internal argument theta role in the N's argument structure. The genitive or locative case is checked within the nominal projection which is formed by MERGE<sup>14</sup>. The theta grid with one discharged role is projected up to the phrasal node (46):



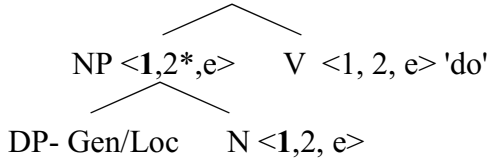
I will assume that the N *khooj* 'search' has no lexical case specified in its lexical entry, but *afsoos* 'regret' as in (38), does assign a lexical case (47a):



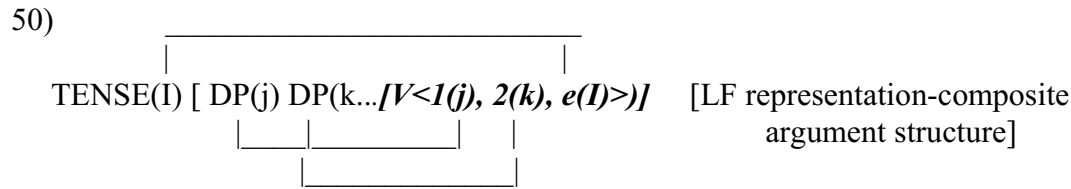
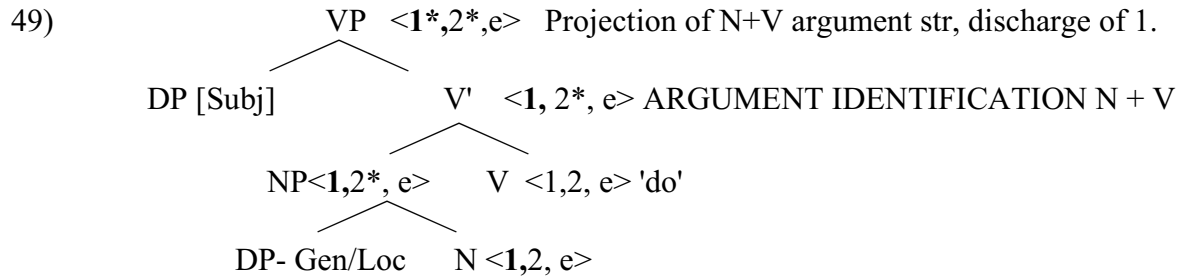
If no lexical case is supplied by the lexical entry of N, then structural genitive is the default.<sup>15</sup>

At this point the derivation can take two distinct paths, one of which is shown in (48). Here the NP with an open theta position undergoes theta identification with V, and does *not* discharge the internal theta role of V. The discharged 2 position of N is identified with the 2 open position of V. The open external argument position of N is identified with the open 1 position of V, as are the event arguments of N and V.

48) V' <1,2\*, e> ARGUMENT IDENTIFICATION<sub>NN+V</sub>

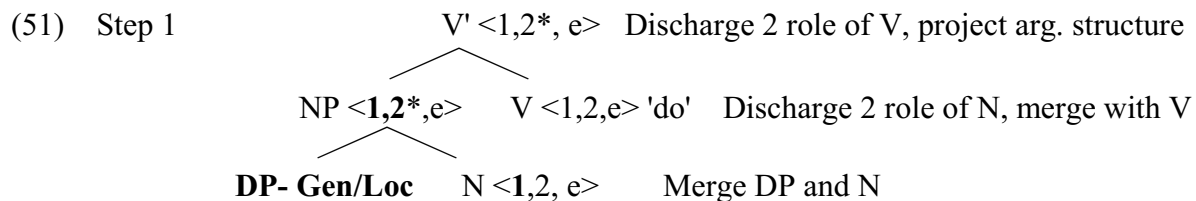


Finally, in (49), the DP external argument merges with the V projection, discharging the external argument theta role. The theta role of the external argument specified by N (**bold**) is available in the composite N-V structure. The LF (50) contains the composite argument structure (**bold italics**):

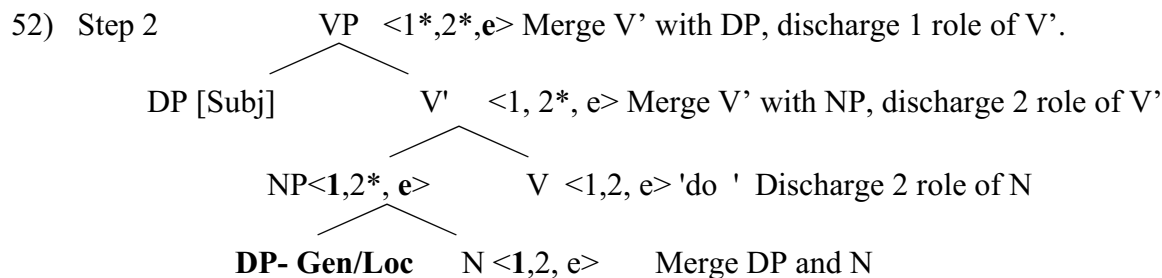


The main feature of this account is that complex event NP itself does not discharge a theta role of V. It does not saturate a position in the argument structure of V. Higginbotham's Saturation Principle (45) would be violated if the N, which has an open theta position in its argument structure, did discharge a theta role of V. Rather, information from the argument structure of NP at that point in the derivation is transferred to V and combined with the specification of the argument structure in V.

