

**ARGUMENT LICENSING AND INTERFACE CONDITIONS:
THE MOTIVATION FOR COMPLEX PREDICATES IN HINDI/URDU**

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

April 13, 2004

1. Introduction

Complex predicates are widely used in human languages. Many languages form new predicates by combining a verb with some other category, N, A, P or even another V. The resulting combination has the properties of a single predicate with one argument structure, but this argument structure is a combination of properties of both the N and the V. The two component parts normally retain enough syntactic independence that they cannot be regarded as single lexical entry made up of two words. In a language like Hindi/Urdu, the subject of this paper, complex predicates formed from N or A plus V represent a very productive source for new predicate formation, superceding other ways of deriving a verb from N or A which were productive at earlier stages of the language.

- 1) N yaad ‘memory, a memory, a recollection’ + V kar-naa ‘do-inf’
=TRANSITIVE VERB ‘to remember, recollect, recall, miss, commit to memory’
> <exp, theme>
- 2) *maiN-nee/*mujhee* -[[*un-kii yaad kii* ‘I remembered/recalled them’
I-erg /*dat 3pl-gen-fs memory-fs do-pf-fs
<Exp> <theme>

Many previous accounts, including those which adopt the assumptions of Principles and Parameters theory (Chomsky 1981) have treated the combinations as lexical entries, the product of some derivational process in the lexicon. These accounts observe the Projection Principle, which preserves the same argument structure at all levels of syntactic derivation. Within Minimalism (Chomsky 1986, 1995, 2000), the Projection Principle constraining syntactic derivations is replaced by conditions on the PF and LF Interfaces. This change in assumptions allows merger or identification of two separate argument structures as one to take place in the syntactic derivation.. In fact, interface conditions on LF require it. In this paper, I will propose a syntactic derivational account of these N/A combinations, based on the event structures of N and V and constrained by the LF interface condition.

The productivity of formation and the morphological/syntactic regularity of these combinations suggest that these combinations are derived in the syntax, not derived separately and stored in the lexicon. The combinations regularly and transparently reflect the semantic, syntactic and

condition which prohibits undischarged argument positions. This interface condition is met by merging the argument structures of N/A and V according to a principle of compatibility of argument structures. In keeping with recent views of 'light' verbs, I will assume that the V component is a normal independently occurring and frequent verb which is underspecified for subcategorization and selectional restrictions, but is able to take information from its syntactic arguments. (cf Pustejovsky 1995). The N will bear the detailed semantic information for the combination as well as certain case specifications.

Complex predicates have been the focus of many studies in the generative tradition, including various forms of Chomskyan theories as well as theories without derivational syntax. Many approaches to complex predicates assume that they are derived by lexical processes, such as Grimshaw and Mester 1988, Barjasteh 1979. One of the reasons for choosing a lexical approach is that even very productive combinations are not as completely productive as rules of syntactic combination might predict. My account of Hindi/Urdu, therefore, must try to explain limits on possible combinations as well as differences across languages which show similarities of complex predicates, such as Persian, Turkish and Japanese. Lexical accounts allow much complex information to be stated, but this approach does not encourage a search for generalizations and explanations both within a language and across languages. In particular, lexical accounts do not explain why many languages have complex predicates with similar (though not identical) properties, and why languages differ in how complex predicates are realized morphologically. I will not discount the role of the lexicon; the properties of the N -V combinations will be distributed among the lexical entries for N and for V.

In this paper, I will take a view which extends the ideas of Jackendoff 1976, Cattell 1984, Jayaseelan 1988, Sproat 1986), Doostan-Karimi 2000, Higginbotham 1985, 1999, 2000 and Folli 2001, in which the argument structures of two heads are identified or merge into one argument structure. This approach has some features in common with the argument fusion proposed in Lexical Functional analyses (Alsina 1993, Butt 1995, Karimi 1994).

An alternative to this analysis is a process of incorporation taking place pre-syntactically (Hale and Keyser 1993), in the syntax (Baker 1988, 1995) or at Logical Form (Saito and Hoshi 1994, Karimi 1994) Dubinsky 1997. . LF incorporation is proposed for indefinite NPs by Dayal 1999, 2003, Massam 2001, Farkas and de Swart 2003, Chung and Ladusaw 2003. I will distinguish the incorporation analysis from the argument merger analysis, and show that while Hindi/Urdu does have presyntactic incorporation and LF incorporation of indefinites, these mechanisms do not account for the formation of complex predicates and their properties.

In sum, I will argue for the following points:

- A. Constraints on argument licensing are met only at LF, in particular to allow a head with an argument structure to discharge an argument position of another head only if the (first) head has no open, undischarged argument positions.
- B. This condition forces identification of the argument structures of V and N or A, without

D. Argument identification is not identical to semantic incorporation, which involves restriction of an argument without saturation of the position. Restriction without identification would violate the LF condition A.

E. The LF condition which forces identification has universal application, suggesting that complex predicates are possible in any language, subject to language-specific conditions on case marking, and encoding of thematic rules and verbal aspect.

The LF interface condition

Chomsky 1986 proposes a LF condition, Full Interpretation, which in general requires that all argument roles be discharged, and all sentence constituents be licensed syntactically. A condition on syntactic derivation proposed by Higginbotham 1985 is a specific instance of Full Interpretation. Higginbotham's condition is part of a general set of processes for discharging semantic roles. The condition forbids a phrase to saturate a thematic position in an argument structure if the phrase itself is not completely saturated, and has open thematic positions not satisfied by some syntactic argument. If both N and V have argument structures, then N cannot be a thematic argument for V if it has any open argument positions. An alternative derivation is available, which merges the argument structures of N and V to form one composite argument structure.

I will adopt Higginbotham's 'theta merger', but I will focus on argument positions rather than specific theta roles and hierarchies of theta roles. I will, however, assume that the arguments of N and V are ordered with respect to one another, and that arguments are associated with events and sub-events, from which their semantic roles are derived (Grimshaw 1990, Pustejovsky 1995). This argument merger may have several different syntactic outcomes, as shown by case marking, agreement and interpretation. Different languages may favor one or more syntactic outcomes. The range of different complex predicates in Hindi/Urdu and what is reflected by the case and agreement morphology of this language reinforces some of the central properties of these combinations.

Overview of complex predicates in Hindi/Urdu

There are three types of V + N complex predicate in Hindi/Urdu. These combinations show three solutions to the question of how to combine N with its thematic object and to license its case. The most common solution is to mark the thematic object with a genitive marker (3).

3) <i>maiN-nee/*mujhee</i>	-[[un- kii	yaad]	kii]	'I remembered/recalled them'	
I-erg	/*dat	3pl-gen-fs	memory-fs	do-pf-fs	[Type I]
<Exp>		<theme>			

The genitive postposition *kii*, the default method of linking a N to a possessor or argument (4):.

postposition as case marker.