Suppose, however, that the event N did discharge a theta role of V. Besides the sequence (46),(48)-(50), there is an alternative derivation, with N saturating an argument position of V. This derivation, however, results in an ill-formed LF representation violating (53). In this derivation, the derivation proceeds as in (46), with merger of N and its thematic object. The difference is in the following steps. In step I (51), the NP merges with V, discharging the internal argument theta role of V:

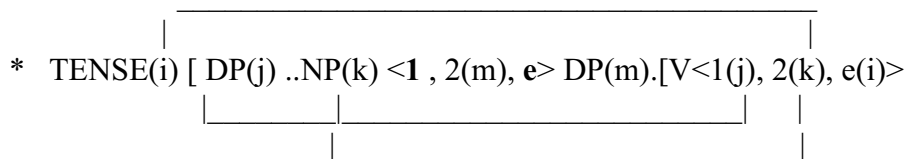


In Step 2 the structure in (51) then merges with the subject DP, discharging the external argument position of V. (52). Note that the argument structure of N, with an open 1 position, remains independent of the argument structure of V.



The result is an ill-formed LF representation (53), with an open, unsaturated argument position. in N, and an unbound event argument (bold)

53) Ill-formed LF representation

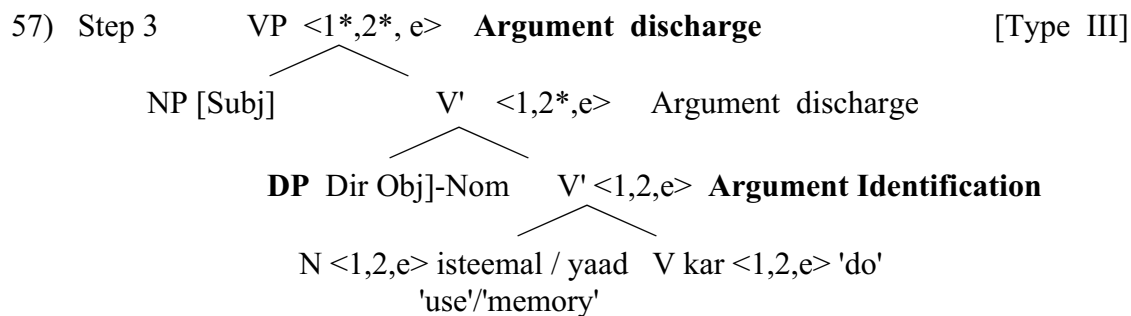
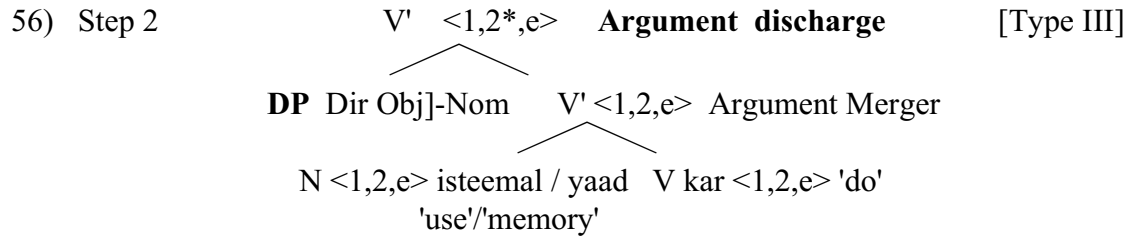
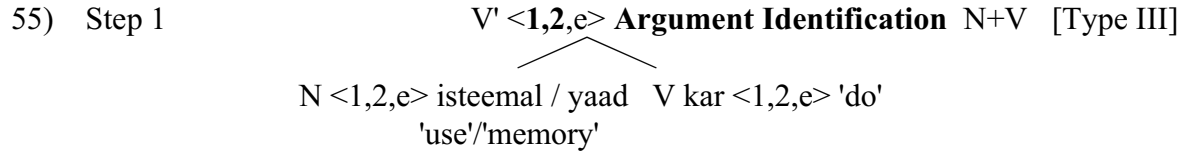


NP(k) is the complex event nominal which discharges the 2 position V, but NP(k) has an open external argument position not linked to a DP. A crucial feature is that there are two separate and independent external argument positions. This result recalls the problematic result of incorporating the event N as a restrictor of the object of V, as in (36) in section 4.4 above. The N's external argument position is still open after the syntactic derivation is completed (53). It is 'left behind', as it is not transferred to a composite argument structure of N+V.<sup>16</sup> The event argument of N is not accessible to binding by TENSE/ASPECT as *e* is within a maximal phrase NP. In the well-formed derivation, the theta merger of N and V means that the external position of N is projected as non-distinct from the external argument position of V, and TENSE binds the common event variable.

Type III complex predicates, which are all variations of Type I predicates, differ in two important ways from Type I. The N does not trigger verb agreement in Type III, and the thematic object has the nominative or dative case of an ordinary direct object. This structure is illustrated by (54).

54) maiN-nee **raam -koo** yaad ki-yaa 'I remembered Ram' (Type III)  
 I-erg Ram-dat memory-f do-pf-ms

The difference is that Argument Identification/ Merger is the first operation, taking place before the DP thematic object is merged (55)-(56). Finally, the subject is merged, as in the other derivations (57):



Argument identification takes place at the X<sub>0</sub> level, before N has combined with a thematic DP object, as it would in Types I and II. The distinction between types I and II predicates follows from the lexical properties of N. The distinction between types I/II and Type III seems to stem from a lexical property of a fairly small number of Ns. They assign either genitive or structural object case. I propose that the Ns which assign nominative or dative structural object case may have a second lexical representation which blocks the assignment of genitive case, based on what Dubinsky 1997 proposes for some complex predicates Japanese. Dubinsky proposes that N loses its full categorial distinction, and is specified only as +N, non-distinct from an adjective. Adjectives may also have argument structures and assign case, but not genitive or lexical case.