The second solution is similar to the first except that a locative postposition *-see* appears on the thematic object instead of the structural genitive:

5) a. N *nafrat* 'hatred, aversion' + V *kar-naa* 'do-inf' =TRANSITIVE VERB 'to hate'
<exp, theme>

b. *woo apnee aap-see nafrat kar-taa hai.*
3s self's self-from hatred do-impf is

'He hates himself.' (Bahl 1974: 19) [Type II]

There is a third possibility. The combination of *yaad* 'memory' + *kar-naa* 'do-inf' 'remember' has one other syntactic realization which is not possible with its synonym *smaraN kar-naa* 'do memory, remember.' In (7), the combination of N - V forms a syntactic unit which can directly license object case on its thematic object:

6) [*suman tum-nee meeree saath joo upkaar ki-yaa hai*]
benevolent you-erg my towards which help do-pf is

un-koo maiN sadaa yaad kar-taa rah-uuNgaa
3pl-dat I always memory do-impf stay-fut-1s [Type III]

'I will always keep on remembering the kind help which you did for me (Bahl 1974:73)

The direct object in (6) has the dative postposition used for animate/specific objects, rather than the default genitive as (3) or the locative as in (5). This option is available for only a few N-V combinations in Hindi/Urdu, the ones which otherwise have genitive objects.. All three of these options must fall out from the account to be proposed below.

A number of verbs participate in N + V combinations as well as *kar-naa* 'do-inf.', varying considerably depending on N. For example, the N *yaad* 'memory' combines with 6 other verbs according to McGregor 1995. The meanings of the combinations are transparent and not unpredictably idiomatic:

7) *yaad* 'memory' +

a. *aa-naa* 'come-inf.' = 'remember, come to mind' [Dative subject]

b. *paR-naa* 'fall-inf.' = 'remember, come to mind' [Dative subject]

c. *dilaa-naa* 'cause to give/cause to get-inf' = 'remind, remember NP to another'

d. *rakh-naa* 'place-inf.' = 'bear in mind'

e. *rah-naa* 'stay-inf' = 'remain in mind, not be forgotten'

as main verbs.

- 9) maiN-nee kyaa ki-yaa
I erg what do-pf ‘What did I do?’

The syntactic and morphological properties of V persist, with few exceptions, in the N + V combinations. N contributes case marking for the thematic object in Types I and II, as well as its argument structure in all three types of N + V.

Syntactic and semantic goals of this account.

Here I sum up the defining features of complex predicates which are demonstrated in the examples above (see also Baker 1996: 353ff. for related discussion).

An account of N-V complex predicates in Hindi/Urdu has to deal with these semantic, syntactic and morphological facts, summarized in (10):

10) Properties of N + V predicates forming a syntactic unit with a single argument structure:

- a. The N acts as a syntactic argument of V, triggering object agreement (3)
- b. The theme of the composite argument structure forms a unit with N, as its thematic object, linked by genitive (4) or locative case (5b). The specific locative case is selected by N.
- c. The case of the external argument of the sentence is determined by V. It is nominative (5b) or ergative (3), or dative (in examples to follow).
- d. The external argument role of the N carries over to the N+V combination; the subject has the role of experiencer, following the meaning of N, rather than agent associated with the subject of V (3), (5b). In effect, the subject of N is the same as the subject of V, requiring that the subject of N is not separately projected.
- e. In a small class of cases, the thematic object of some N-V combinations may have normal direct object case (6), blocking V-agreement with N.

Literal vs complex predicate interpretation

The complex predicates discussed here form a unit with contributions from both N and V. The N has an external argument, but it is identified with the external argument of V, which takes on the theta role specified by N.² This semantic feature is one of the primary reasons for assuming that the argument structures of N and V are identified. In addition, the object of the N + V combination is the N internal argument not the N itself, though N controls agreement. A literal interpretation treats the N as the syntactic and semantic object of V, and the subject of N (if any) is distinct from the subject of V³.

I leave open the possibility that some combinations of N and V do have a literal interpretation,

interest here, the N is an event nominal rather than a result nominal, following Grimshaw's 1990 distinction. But I will assume also that the V has verbal characteristics, including case assigning properties, aspectual properties and a thematic structure, on the assumption to be justified further below that the Vs participating in N-V combinations are not special, stipulated 'light' verbs with no thematic or semantic properties. The variety of Vs in (7) shows some degree of distinct semantic contributions to the N-V combinations.

The interface condition on syntactic derivation.

Below I will outline three possible syntactic derivations found in Hindi/urdu, each based on the MERGE operation according to assumptions of the Minimalist Program, general principles of case and agreement licensing, and lexical information associated with N and V. The MERGE operation will also be subject to interface conditions, which determine well-formed or ill-formed representations in the phonological and semantic/interpretative interfaces. The crucial conditions have to do with the LF representation, as part of the interpretative interface. They may be subsumed under general rubric of Full Interpretation (Chomsky 1986, 1995) or Semantic Completeness (cf. Chung and Ladusaw 2003).

The prohibition against the presence of unchecked uninterpretable features in Phonetic Form is much discussed in the syntactic literature as an Interface Condition (Chomsky 1995, Ura 2000, etc). LF interface conditions have received less explicit focus as such, but many are familiar. For example, referentially dependent expressions such as anaphors must be identified by c-commanding antecedents in the right local domain. For example, an LF condition requires that anaphors in Hindi/Urdu must be locally identified by a subject antecedent.⁵ Other interface conditions constrain the distribution of wh- phrases and empty categories which they identify.

The Interface Condition which I will use in the explanation of complex predicates is based on a semantic condition on argument saturation proposed by Higginbotham 1985, reinterpreted as a Interface Condition on argument/event structures at LF.

10) Higginbotham's Principle,

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

Following Higginbotham's conception of theta operations including theta discharge, I will assume that if a theta role is discharged in an argument structure, then a syntactic constituent is identified with the theta role (see discussion in Speas 1990). For expository purposes, I will use subscripts to show which phrases are linked with which argument positions, even though LF representations are meant to be free of diacritic devices such as indices

This principle applies to complex predicate constructions. The N has open theta positions at

creates a new composite argument structure (Some other kinds of examples are discussed in Folli 2001).

Higginbotham's (1985) theta role-related operations, discussed in somewhat different form in Speas 1990, are summarized in (14):

11) a. *Theta discharge*- a theta role from the argument structure $\langle x,y \rangle$ is linked (by coindexing) with a phrase in the sentential structure. I assume this association is a consequence of MERGE and PROJECT (Chomsky 1995). Discharged theta roles are represented as 2^* and the position of the 2^* argument is *saturated*.

b. *Theta projection*: the argument structure is projected upward to the dominating node.

c. *Theta identification* identifies one argument position in an argument structure with an argument position in another argument structure; *theta merger* creates a composite argument structure by linking or identifying argument positions from separate argument structures.

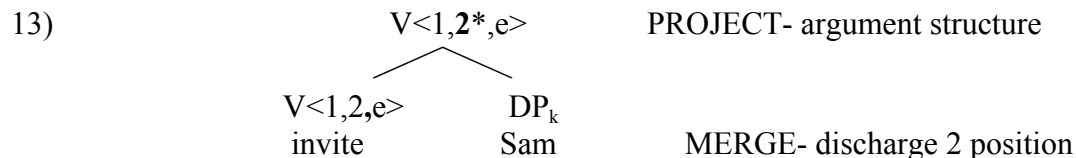
An example of a well-formed derivation

I begin by demonstrating how the operations of MERGE and Theta Discharge apply in a sample derivation of an English sentence, in which the arguments are saturated constituents. A sentence like (11a) is built from the merger of the transitive verb *invite* (11b) with its arguments. In addition to internal and external arguments, the argument structure will include the Davidsonian event argument *e*.