In fact, the structural case-marking of objects in the Type III N-V predicates is exactly like what we would expect from adjective-verb combinations (McGregor 1995) Hindi/Urdu has an A+V complex predicates, such as *mazbuur kar-naa* 'make forced, force, insist upon', which assign structural object case, as in (58a). This dative object case is structural, as it disappears in the passive (58b):

- 58) a. unhooN-nee DaakTar-*koo* [PRO fiis lee-nee] par **mazbuur ki-yaa**  
 3pl-erg doctor-dat fees take-inf-on forced do-pf  
 'They forced the doctor [PRO to take his fees].' Bahl 1979: 37
- b. us-kaa jiiivan phir caTTanoon-par [PRO Takaraa-nee] -kee liyee **mazbuur**  
 3s-gen life again rocks -on dash-against-inf-gen-for forced  
**ki-yaa ga-yaa hai**

do-pf go-pf is  
'Her life again was forced to dash against the rocks.' Bahl 1979:37

These case-marking facts follow from the inability of adjectives to assign object case, so that Argument Identification is forced to form an A-V unit which assigns structural case. Early argument merger is therefore forced for adjectives in general and for Ns which have a defective category specification. Early merger is a consequence of category specification in the lexicon. Some Ns, like *yaad* 'memory' and *isteemaal* 'use', have dual lexical representations, allowing them to combine in both Type I and Type III combinations. Optionality is a consequence of dual lexical representation, in this view.

This dual representation of N as [+N-V] (Types I, II) and [+N] (Type III) provides an explanation for the differences of case and phrase structure. The object merges with N low in the structure for Types I/II, while it is higher in the structure for Type III. The object position follows from the point in the derivation at which argument merger takes place. Early argument merger is forced when N or A cannot assign case, deriving type III structures in which the DP object merges higher up with N-V, while type I/II structures are derived when N does assign case to its object, allowing the DP with N, so that in the resulting clause structure it is quite low. It is not necessary to stipulate the high versus low position of the object, the criticism leveled at the structures (26b)-(27b). I've argued above that such stipulations are required if complex predicates have the same structure as incorporation/conflation structures.

#### 5.4 Broader applications for early versus argument identification

The formation of Arabic masdars, or verbal nouns (Fassi Fehri 1993: 232ff) exhibits different syntactic results from early or late merger and theta/argument identification. The Arabic masdars form a nominal from a verbal root, and are in some sense the inverse of the N + V combinations which form a predicate from a N argument structure with the addition of a verb. Verbal roots combine with a nominalizing suffix, forming something like a gerund with a genitive subject and an object. The object may have accusative case, which Fassi Fehri accounts for as the merger first of the object with the verbal root assigning case, then merger with the nominalizing affix. The result is like the Type III N + V compounds in Hindi/Urdu. In the second option, the verbal root merges first with the nominalizing affix, then with the direct object, which cannot have accusative case because the case licensor is a nominal. For that reason, the object is marked with a preposition, like the object of a noun. In both derivations, when the nominalizing affix combined with V, the event nominal of the verbal root is theta-identified with the event argument of the nominalizing affix (Higginbotham 1985), the sole argument of this affix. Fassi Fehri uses the evidence of Raising to Object within the masdar phrase to show that case is assigned to a DP in a derived position. If so, then masdars are formed by syntactic derivation rather than in the lexicon (Fassi Fehri 1993: 220)

The early or late options for argument identification also account for the two forms of 'argument transfer' in Grimshaw and Mester 1988; it can be partial or complete. If N combines only with its thematic argument, theta identification follows this combination and is 'late' or higher in the phrase structure; this is partial argument transfer. Complete argument transfer can be viewed as the same as 'early' V-N merger with argument/theta identification, so that the V-N

combination case marks all the arguments. In my view, the syntactic derivation in both cases uses the same mechanism of theta identification, which may take place early or late. These cases offer support for the claim that theta identification is a universally available possibility

## 6. How N and V combine to form a single predicate

I return here to the question of how V and N contribute to the meaning and aspectual properties of the N-V combination formed by argument merger motivated by the saturation Principle (45), focussing in greater detail here about lexical information and how it is combined when argument structures are identified. A somewhat similar proposal based on argument merger has been made for Persian (Karimi-Doostan 1997), though the structural differences of case and lexicon between Hindi/Urdu and Persian entail some formal differences from the account given here, which derives the variety of complex predicates in Hindi, counterparts of which are apparently not found in Persian. The argument merger approach was explored in earlier work by Jackendoff (1970) and Cattell (1984).

I assume that V in N-V combinations is as fully specified as in independent occurrence without the event N, in contrast to earlier accounts, Grimshaw and Mester 1988 in particular, who represented V as ‘light’, devoid of thematic properties. But more recent discussions such as DiSciullo and Rosen 1990 point out that V in N-V is far from empty of specification. My position is that the kind of verbs which enter into N-V combinations (Appendix III) have thematic properties and aspectual properties, but are inherently *less* specified than other verbs, such as those which incorporate indefinite objects like *samhaal-naa* ‘look after’, discussed above. The complex event nouns, on the other hand, are inherently more fully specified, not only for event structure and arguments but also for various specific default properties for the arguments. This asymmetry of lexical content accounts for the generally greater contribution of N’s properties when the argument structures are merged. The V and complex event N have similar lexical structures, so that corresponding information can be matched. I will use the lexical structures of Pustejovsky 1995, somewhat simplified, to focus on a representation of the lexical content of N and V.

For a concrete example of the same N combined with different Vs, compare the complex predicates in (59):

- (59) aur us-see mujhee bahut-sii baateeN **yaad aa gayii thiiN** jinheeN maiN  
 and 3s-from I-dat many-rather matter-pl memory come go-pf were which-dat I  
**yaad nahiiN kar-tii**  
 memory not do-impf  
 ‘From that I remember many things which I do not want to recall/think about.’  
 Bahl 1974: 73

This sentence contrasts two versions of ‘remember’, the dative subject version with *aa-naa* ‘come’, and the nominative/ergative subject with *kar-naa* ‘do’. This contrast illustrates a subtle aspectual difference. The dative subject version refers to an involuntary process, instigated by another event *us-see* ‘from that’, and leading immediately to a resulting state of remembering. This version has the properties of an achievement. The nominative/ergative version is more an

accomplishment, focussing on a process which takes some time. Some volition may be conveyed.