12) a. *A friend invited Sam*

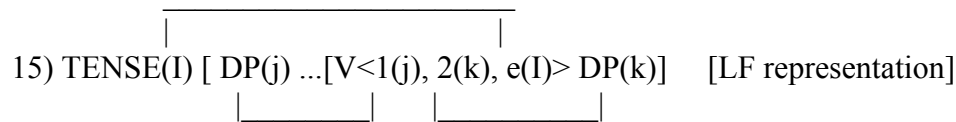
b. *invite* $\langle 1,2,e \rangle$ MERGES with *Sam*, discharging the 2 argument position

The verb *invite* merges with the DP *Sam*, and the V head PROJECTS upward, with its argument structure. The DP discharges one semantic role, saturating the internal argument position. Saturated positions are indicated by *, and we also assume a coindexing between this position and the DP argument. The result is (13):





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 (15):



Coindexing is represented by lines linking arguments and argument positions. 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.

Derivation of sentences with complex predicates in Hindi/Urdu

I will extend the syntactic derivation given above for a simplex predicate in English to the three complex predicate types in Hindi/Urdu, summarized as Types I-III in (16). They are distinguished by the case of the thematic object, and the agreement possibilities which follow from the choice of case.

16) Three types of complex predicate construction N + V

- a. Type I: [DP-**gen** N] + V = examples (3), (17)
- b. Type II [DP-**loc** N] + V = example (5b), (18)
- c. Type III DP-**nom/dat** [N+V] = examples (6), (19) -- N and V function as single constituent

17) Type I - The thematic object gets genitive case: *khoj kar-naa* 'search for, discover'

- a. ganapat singh-nee [*eek naai bimaarii-kii khooj*] **kii** hai kar-pf-f =kii
 Ganpat Singh-erg one new illness-**Gen** search(fs)-Nom do-pf-fs. is

'Ganpat Singh has discovered a new disease.' (Bahl 1974:222)

c. diidii-nee [[apnii aqal - kaa] isteemaal] ki-yaa [Type I]
 elder-sister-erg. self's wits-fs of use(ms) do-pf.ms

'Sister used her wits'. (V agrees with DP-kaa *isteemaal* 'use of DP') (Hook (1979:158))

18) Type II - The thematic object is locative *afsoos kar-naa* 'regret, repent' *afsoos* 'regret'

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)

19) Type III Structural nominative or dative case on thematic direct object N = *isteemaal* 'use'

bhaiyaa-nee [apnii taaqat] (*sahii) isteemaal kii [Type III]
 brother-erg. self's strength-nom-fs right use-ms do-pf-fs.

'Brother used his strength' (Hook (1979:158)) (V agrees with the thematic object; compare with (16b,c))

The N component of the complex predicate determines verb agreement, and in both types I and II, N can be modified by an adjective (17b). These properties are evidence for the phrase nature of the nominal-verbal combination. Type III does not allow adjectival modification; an adjective is impossible in (18) and in the type III equivalent of (17b) (Hook 1979). The N of Type III predicates never controls agreement on the verb. These properties argue for some kind of closer syntactic connection between N and V than is found in Types I and II.

The derivation of complex predicate sentences

I begin by deriving the sentence structures associated with Types I and II, which are similar except for the presence of lexical or genitive case on the thematic object. I will propose derivations for the structures of (20)-(21)

20) Type I - The thematic object gets genitive case: *khooj kar-naa* 'search for, discover'

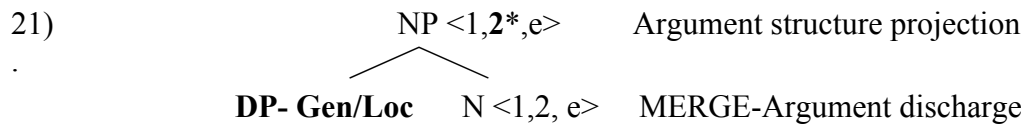
. ganapat singh-nee [*ek naii bimaarii-kii khooj*] kii hai
 Ganpat Singh-erg one new illness-Gen search(fs)-Nom do-pf-fs. is

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

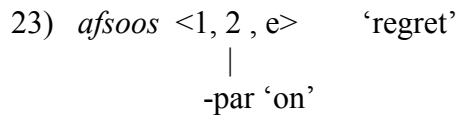
baad][*apni haalat* -par afsos kar-tii haiN.
 after self's condition-on regret do-impf are

'These wayward women, not at first, but after coming down from the intoxication of liquor, *feel regret about their condition.*' (Bahl 1974:98)

The N first merges with a DP, which gets the internal argument theta role. The genitive or locative case is checked within the nominal projection which is formed by MERGE. The theta grid with one discharged role is projected upward to the phrasal node (21):

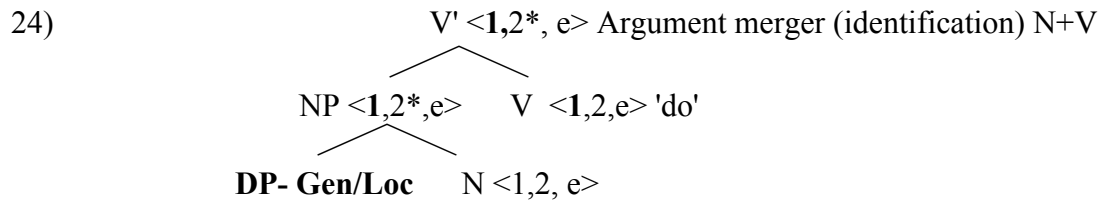


I will assume that the N *khoj* 'search' has no lexical case specified in its lexical entry, but *afsoos* 'regret' does (23):

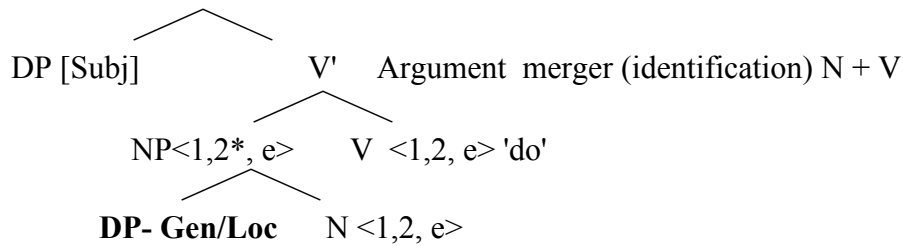


If no lexical case is supplied by the lexical entry of N, then structural genitive is the default.