The aspectual property of the N-V follows from the specification of the V involved. Following Pustejovsky 1995, I represent the event denoted by V in terms of subevents, which are in a precedence relation. A subevents may be the head of the whole event (distinguished by \*), representing the focus of the event. Accomplishments focus on the initiating process, achievements on the resulting state (60a,b).

- (60) a. achievement: E = e1:process e2: state precedence: e1 < e2\* (<= precedes)  
**aa-naa** ‘come, arrive, happen’
- b. . accomplishment E = e1: process, e1:state precedence:: e1\* <e2  
**kar-naa** ‘do, make, bring about’
- (61) underspecified: E = e1: process e2:state precedence: e1 <o e2 (<o =precedes  
**yaad** ‘memory’ and overlaps)

If *yaad* ‘memory’ is ambiguous between the state of remembering and a process of bringing something to memory, the N lacks a specified head. In combination of N with a V with a specified head, the N-V takes on head specification, with the results exemplified in (59). The accomplishment meaning of *khooj kar-naa* ‘discover’ stems from the property of V *kar-naa* which focusses on the process leading to a result.

The argument structure of N *yaad* ‘memory’ specifies the arguments of the two subevents and places some restrictions on their denotations.

- (62) N *yaad* ‘memory’
- a. Argument structure: Arg 1 = x: proposition<sup>17</sup>  
Default argument 1 = y: human<sup>18</sup>
- b. Qualia = Formal = ‘remember’ (e2, y, x)  
Agentive= exp\_act (y, e1)

The N has a syntactic argument x, the internal argument. The default argument of *yaad* ‘memory’ is the human y who remembers; a default argument need not be projected syntactically but is implied. The *qualia* are descriptors of the word’s meaning; the formal quale here defines the relations between the arguments. The experiencer role for the y external argument follows from the restrictions that y is human and x is a proposition, and the agentive quale. This quale defines how the event comes about; here by the experiencer undergoing some process event (which may be expressed separately, by *un-see* ‘from that’ in (59) above, or implied by the context.

The argument structures of the Vs are similar to the N, but the restrictions on the arguments are less specific than those of the N:

- (63) V *aa-naa* ‘come, arrive, happen’
- a. Argument structure: Arg1 = x: individual entity  
Default arg1 = y: location
- b. Qualia = Formal = to (e2, x, y)

Agentive = ‘come’ (e1, x)

In combination V (63) with N *yaad* ‘memory’ the Arg 1 of (63) is identified with the proposition of (62), while the location of (63) is identified with the human experiencer default argument of N. The qualia of the N are more specific than the change of (abstract) location of (63); the ‘location’ default argument is projected as a dative NP.

The V *kar-naa* ‘do, make, bring about’ has the properties of a very general causative predicate:

- (64) V *kar-naa*
- a. Argument structure: Arg 1 = x  
Arg 2 = y
  - b. Qualia = Formal = result (e2, y)  
Agentive= make\_act (x,y)

The more specific argument restrictions and qualia of N narrow down the very general definition of V, once the arguments of N and V are identified. The default experiencer argument of N is identified with the syntactically projected Arg 1 of V. This combination has a range of interpretations, some agent-like and others less volitional. For example a sentence like (4) can be used to refer a volitional or involuntary act. This lack of fixity stems from the contributions of somewhat different but compatible information from N or V. It also explains how the experiencer interpretation arises in combination with a verb which normally has an agentive interpretation elsewhere.

The explanation I have proposed for complex predicates is that they are formed by completely general principles, syntactic operations of MERGE and theta related processes. Argument identification is forced in complex predicates by Higginbotham’s Saturation Principle, which guarantees well-formed syntactic objects which can be interpreted at the LF interface. There is no option for avoiding N-V argument identification, given a complex event N with an array of multiple arguments, and a verb which might assign it a theta role.

## 7 The optionality of complex predicate meanings

Should an account of complex predicates allow for a derivation *without* argument identification/theta merger? There are actually two questions here. One is whether complex predicates are always syntactically derived, with predictable, transparent meaning derived from the component parts. The alternative would be an idiomatic meaning associated with a phrasal rather than Xo category, in effect preempting syntactic derivation in which argument identification takes place. The other question is whether a complex predicate may be ambiguous between a unitary predicate meaning, and a literal, phrasal meaning.

Some N-V combinations have metaphorical or non-literal meanings. Many examples are found in Persian, such as *dast keshidan* ‘hand pulling = to touch’ (Folli et al. to appear). Barjasteh 1983 proposes that all N-V combinations are lexical units, since it is not possible to draw a clear line between the metaphorical and the syntactically formed combinations. Nunberg et al 1994, however, argue that even idiomatic phrases are formed in a syntactically transparent

way, reflecting the general patterns of the language in which they occur. The closest analogues to the Persian example which I can find in Hindi/Urdu are combinations of N with *maar-naa* ‘beat’. This verb combines with a N referring to some manner of striking, such as *thappaR* ‘slap’ or instrument *goolii* ‘bullet’, *juutaa* ‘shoe’, and then takes a referential argument x. The results mean ‘slap x, hit/kill x with a bullet’, ‘beat x with a shoe’ (Appendix III) These Ns are bare indefinites, and it is plausible that they refer to instruments, not events, and do not have argument structures of their own. Instead, the N is incorporated as a modifier of V.<sup>19</sup> So these examples can be explained as indefinite N incorporation like (36), rather than a complex predicate formed with an event N.