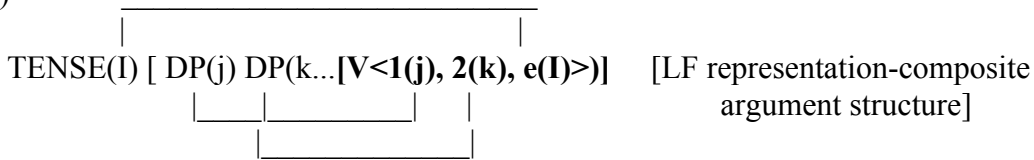
At this point the derivation can take two distinct paths, one of which is shown in (24). Here the NP with an open theta position undergoes theta identification with V, but does *not* discharge the internal theta role of V. The discharged 2 position of N is identified with the 2 open position of V. The open external argument position of N is identified with the open 1 position of V. The theta role specified by N is available in the composite N+V structure.



Exactly how the identification occurs will be discussed below, where the range of possible N and V combinations of argument structures are compared.

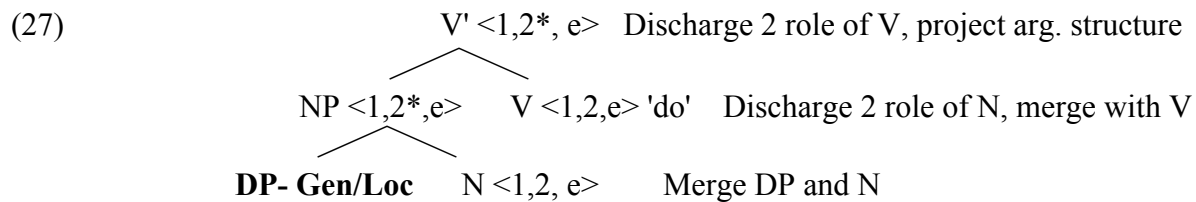


26)

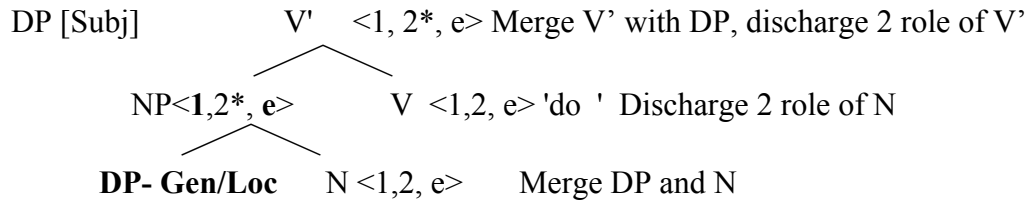


The main feature of this account is that NP itself does not discharge a theta role of V. It does not saturate a position in the argument structure of V. Higginbotham's condition would be violated if N with open theta positions in its argument structure did discharge a theta role. Rather, information 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. There are parallels with N incorporation, which I will explore below, concluding that N incorporation is distinct from theta identification. This process of identification is not strictly lexical; rather it is a lexical process which occurs in the sequences of MERGE and theta discharge. This process in effect reduces the total number of argument positions, so that the requirements of the composite predicate can be satisfied with the same number and kind of syntactic arguments as a simplex predicate.

Besides the sequence (22),(24)-(26), there is an alternative derivation, with N discharging an argument position, resulting in an ill-formed LF representation. In this derivation, the derivation proceeds as in (22), with merger of N and its thematic object. The difference is in the next steps. In (26), the NP merges with V, discharging the internal argument theta role of V:

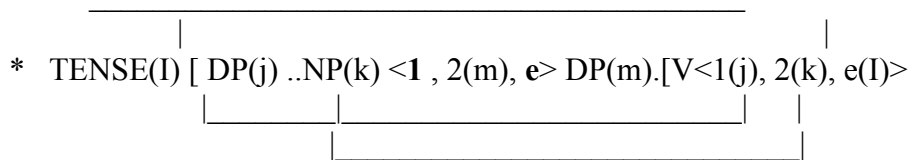


The structure then merges with the subject DP, discharging the external argument position of V. (27). Note that the argument structure of N remains independent of the argument structure of V.



The result is an ill-formed LF representation (29), with distinct N and V argument structures; an open argument positions. DP(j) is the external argument of V, NP(k) is the event nominal with an open external argument position. A crucial feature is that there are two separate and independent external argument positions, and this position in the argument structure of N is still open. It is 'left behind', as it is not transferred to a composite argument structure of N+V. The argument structures remain distinct, and there are two event arguments as well.

29) Ill-formed LF representation:



Here the NP counts as the internal object of V, projecting V's theta grid. The NP has an open theta position for the external argument. Its event argument is not bound by TENSE. Arguably, the sentence TENSE cannot bind into NP with a separate argument structure from V. 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.

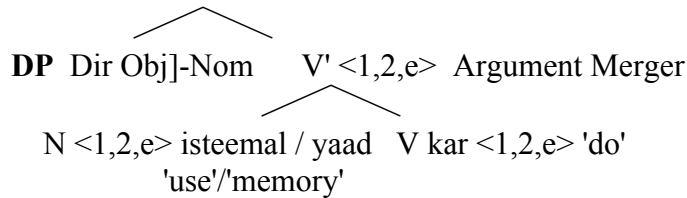
Type III complex predicates, which are all variations of Type I predicates, are derived in the same way as above, with two important differences: the N does not trigger verb agreement, and the thematic object has the nominative or dative case of an ordinary direct object.. This structure is illustrated by (30).

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

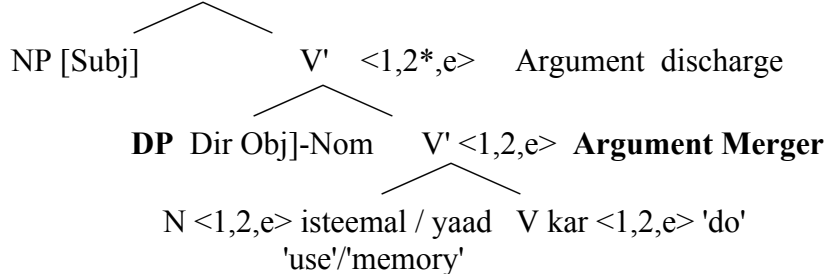
The difference is that Theta Merger is the first theta operation, taking place before the DP thematic object is merged (31)-(32). Finally, the subject is merged, as in the other derivations (33):

'use'/'memory'

32) V' <1,2*,e> **Argument discharge** Type III



33) VP <1*,2*, e> **Argument discharge** Type III



Argument identification takes place at the Xo level, before N has combined with a thematic DP object, as it would in solutions I and II. I leave open the question of whether identification takes place in the lexicon or early in the syntactic derivation⁶. This is an exceptional option in Hindi/Urdu, restricted to certain Ns. (Appendix). It is more the norm, however, in Persian and Turkish (see ___ below).

Early or late argument identification?

I have been assuming that for both types I and II, that theta discharge takes place 'early' in derivation, when DP merges with N. In both cases, N is able to case-mark the DP thematic object. The locative postposition in Type II combinations is an example of lexical, theta-related case, which is assumed to be licensed at theta discharge (Chomsky 1995, Ura 2000). Theta discharge of the internal argument of N would then take place before Theta Identification, as in Type I. This course of the derivation would distinguish Types I and II from Type III.

Type III combinations resemble Adj-V complex predicates (McGregor 1995), in that adjectives cannot assign genitive or locative case.⁷ Adjectives also do not trigger verb agreement. Type III Ns (Appendix) include some Perso-Arabic borrowing which are used both as N and Adj. I extend this property to all Type III Ns. They are underspecified for category, as +N alone. (Dubinsky 1997). The ability to assign genitive structural case is associated with the feature -V. There will be dual lexical entries for fully specified N *yaad* 'memory' +N -V, which enters

There is some counter-evidence to the claim that locative case is assigned in the same way as genitive case on the thematic object. 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:

34) 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, 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, this course of derivation would violate the Higginbotham condition on LF, while Jayaseelan was concerned about violations of the Projection Principle at earlier levels than LF.⁸ I will continue to assume that theta identification is not postponed, and instead note an unresolved problem: N cannot assign a locative postposition to its thematic object in the absence of a V host.

The contrast between the ‘late’ merger in types I,II and the ‘early’ merger in Type III also recalls the formation of Arabic *masdars*, or verbal nouns (Fassi Fehri 1995: 232ff). 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. The object may also be marked with a preposition, like the object of a noun. 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. In both derivations, 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. The *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. The syntactic derivation in both cases uses the same mechanism of theta identification, which may take place early or late.⁹ These cases offer support for the claim that theta identification is a universally available possibility.¹⁰

Identification of argument structure.

Above I have outlined the general course of syntactic derivation which I propose for the formation of complex predicates in Hindi/Urdu, which involves argument identification of two argument structures in the first or second steps in the derivation.. In this section I want to give a more complete picture of what identification does in specific cases, especially given differences in the argument structures of N and V. I will focus on the argument positions of N and V, assuming an ordering from innermost to the external arguments. In a later section I will give a more-fine grained account based on event structure and sub-events, assuming that ‘theta’ roles are not syntactic primitives associated with N or V, but rather are defined with respect to the sub-events (Grimshaw 1990, Pustejosky 1995).

- 35) [nominative or dative subjects]¹¹
- a. aa-naa 'come (to), become, happen'
 - b.. hoo-naa 'be, become'
 - c, lag-naa 'be attached to, strike'
 - d. cal-naa 'go, move, start, advance'
 - e. pahuNc-naa 'arrive, reach'

- 36) Transitive [ergative subjects]
- a. kar-naa 'do, make'
 - 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. khaa-naa 'eat' (in combinations with the meaning 'suffer, endure, go through N')
 - f. lag-aa-naa 'cause to strike, put'
 - g.. uThaa-naa 'raise'
 - h. pahuNc-aa-naa 'cause to arrive'
 - i. paa-naa 'find'
 - j. maar-naa 'beat'.
 - k. dee-naa 'give'

The complex predicates formed from these verbs preserve the cases which are associated with them as main verbs. The verbs in (36) preserve the nominative/ergative subject marking of the V component.¹² The type III predicates with a V from this category have dative direct objects, if the thematic direct object is animate or specific. The complex predicates formed with verbs in (35) never have ergative subjects. The subject is either nominative or dative, and the thematic object has genitive case. In Type III predicates, only nominative case is possible.

Diatheses and complex predicates

The verbs entering into complex predicates often relate complex predicates in the same way that simplex verbs are related in diatheses.. In Hindi/Urdu, many transitive verbs have derived detransitive counterparts, as well as causative verbs formed from the reduced stem of the detransitive. Complex predicates represent a very productive source of new predicates, as the older denominal verb formation became less productive (Gambhir 1994). In fact, as Gambhir (1993) notes, complex predicates can be formed to fill out paradigms of derived verbs:

- | | | | |
|--------------------------|-----------------------|---|---|
| 37) a. detransitive | b. transitive | c. causative 1 | d. causative 2 |
| khul-naa
“to be open” | khol-naa
“to open“ | khul-aa-naa
“to cause to be
open” | khul-vaa-naa
“to get opened
by another” |
| phir-naa | pher-naa | phir-aa-naa | phir-vaa-naa |

38)	intransitive	transitive	causative
a..	talaash hoo-naa 'be searched'	talaash kar-naa 'search for'	talaash kar-vaa-naa 'have someone search for'
b.	* <i>talash-naa</i> 'be searched'	talaash-naa 'search for'	* <i>talash-vaa-naab.</i> 'have someone search'

(Gambhir 1993:89)

The intransitive is formed from N+ *hoo-naa* 'be, become', and the causative is made up of N and *kar-vaa-naa* 'to cause to do', one of the causative forms of *kar-naa* 'do'. In this way, the complex predicates of Hindi/Urdu mirror the relations among simplex verbs in (37), (38a).

The argument structures of N and V do not match as perfectly as examples with transitive verbs such as *kar-naa* 'do'. I assume that all of the instances of *talaash* 'search' in (38)a., the argument structure is as in (39)a:

- 39) a. *talaash* 'search' <1, 2, e> 1: agent 2: theme + *kar-naa* 'do' <1,2>, Agent, theme
 b. *talaash* 'search' <1, 2, e> 1: agent 2: theme + *hoo-naa* 'become' <1> theme
 c. *talaash* 'search' <1, 2, e> 1: agent 2: theme + *kar-vaa-naa* 'cause X to do/make Y
 <1, 2, 3, e> 1: Agent 2: sub-agent 3 theme

The argument structures match in (39a) There are two mismatches: in (39b,c), the N has one more argument than V in (39b), while in (39c), the V has one more argument than N. In the first instance, the theme argument of N is identified with the theme argument of V, and the external argument of N is not syntactically projected (40b). In the second case, another argument is syntactically projected in addition to the agent and theme of *talaash* 'search'. The argument structure of N is identified with the two internal arguments of *kar-vaa-naa* 'cause to do', but the causative verb itself supplies the outer agent and licenses ergative case for it (40c).

- 40) a. *pulis-ne sargarmii-see apraadhi-kii talaash kii*
 police-erg actively criminal -gen search do-pf

'The police searched actively for the criminal.' (Gambhir 1993: 92)

- b. *apraadhi-kii talaash nahiiN huii*
 criminal-gen search not be-pf

'The criminal was not searched for.'

- c. *sarkaar -nee pulis-see apraadhi-kii talaash kar-vaa-ii*
 government-erg police-from criminal-gen search do-cause-pf

Combinations of argument structures

In order to demonstrate the process of argument identification, I focus on the syntactic argument structures of N and V, leaving aside their event structures until the next section. The syntactic arguments are ordered, distinguishing internal arguments from the external argument (Dowty 1984). The principle which governs a well-formed combination matches arguments structures beginning with the innermost argument, then successively higher arguments without skipping any argument or setting up ‘crossed’ associations.

I will also assume that the N which enters into a complex construction is an event nominal with an argument structure like a verb. But unlike V, the event N has no structural case assigning properties except for genitive; it may also have lexical case associated with a specific argument. These properties mean that it can license a syntactic object, but not a subject, independently of the N + V combination.

In these examples, I note three different combinations of N and V (41)-(43). In these examples, the argument structure of N has been kept constant, while the V varies in the number of arguments it has. I include an event noun with an agentive external argument role (*talaash* ‘search’), and another event N with the experiencer role for its external argument (*yaad* ‘memory’).