Literal meanings of N-V are possible if N is ambiguous between is a result N, without an argument structure, and a complex event N with an argument structure including an external argument. The ill-formed LF structure in (53) has an open, unsaturated external argument position. A result N lacks obligatory arguments (Grimshaw 1990), allowing result Ns to occur without argument merger where it would be forced for complex event Ns. The N *bhuul* ‘oversight, omission, mistake’ is a possible example of a result N which combines with *kar-naa* with the literal meaning ‘make a mistake’ (not synonymous with V *bhuul-naa* ‘forget’). In (60a), the N *bhuul* ‘mistake’ is inflected for plurality, and modified. The NP-V phrase in (60b) refers to the NP in (60a).

- 60) a. *bhaaSaa-meeN ling - sambandhii bhuul-eeN sab-see burii samjh-ii jaatii haiN.*  
 language-in gender-agreement mistake-pl all-from bad understand-pf go-impf are  
 ‘Gender and agreement mistakes are considered the worst in language.’ Bahl 1974:16
- b. [*kisii-koo is prakaar-kii bhuul-eeN kar-tee*] *deekh-kar loog praayah kah*  
 some-dat this kind-gen mistake-pl do-impf see-prt people often say  
*baiTh-tee haiN -- uNh, unheeN too striiling aur puling -tak -kaa jnaan nahiiN hai*  
 sit-impf are Oh 3pl-dat emph feminine and masculine-to-gen knowledge not is  
 ‘Seeing someone make this kind of mistakes, people can’t help saying, ‘Oh, they don’t  
 even know feminine and masculine.’ Bahl 1974:16

The N *baat* ‘what is said, piece of information, conversation’ is another result N which occurs independently of V, but in combination with V *kar-naa* it forms a complex predicate ‘to talk with (about)’ with a locative object suggesting the complex argument structure of a complex event reading. It does not seem to be the case that N-V combinations are ambiguous in Hindi/Urdu between a literal result meaning and a complex predicate meaning; either they have indefinite object readings (eg. *kaam kar-naa* ‘do work, work’) or they form composite readings with combined N and V argument structures.<sup>20</sup>

## 8. Summary and conclusions

In the course of arguing for a syntactic derivation of complex predicates in Hindi/Urdu, I have compared two different derivational methods for which there is much evidence in human languages. I have contrasted different predicate types in Hindi/Urdu, including three types of complex predicate, defined by the thematic object case. Combinations of complex event Ns can be related to V in several different ways in this language. Simplex verbs related to a N are

plausibly derived by Hale-Keyser (1993) conflation or incorporation in I-syntax, using a simple or complex VP structure. Application of the same structure to complex predicates fills the V position with a lexical verb and yields only one syntactic subclass (Type III). It predicts the wrong phrase structure for the other types, and does not explain why the composite predicate has the argument structure and semantic roles of N. In particular, it does not explain the large class of N-V predicates with experiencer subjects. Covert incorporation is freely available for indefinite common noun objects, which are predicates restricting the meaning of the V. Complex event Ns with their own argument structure (Grimshaw 1990) cannot coherently be interpreted as V with its array of arguments. This kind of incorporation does not explain why N-V combinations have a single argument structure formed from the argument structures of N and V.

This kind of combination can be derived by completely general processes, the syntactic operation of MERGE (Chomsky 1995) and argument/theta identification, which is required by an LF interface condition which has the effect of disallowing a NP with open argument positions from saturating an argument position of V (Higginbotham 1985). Argument identification may follow combination of N with its thematic object (Types I and II), or precede it (Type III), yielding the differences of thematic object and verb agreement noted above. It is not necessary to postulate special argumentless 'light' verbs. The kind of verbs which participate in these combinations in HU are inherently less specified or restricted by default information than common event nominals (cf. Pustejovsky 1995). In the course of argument identification, N contributes its richer information to the combination of corresponding N and V arguments. V contributes its aspectual type and the case required for the subject. The lexicon of Hindi/Urdu determines some options. It distinguishes between argumentless result nouns, and complex event nominals with argument arrays. Some event nouns which participate in Type III predicates have dual lexical forms, full Ns and N indistinguishable from A. Language specific properties of lexicon and the case/agreement features derive differences of complex predicates in different languages. Otherwise complex predicates are derived by universally available and completely general syntactic and semantic principles. In Hindi/Urdu, a nominal may combine in three ways with V. If the nominal is a common noun, it may discharge an argument position of V. Alternatively it may restrict V semantically, by (pseudo) incorporation. If, however, N is a complex event nominal, it must undergo argument identification with V.

#### Appendix I- Ns which participate in Class III combinations

- a. yaad (f) memory
- b. isteemal (m) use' prayoog (m) use'
- c. qatal (m) 'murder'
- d. anubhav (m) 'experience, feeling'  
ahsaas (m) perception, feeling'
- e. talaash (f) 'search' cf. V talaash-naa 'search for'
- f. kharc (m) expense V. kharc-naa 'spend'
- g. niilaam (m) auction
- h. naql (f) copy, imitation
- i. kalpanaa (f) imagination

- j. daan (m) gift
- k. pariikSaa (f) examination; niriikSan (m) test
- l. praNam (m) 'respectful greeting'
- m. kSamaa (f) forgiveness, pardon
- n. shuruu (m), arambh(m) 'beginning'
- o. xatm (m) and Adj . 'end/finished'
- p. sviikaar (m) 'acceptance'
- q. jamaa (f) and Adj 'collection'
- r. mool m. price, value + lee-naa 'take' = 'sell'
- s. qarz (m) debt + lee-naa 'take' = 'borrow'

#### Appendix II- Verbs in N-V combinations (not exhaustive)

##### 2) [nominative or dative subjects]

- a. aa-naa 'come (to), become, happen' (Type I and III predicates possible)
- b. hoo-naa 'be, become'
- c. lag-naa 'be attached to, strike'
- d. cal-naa 'go, move, start, advance'
- e. pahuNc-naa 'arrive, reach'