If the argument structures of N and V have the same number of arguments, then the arguments are identified according to hierarchical ranking

41) a. *talaash kar-naa* ‘search for’

$$\begin{array}{c} N \langle 1, 2, e \rangle \\ | \quad | \quad | \\ V \langle 1, 2, e \rangle \end{array}$$

b. *yaad kar-naa* ‘remember, recall’

$$\begin{array}{c} N \langle 1, 2, e \rangle \\ | \quad | \quad | \\ V \langle 1, 2, e \rangle \end{array}$$

c. composite N + V

$$\langle 1, 2^*, e \rangle$$

In (41a,b) the argument structures of N and V correspond rather closely. In (41b), the N assigns an experiencer role to the external argument, in italics, which is projected into the combination of N+V. The internal arguments of N and V have similar theme theta roles. The internal argument position of N is more fully specified than the corresponding position of V, because the 2 position of N is marked as saturated and indexed to a particular DP. This information is included in the composite N+V argument structure.

If the V has fewer argument positions than N (42), then the corresponding arguments are

b. *yaad hoo-naa* ‘remember, have in mind

$$\begin{array}{c} N < \mathbf{I}, 2, e > \\ | \quad | \\ V < 2, e > \end{array}$$

c. *yaad aa-naa* ‘remember, recall, miss’

$$\begin{array}{c} N < \mathbf{1}, 2, e > \\ | \quad | \quad | \\ V < 2, e > \end{array}$$

The intransitive V in (42) has no structural case licensing properties for this extra argument. It can be projected syntactically with lexical case, either the ‘instrumental’ *-see* ‘with’ for the agent in (42a), or the dative *-koo* for the experiencer in (42b,c).¹³

V may have has one more argument position than N (43):

43) a. *talaash karvaa-aa* ‘have X search for Y’

$$\begin{array}{c} N < \mathbf{1}, 2, e > \\ | \quad | \quad | \\ V < \mathbf{1}, 2, 3, e > \end{array}$$

b. *yaad dilaa-naa* ‘remind X of Y’

$$\begin{array}{c} N < \mathbf{I}, 2, e > \\ | \quad | \quad | \\ V < \mathbf{1}, 2, 3, e > \end{array}$$

The arguments of N and V are identified according hierarchical ranking beginning with the lowest. In the resulting complex predicate, the causative verb adds its external argument to the linked arguments of N and the inner arguments of V. This argument is merged with the composite argument structure formed by identification. The syntactic and case licensing properties of V support the syntactic projection of the merged argument, and license ergative case for it.

All of the combinations involve two argument structures which are compatible; the arguments can be matched, according to the ordering which distinguishes internal and external arguments. The identification begins with the innermost arguments and proceeds upwards. Arguments at the upper edge may be omitted or added. There are no crossed relations or lower arguments which are skipped.¹⁴

Some problems arise if N is transitive, but V is ditransitive. If the external arguments match, then either the 2 argument of V is skipped (44a), or the N itself saturates the 2 position of V (44b), violating the Interface condition on N arguments with unsaturated positions.

44) a. N <1, 2, e> b. N <1, 2, e>

the V *dee-naa* ‘give’.¹⁵ Interestingly, there are few complex predicates in HU formed with *dee-naa*, unlike English *give*.¹⁶

The verb *dee-naa* ‘give’ appears to be reanalyzed in combinations with *dikhaaii* ‘sight’, *sunaii* ‘hearing.’ The verb *dee-naa* normally takes an ergative subject in perfective sentences, but in these combinations it does not, and the only possible subject case is dative, and the theme controls agreement:

45) a. un-koo naaNw dikhaaii nahaiiN dii
3pl-dat boat-f sight-f not give-pf.

‘They didn’t see the ship, the ship was not visible to them.’

b. acaanak un-koo yah taaraa dikhaaii di-yaa ‘
suddenly 3pl-dat this star-m sight-f give-pf

‘Suddenly they saw, caught a glimpse of this star.’

If we were to look for a lexically reanalyzed ‘light’ verb in Hindi/Urdu, this is clearly one (others include *paa-naa* ‘find’ which loses ergative case as the auxiliary ‘manage to VP’). On the other hand, the combination NP-koo dooS *dee-naa* ‘to give blame to NP (for X) matches ditransitive *dee-naa* ‘give’ with ditransitive *dooS* ‘blame’. Unless the N and V are both ditransitive, the ‘give’ type of verb seems to require some adjustments. Another ‘light’ version of *dee-naa* retains ergative subject case but loses the dative indirect object (Karimi 1997, Karimi-Doostan 1997 make similar observations about ‘give’ in Persian complex predicates):

46) NP-par dhyaan *dee-naa* ‘to give thought to, reflect on’

maiN-nee udhar dhyaan nahiiN di-yaa
I-erg there thought not give-pf

‘I did not pay attention there.’ (Bahri 1992: 332)

Many complex predicates in HU are formed with *dilaa-naa* ‘cause to give, cause to possess’, which in morphological form is the causative of *dee-naa* ‘give’. In combination with N, it functions as the causative of transitive complex predicates:

47) a. yaad hoo-naa, yaad, aa-naa ‘remember’ b. yaad dilaa-naa ‘remind’
c. bharoosaa kar-naa ‘rely on, have confidence in’ d. bharoosaa dilaa-naa ‘encourage’¹⁷

These combinations (47b,d) link the N arguments to the internal arguments of V, conforming to the conditions on well-formed identification.

FORMAL: remember (e2, [1], [2])
AGENTIVE: = experiencer-act (e1, [1])

Nominals of this sort have both result and process interpretations, depending on whether e2 or e1 is the head of the whole event (49a). The N does not have the syntactic means to license two arguments with structural case, so the external argument is represented as a default argument (49b) which can be projected with a lexical, postpositional case.¹⁹

Verbs are specified for the head of the event. Let us consider *kar-naa* ‘do/make’ to be a generic verb of causation projects the initiating event e1 (50) (cf. Pustejovsky 1995: 187):

50) *kar-naa* ‘do, make’

- a. Event Structure: E1 = e1: process E2 = e2: state
[Ordering] restriction < " [event 1 precedes and overlaps event2] **Head = e1**
- b. Argument structure = Arg 1 = [1] , Arg 2 = [2]
- c. Qualia structure: Default causative lexical-conceptual paradigm.
FORMAL: do_result (e2, [2])
AGENTIVE: do_act (e1, [1], [2])

The process of theta identification I have proposed gets realized in this notation as the unification of event, argument and qualia structure^{4s}. The more fully specified values of N or V fill out the more general or less specified values of the other (51).

51) *yaad kar-naa* ‘remember’

- a. Event Structure: E1 = e1: process E2 = e2: state
[Ordering] restriction < " [event 1 precedes and overlaps event2] **Head = e1**
- b. Argument structure = Arg 1 = [1] animate being , Arg 2 = [2] proposition
- b. Qualia structure: Experiencer lexical-conceptual paradigm.
FORMAL: remember (e2 [1], [2])
AGENTIVE: exp-act (e1, [1], [2])

The N specifies two subevents, but not which one is the head. The V supplies the head of the event structure, so that the arguments associated with this event must be projected syntactically, including the external argument which is only a default argument of N. The qualia of N supplies selectional specification for the arguments.

a, Event structure: E1 = e1: process E2 = e2: state Ordering: e1 < " e2
head = e2

b. Argument structure: Default arg1 = [1] : location Arg1 = [2]

c. Qualia = Formal = **to** (e2, [1] [2]) Agentive: come-act (e1, [2])

Combined with *yaad* ‘memory’ it yields (53), with a change/resulting state interpretation:

53) *yaad aa-naa* ‘remember’

a) Event Structure: E1 = e1: process E2 = e2: state Ordering: e1 < " et2]
head = e2

b) Argument structure = Default Arg 1 = [1] animate being , Arg 1 = [2] proposition

c) Qualia structure: Experiencer lexical-conceptual paradigm.
FORMAL: remember (e2 [1], [2])
AGENTIVE: exp-act (e1, [1], [2])

The two complex predicates *yaad kar-naa* ‘remember, recall’ (51) and *yaad aa-naa* ‘remember, recall, miss’ (53) have similar qualia structures, but differ in range of meaning. The explanation derived from this representation lies in the event structure. The combination which has the e1 or process event as head may have an agentive interpretation, but need not (3), (6)). The combination with e2 as the head allows only a non-agentive interpretation. The V quale specification ‘to’ is realized as dative case on argument [1] in the combinations (53).

These examples allow us to make some generalizations about the kind of verb which may participate in complex predicate combinations with N. The verbs in (35)-(36) have several properties in common, which distinguish them from the vast majority of verbs which do not participate in N + V complex predicates. First, their qualia structure lacks very specific selectional properties for their arguments. Second, they are verbs with a telic interpretation, as accomplishments or achievements. Their event structure involves a change of some sort, leading to a resulting a state of affairs. A verb is unlikely to form a complex predicate with an event N if it refers to an activity specified for manner or direction of motion (eg. *naac-naa* ‘dance’ *uchal-naa* ‘jump up, leap’) or an inherent state (eg. *jaan-naa* ‘know’).

The Ns which participate in these combinations all allow an event interpretation, and their qualia structures contain more specific information than the corresponding properties of V. These verbs need not be lexically derived ‘light’ verbs of the kind proposed in Grimshaw and Mester 1988, but rather verbs which are inherently ‘lighter’ in qualia specification than N. Rosen 1990 and Butt 1995 propose that ‘light’ verbs may have different degrees of semantic

The combinations of N + V in this language function as predicational units. This unitary nature is explained in this analysis as the result of argument identification between N and V in the course of syntactic merger in the derivation of a sentence. The information encoded in N is transferred to the whole combination without requiring that N discharge a theta role in the argument structure of V.

Incorporation of N into V might be seen as an alternative mechanism to argument identification. Incorporation is understood in several different ways. Baker 1988, 1996 argues for syntactic incorporation of direct objects in polysynthetic languages as a way of assigning case and discharging a theta role. The N is contained within the V. Hale and Keyser 1993 explain verb formations of various sorts by lexical/syntactic incorporation of nouns and other categories in the projection of 'light' verbs (conflation). A kind of semantic incorporation is proposed for non-polysynthetic languages, allowing the N and V components to retain syntactic and morphological independence. This process is also called 'Pseudo N incorporation' (Massam 2003).

As Hindi/Urdu is not a polysynthetic language, complex predicates in this language would have to be instances of semantic incorporation. This kind of operation applies in many languages to indefinite NP (not DP) objects, restricting the meaning of the V.

Mohanan 1994, 1995 provides a number of arguments that Hindi/Urdu has incorporation of indefinite, not necessarily referential objects. Dayal 2003 works out a theory of incorporation which retains the syntactic independence of the components, but which interprets the bare singular N object as a predicate restricting the meaning of the verb, rather than a referential phrase discharging a theta role. The category N is given a set interpretation of type $\langle e, t \rangle$, the kind of set-denoting expression shared by intransitive verbs and adjectives. The nominative object is not necessarily singular in interpretation:

54) . 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)

Various proposals are currently being discussed which combine a bare N with a verb (Dayal 2003, Farkas and deSwart 2001, Chung and Ladusaw 2003). Chung and Ladusaw 2003 treat the N predicate as a modifier of the predicate, which restricts the predicate without discharging/saturating an argument position. The combined structures for the V (55a) and the object N (55b) undergo the operation Restrict (55c), yielding (55d).²⁰

55) a. *samhaal-naa* 'look after' $\delta y \delta x \delta e$ [look after' (y)(x)(e)]
 b. *baccaa* 'child'
 c. Restrict ($\delta y \delta x \delta e$ [look after' (y)(x)(e), child')

N+V combination *khooj kar-naa* means ‘search, discover’.

58) ganpat singh haasTal-kii supeeerinTeeNDaNT-nee eek naai bimaarii-kii khooj kii hai
Ganpat Singh hostel-gen superintendent-erg one new disease-gen-search do-pf is

‘Ganpat Singh the superintendent of the hostel has discovered a new disease.’ Bahl 1974:222

This difference in range of meaning suggests that simplex verbs and complex predicates are not derived by the same process. If we accept conflation as the incorporation process which derives V from roots/nouns (51), then it is not the process for deriving complex predicates. Conflation creates a single lexical word; importing the argument structure of the root into an empty V without morphological properties;. For this reason conflation is not the process involved in Type III complex predicates. In these complex predicates, the V contributes aspectual and morphological information, which may distinguish complex predicates with similar meanings, such as *yaad kar-naa* ‘remember, recall’, *yaad aa-naa* ‘remember, miss, recall’.

The preferred account for complex predicates, then, is theta identification if both components have lexical content in addition to category information.

I conclude this section on incorporation with the view of lexical incorporation in Baker 1996, 1997. Baker considers causative formation as syntactic incorporation of a VP projection with an argument structure for V into a ‘light’ causative verb, forming a composite predicate, provided that the argument structures are non-distinct:

59) [Light Verb Causative]

If the argument structure of X is non-distinct from the argument structure of Y, then the head chain ($X_k + Y t_k$) may assign theta roles as a unit. (Baker 1996: 355)

Crucially, the head chain is derived by syntactic adjunction of X to Y, the kind of incorporation found in polysynthetic languages like Mohawk. Such languages are subject to a constraint relating agreement relations to argument structures, assuming that incorporation is a form of object agreement side by side with coindexing a phrase with agreement morphology.:

60) [Condition on Agreement Relations]

An agreement relation X can be coindexed with a phrase Y only if Y is coindexed with a position in the argument structure of Z and X is adjoined to Z (Baker 1996:344).