##### 3) [nominative/ergative subjects]

- a. kar-naa 'do, make' (Type I, II and III predicates possible)
- b. dil-aa-naa 'cause to give, assign, put in possession of.'
- c. rakh-naa 'place, put, keep'
- d. dhar-naa 'hold, keep, possess, detain, apply'
- e. lag-aa-naa 'cause to strike, put'
- f. uThaa-naa 'raise'
- g. pahuNc-aa-naa 'cause to arrive'
- h. paa-naa 'find'
- i. maar-naa 'beat'.
- j. dee-naa 'give'

##### 4) yaad 'memory' +

- a. aa-naa 'come.' = 'remember, come to mind' [Dative subject]
- b. paR-naa 'fall.' 'remember, come to mind' [Dative subject]
- c. dilaa-naa 'cause to give/cause to get' = 'remind, remember NP to another'
- d. rakh-naa 'place' = 'bear in mind'
- e. rah-naa 'stay' = 'remain in mind, not be forgotten'
- f. hoo-naa 'be/become' = be in mind, be remembered, be learned by heart [Dative subject]

#### Appendix III - Incorporation idioms

##### 5) *N. khaa-naa* 'eat N' [ergative subject]

- a. dhookaa khaa-naa 'eat deception, be deceived',

- b. gam khaa-naa 'eat sorrow, suffer'  
 c. Thookar khaa-naa 'eat stumbing, stumble'  
 d. maar khaa-naa 'eat a beating, get beaten',  
 e. GoTaa khaa-naa 'eat a dive, get a ducking, nearly drown' (Bailey 1962, 59).  
 f. cooT 'eat injury, get injured'  
 g. dhakkee khaa-naa 'eat shoves, suffer indignities, misfortunes' (Nespital 284)  
 h. qasam khaa-naa 'eat an oath, swear an oath'  
 i. haar khaa-naa 'eat defeat, get defeated'  
 j. jeel-kii hawaa khaa-naa 'eat the jail air, serve a term in jail'
- 6) us-nee patthar-see **Thookar khaa-yii**  
 3s-erg stone-from stumble eat-pf  
 'He/she stumbled on a stone.' (Porizka 1963 439)
- 7) *N maar-naa* 'beat N' [ergative subject]  
 a. goolii maar-naa 'beat a bullet' (-koo/-par) 'shoot/shoot at'  
 b. juutaa maar-naa 'beat shoe(s) (-par) 'give a shoe-beating, insult'  
 c. thappaR maar-naa 'beat a slap' (-par) 'slap'
- 8) . gabbar singh-nee din dahaaRee laalaaaram-koo **goolii maar dii**  
 Gabbar Singh-erg day growl-pf-obl Lalaram-dat bullet beat give-pf  
 'Gabar Singh shot Lalaram in broad daylight.' Barz and Yadav: 186

## References

Bahl, Kali Charan

1974 *Studies in the semantic structure of Hindi*. Vol. I. Delhi: Motilal Banarsidass.

1979 *Studies in the semantic structure of Hindi*, Vol. 2. Delhi: Manohar.

Bahri, Hardev

1992 *Learner's Hindi Dictionary*. Delhi: Rajpal and Sons.

Bailey, T. Grahame

1956 *Urdu*. London: English Universities Press.

Baker, Mark

1996 *The polysynthesis parameter*. Oxford: Oxford University Press.

Barjasteh, Darab

1983 Morphology, syntax and semantics of Persian compound verbs: a lexicalist approach. [Unpublished dissertation University of Illinois, Champaign-Urbana

Barz, Richard K. -- YogendraYadav

1993 *Introduction to Hindi and Urdu*. New Delhi: Munshiram Manoharlal.

- Bhaskararao, Peri -- K.V. Subbarao (eds)  
 2004 *Non-nominative subjects*, vol. 1., Amsterdam: John Benjamins
- Butt, Miriam  
 1995 *The structure of complex predicates in Urdu*. Stanford: CSLI Publications
- Butt, Miriam -- Gillian Ramchand  
 2001 Complex aspectual structure in Hindi-Urdu. In B Jensen at al, *Oxford Working Papers in Linguistics 6*. Oxford: Committee for Comparative Philology and General Linguistics
- Cattell, Ray  
 1984 *Composite predicates in English*. Syntax and semantics, 21. Sydney: Academic Press.
- Chomsky, Noam.  
 1995 *The Minimalist Program*. MIT Press.
- Chung, Sandra -- William Ladusaw  
 2003 *Restriction and Saturation*. Linguistic Inquiry Monograph 42. Cambridge: M.I.T. Press.
- Davison, Alice  
 2000 "Lexical pronouns and anaphors in Hindi/Urdu." In B, Lust --K. Wali--J. Gair--K.V.Subbarao (eds.), 397-471  
 2004 "Non-nominative subjects in Hindi/Urdu: VP structures and case parameters." In P. Bhaskararao and K.V. Subbarao (eds) 141-168  
 to appear "Structural case, lexical case and the verbal projection." In V. Dayal A Mahajan, (eds)
- Dayal, Veneeta  
 2003 A semantics for pseudo-incorporation. [Unpublished MS], Rutgers University.
- Dayal, Veneeta -- AnoopMahajan (eds)  
 to appear *Clause structures in South Asian languages*. Dordrecht: Springer/Kluwer Academic Publications.
- DiSciullo, Anna-Maria --Sarah Thomas Rosen  
 1990 "Light and semi-light verb constructions." In K. Dziwirek-- P. Farrell -- E. Mejias-Bikandi (eds) 109-127.
- Dubinsky, Stanley  
 1997 "Syntactic underspecification and light-verb phenomena:," *Linguistics* 35, 627-677.
- Dziwirek, K-- Patrick. Farrell -- E. Mejias-Bikandi (eds) *Grammatical relations: a cross-*

*theoretical perspective*. Stanford: CSLI,

Fassi Fehri, Abdelkader

1993 *Issues in Arabic clauses and words*. Dordrecht: Kluwer Academic Publications.

Folli, Rafaella

2001 Constructing telicity in English and Italian. [Unpublished dissertation, Wolfson College, Oxford]

Folli, Rafaella-- Heidi Harley --Simin Karimi

to appear "Determinants of event type in Persian." *Lingua*.

Gambhir, Vijay

1993 "Complex verb phrase: a diachronic and synchronic view." In Manindra Verma, (ed), 77-96.

Grimshaw, Jane

1990 *Argument structure*. Cambridge: MIT Press.

Grimshaw, Jane -- Mester, Armin

1988 "Light verbs and  $\theta$  marking." *Linguistic Inquiry* 19, 205-232

Hale, Kenneth --Samuel Jay Keyser

1993 "On argument structure and the expression of syntactic relations." In K. Hale -- S.J.Keyser (eds), 53-109

Hale, Kenneth --Samuel Jay Keyser Samuel Jay (eds)

1993 *The view from Building 20: essays in honor of Sylvain Bromberger*. Cambridge: M.I.T. Press

Higginbotham, James

1985 "On semantics." *Linguistic Inquiry* 14, 545-597.

1999 Accomplishments. [Unpublished paper, University of S. California]

Hook, Peter

1973 *The compound verb in Hindi*. Ann Arbor: University of Michigan Center for South and Southeast Asian Studies.

1979 *Intermediate Hindi Structures*. Ann Arbor: University of Michigan Center for South and Southeast Asian Studies.

Jackendoff, Ray

1974) "A deep structure projection rule." *Linguistic Inquiry* 5, 481-506.

Jayaseelan, K.A.

1988 "Complex predicates and theta -theory." In.Wendy Wilkins Wendy (ed.), 91-111.

Kachru, Yamuna

- 1982 "Conjunct verbs in Hindi and Persian." *South Asia Review* 6.3.
- Karimi, Simin  
1997 "Persian complex verbs: idiomatic or compositional?" *Lexicology* 3.1 273-318
- Karimi-Doostan, Gholamhossein.  
1997 Light verb constructions in Persian and Kurdish. [Unpublished dissertation, University of Essex].
- Lust, Barbara--Kashi Wali-- James Gair--K.V. Subbarao, (eds.)  
2000 *Lexical anaphors and pronouns in selected South Asian languages*. Berlin: Mouton deGruyter
- McGregor, R.S.  
1995 *Outline of Hindi grammar* (3rd edition). Oxford: Oxford University Press.  
1997 *Oxford Hindi-English dictionary*. Oxford: Oxford University Press.
- Mohanan, Tara  
1994 *Arguments in Hindi*. Stanford: CSLI Publications.  
1995 "Wordhood and lexicality: Noun incorporation in Hindi." *Natural Language and Linguistic Theory* 13.1, 75-134.
- Montaut, Annie  
to appear *Hindi Grammar*. Munich: LINCOM-Europa
- Nespital, Hellmut  
1997 *Dictionary of Hindi verbs*. Allahabad: Lokbharati Publishers.
- Nunberg, Geoffrey --Ivan Sag--Thomas Wasow  
1994 "Idioms." *Language* 70: 491-538.
- Porizka, Vincenc  
1963 *Hindstina, Hindi Language Course*. Prague: Stani Pedagogike Nakladatelstvi.
- Pustejovsky, James  
1995 *The generative lexicon*. Cambridge: MIT Press.
- Rosen, Carol Thomas  
1990 *Argument structure and complex predicates*. New York: Garland.
- Saxena, Anju  
1985 "Reflexivization in Hindi: a reconsideration." *International Journal of Dravidian Linguistics*. 14: 225-237.
- Speas, Margaret  
1990 *Phrase structure and the lexicon*. Dordrecht: Kluwer Academic Publications.

- Smith, Carlota  
1997 *The parameter of aspect*. Dordrecht: Kluwer Academic Publications.
- Terada, Michiko  
1990 Incorporation and argument structure in Japanese. [Unpublished dissertation, University of Massachusetts]
- Verma, Manindra K  
1993 "Complex predicates and light verb in Hindi." In M. Verma (ed.) 197-215.
- Verma, Manindra (ed)  
1993 *The complex predicates in South Asian languages*. Delhi: Manohar Publications.
- Wilkins, Wendy (ed.)  
1988 *Syntax and semantics 21, Thematic relations*. San Diego: Academic Press

## NOTES

1. Previous versions of this paper were presented at the University of Iowa Department of Linguistics Colloquium, the Conference on Argument Structure, Delhi University (January 2003), and in colloquia at the Universities of Potsdam and Konstanz in 2004. I thank the audience at these presentations for their discussion and helpful comments. The Obermann Center for Advanced Study, University of Iowa, provided a supportive atmosphere for the writing of previous drafts of this work. I am grateful to Carlota Smith, Paula Kempchinsky, Stanley Dubinsky and Simin Karimi for commenting on earlier manuscript forms of this paper. Rashmi Gupta, Rajiv Sahay, Manish Srivastav, and Sanjay Gupta generously contributed native speaker judgements.
2. The N *yaad* combines with a genitive-marked theme 'memory about X', not 'X's memory about something.'. The ability to remember is expressed with a different N, *yaad-daasht*.
3. Idiomatic passive sentences which express the agent are negative and have the 'ability' interpretation.
4. See Karimi 1997 for other kinds of argument from Persian that the N component of N-V combinations is a syntactic direct object.
5. The voiced velar fricative G is not found in the Indo-Aryan vocabulary, only in Perso-Arabic borrowings.
6. The simplex verb may have the accomplishment sense with the addition of a vector verb like *lee-naa* 'take' (Hook 1974, Butt 1995). The addition of a compound or vector verb in realis contexts (ie. not negative, interrogative, etc.) introduces a telic sense; see Butt and Ramchand 2003 for an event-structure account of this property.

7. What this intransitive structure does predict is the projection of an agent as an adjunct, suggesting that the reduced stem still encodes the transitive argument structure of the full stem of the V or N *kaaT* ‘cut’:

- (i) *maalii* -see *yah peeR kaT ga-yaa*  
 gardener from this tree be-cut go-pf  
 ‘This tree got cut by the gardener (unintentionally).’

8. Analysis of events as consisting of sub-events in some specific temporal relation has proven very productive; see section 6 and Pustejovsky 1995.

9. The kind of N incorporation proposed in Baker 1996 is motivated as an option because it provides an alternative to case assignment for a syntactic object. This view of incorporation explains the licensing of N, but says nothing about the licensing of N’s DP thematic object. By many criteria, Hindi/Urdu is not a polysynthetic language with optional syntactic N incorporation.

10. The N *baccee* is translated as a predicate which modifies the direct object argument (y) (Chung and Ladusaw (2003):

- (i) a. *samhaal-naa* ‘look after’  $\lambda y\lambda x\lambda e$  [look after’ (y)(x)(e)]  
 b. *baccaa* ‘child’  
 c. Restrict ( $\lambda y\lambda x\lambda e$  [look after’ (y)(x)(e)], **child**)  
 d. =  $\lambda x\lambda y\lambda e$  [look after’ (y)(x)(e)  $\wedge$  **child**’(y)]

11. The N *yaad* comes with its own argument structure (ib). The operation of restrict would seem to apply as in (ic,d), so that ‘w remembering z’ is a property of the (y} object of ‘do’:

- (i) a. *kar-naa* ‘do/make’  $\lambda y\lambda x\lambda e$  [do’ (y)(x)(e)]  
 b. *yaad* ‘memory’  $\lambda w\lambda z\lambda e$  [remember’ (w)(z)(e)]  
 c. Restrict ( $\lambda y\lambda x\lambda e$  [do’ (y)(x)(e), **remember**’ (y)(x)(e))  
 d. =  $\lambda x\lambda y\lambda e$  [do’ (y)(x)(e)  $\wedge$ , (**remember**’ (w)(z)(e))(y)]

12. This discussion does not assume a finite list of discrete theta roles. ‘Theta discharge’ will be taken to mean an association of a phrase with an argument position which has some semantic role derived from lexical information. See section 6 for a characterization of lexical information.

13. To simplify the exposition in what follows, I have not shown the verbal projection in detail, omitting the vP projection and glossing over case-marking features for ergative subject case and dative/nominative object case.

14. The genitive case in HU is somewhat different in nature from the locative cases, which are clearly instances of lexically selected, theta role-related cases. Genitive seems to be the default case which links N and some DP, regardless of whether there is a possessor relation or some

other relation (like thematic object). But I am assuming that genitive, like locative, is checked early at the first instance of MERGE.

15. There is some counter-evidence to the claim that locative case in Type II predicates is exactly comparable to genitive case. As noted in Verma 1993, a locative PP in isolation does not seem to form a grammatical combination with N, in contrast to genitive PP:

- i) a ??siitaa-see preem ‘love from Sita’      b. siitaa -kaa preem ‘love of/for Sita’

Jayaseelan 1998 makes the same observation about Malayalam, arguing that the theta roles of N are not discharged early in derivation, creating an ill-formed constituent. Instead N discharges a semantic role of V and then projects its argument structure up to V', where it undergoes union with the open argument positions roles of V to form a single argument structure. In my proposal, case assignment is not delayed until after merger with V. I leave this question open. The answer may be related to the question of why complex event Ns in Hindi/Urdu do not have their full array of syntactic arguments without the presence of a verb (Davison 2000).

16. Terada (1990) avoids this problem for Japanese by coindexing a PRO subject of N with the subject of V. Each subject would have its own theta role, so that there would be an agentive subject interpretation in all instances as well as whatever role N assigns to its subject.

17. Pustejovsky (1995:118) argues that verbs of mental state or propositional attitude (and presumably the related nominals) select semantically only for a proposition. Its complement may be a finite sentence, directly satisfying this requirement, or by an infinitive or a NP. A proposition can be derived (by coercion) from NP; the proposition contains the referent of NP.

18. This explains the interpretation of (i) focussed on the internal argument, and the ill-formedness of (ii):

- i) raam-kii yaad  
Ram-gen memory ‘The memory of/about Ram.’

- ii) ?? meerii raam-kii yaad  
my Ram-gen memory ‘My memory of Ram’

19. The result of combining *juutaa maar-naa* ‘shoe-beat’ as a modifier (Chung and Ladusaw 2003) would be (i)

- i)  $\lambda y \lambda x \lambda e$  {beat' (y) (x) (e)  $\wedge$  with-shoe'(e)}

20. Dubinsky (1997) discusses the differences in Japanese between the literal or ‘heavy’ meaning and the complex predicate meaning of some N-V combinations. The literal meaning is always agentive, because of the lexical meaning of the verb *suru* ‘do’, and the N-V is

paraphrased as 'x accomplishes an act of a specific type'. It refers to a single instance of the event at one time, not indefinite multiple events; this feature recalls the difference between referential and incorporated indefinite Ns in section 4.4. Dubinsky refers to the heavy or literal reading of N as not having aspectual structure, a characteristic of result Ns.

The specific properties of Japanese case marking limit possible readings. If the N is marked with the -o object marker and combines with P-gen with an argument it forms a NP, and has only the 'heavy' or literal meaning. Ns which have no case or -o and are not linked by genitive case to postpositionally marked arguments, may have either the literal or complex predicate meaning. It is hard to compare Japanese and Hindi/Urdu, because of the differences of the case marking principles, and in the different choices of Ns which participate in N-V compounds in the two languages. It is not clear whether the event/result N ambiguity would account for the Japanese facts.