These conditions exclude most kind of causative formations from polysynthetic languages like Mohawk. Only unaccusative verbs combine with a causative suffix in these languages. As the VP argument structure is a subset of the matrix, the theme of the embedded unaccusative VP is non-distinct from the theme of (transitive) CAUSE. Transitive verbs and unergative verbs are excluded, because their argument structures are distinct from the argument structure of the

Baker's principles offer an elegant explanation for the properties of polysynthetic languages and what is allowed or systematically excluded in morphologically complex words. It is unclear, however, whether either (59) or (60) applies to Hindi/Urdu, whose properties are quite different from what is found in polysynthetic language. It allows causative formations for transitive and unergative verbs, and even double causatives of transitive as well as intransitive verbs:

- 61) a. [Unaccusative] pahuNc-naa 'reach, arrive' pahuNc-aa-naa 'cause to arrive. reach'
 b. [Unergative] naac-naa 'dance' nac-aa-naa 'cause to dance, harass, spin (a top)
 c. [Unergative] kheel-naa 'play' nac-vaa-naa to cause NP to be danced by NP
 khil-aa-naa 'cause to play, amuse'
 khil-vaa-naa 'cause NP to cause NP to play'
- 62)a. [Transitive] beec-naa 'sell' bik-aa-naa 'to have NP sell NP'
 bik-vaa-naa 'to have NP make NP sell NP'
- b. khaa-naa 'eat' khilaa-naa 'cause to eat, feed'
 khil-vaa-naa 'have NP feed NP'

Agreement in Hindi/Urdu does refer to nominative direct objects (if the subject is not nominative. But agreement is not confined to the local minimal clause: An option for agreement is triggered by the object of the embedded clause object and expressed on the matrix verbal complex as well as the embedded infinitive:²²

- 63) baccooN-koo [PRO saaikal calaa-nii] aa-tii hai
 children-mpl-dat bicycle-fs drive-inf-f come-impf-fs is-3s

'The children know [(how) to ride a bicycle].'

The theta identification procedure I propose here allows the N + V combination to inherit properties from both N and V. The combination is like a V, but the thematic object case is determined by N, and the external argument has a theta role determined by N. On many counts, Hindi/Urdu instantiates a different linguistic type from the polysynthetic languages, and it does not rule out argument composition.

Other accounts of complex predicates in Hindi/Urdu

Many of the important issues raised by complex predicates in this language have been discussed in previous work. Writers such as Kellogg 1938, Bailey 1956, Porizka 1963, Sharma 1975, Kachru 1968, 1980, 1982 Masica 1991 and McGregor 1995 have contributed many valuable insights and examples. Bahl used an early corpus study of books in Hindi to define the properties of N + *kar-naa* 'do' 1974, 1979. These volumes represent a valuable resource for studying complex predicates, independently of the theory employed. Transformational grammatical

devices used and the kind of explanation which emerges.

Mohanan 1994, 1997 makes the same tripartite distinction among complex predicates which I have adopted here, based on the case of the thematic object. She makes a further distinction based on the agreement of V with the event noun (my Types I and II), or with the thematic object (Type III, which also includes default agreement if the object is dative). Both the N and V components have argument structures, as in this account. She also gives evidence for some syntactic independence of N, particularly in the N + V combinations with genitive and locative object case.

Her account of the complex relation between N and its arguments and V and its arguments is based on the mapping between levels of representations, which in her Lexical Functional Grammar account include semantic structure SEM STR, argument structure ARG STR, grammatical function structure GF STR, the word string and grammatical category structure GC STR. The variations among complex predicates lie in mismatches or rather adjustments in the mapping between SEM STR and ARG STR, which have different consequences for case marking and agreement.

The relation between the N and V at SEM STR is the same for all the outcomes. The N and its arguments are embedded within the project of V and its arguments. The subject argument of N is linked with the subject of V, and this linking is supported by evidence that subject-oriented reflexives cannot refer to two distinct subjects (1994:218). The thematic role of the N-subject prevails. The N+V combinations which mark the thematic object as genitive or locative The subject of N and V is linked to the ARG1- SUBJ. The event N is linked to ARG2-OBJ, and the thematic argument is linked with ARG3 OBL (1994: 224,232). The object status of the event N accounts for why N agrees with the V when the subject is not nominative.

Type III complex predicates differ from the other types in how the composite SEM STR arguments are mapped into ARG STR arguments. There is a reanalysis or restructuring of the N arguments in composition with the V. The linked subjects are mapped into the ARG1-SUBJ, as in the previous cases. But the logical object (what I call the thematic object of N) is also the Iobject of V, and is mapped into ARG2-OBJ. The N itself is not an argument of V and is not mapped into an ARG-OBJ at ARG STR. In the terms of the present analysis, the argument structures of N and V are fused, reducing four possible arguments to two (by fusion of the subjects and making the N object into the object of the whole). This is the analogue of my early or lexically required theta identification. In Mohanan's analysis, Type I/II predicates have an 'extra' oblique argument, as N is an argument along with the subject and object arguments of V. The object of N is demoted to an OBL argument.

Mohanan's analysis is very comprehensive and based on extensive and well-chosen evidence. It derives the right surface results, but it depends on solutions to SEM/ARG mismatches which are not clearly motivated, especially with respect to the definition of the objects which gives the

Complex predicates in Persian, Turkish and Japanese

The explanation I have proposed for complex predicates is that if two heads with distinct argument structures are to combine by MERGE, the only well-formed way this could happen is by argument identification. This claim implies that in principle any language could form complex predicates if N and V have the right compatible lexical properties. There should be some syntactic and semantic commonalities among languages with large numbers of complex predicates.

It is beyond the scope of this paper to discuss every language with complex predicates (see Cattell 1988 for a comprehensive discussion of English). But this implication seems to be true for Turkish and Persian, in which very productively form complex predicates from the same Perso-Arabic and other borrowed words as Hindi/Urdu, and Japanese, which has head-final structure and overt case marking in some ways similar to Hindi/Urdu.

I will discuss these languages briefly, asking if (a) there are distinct ways of case-marking the thematic object, paralleling the three realizations in Hindi/Urdu, (b) the same Ns and Vs enter into combination as in Hindi/Urdu, and (c) argument identification and aspectual specification are the same across languages.

In Turkish and Persian, most complex predicates fall into what I have called Type II (with a non-genitive adposition) and Type III, with normal object marking on the thematic object. What is rare is genitive marking (Type I).

Turkish has complex predicates with accusative, dative and locative marking on the thematic objects (examples from Kornfilt 1997):

64) kristof kolomb amerika-yi keS et-ti
Christopher Columbus America-acc discovery do-past

‘Christopher Columbus discovered America.’ (1997:323)

64) hasan ben-den nafrat ed-er
Hasan I-from hatred do-aor

‘Hasan hates me.’ (1997:323)

66) hasan ban-a yardim et-li
Hasan I-dat help do-past

‘Hasan helped me.’ (1997: 323)

Persian shows complex predicates of all three types. Types II and III are the most common:

67) Type II PP object

javaad [*be maa*] [hoge xad]
Javad to us cheating struck 'Javad cheated us. Barjateh 1983: 257.

68) Type III Nom or -ra object

javaad ali-raa sarzaneš kard
Javad Ali-OM reproach did 'Javad reproached Ali.' Barjasteh 1983: 271

The N may take an infinitive object marked by the (*y*)*e* conjunction which marks genitives.

68) Type I genitive object of N (but N has object marker)

polis [az zendaaniyaan-e farari] [*taqaazaa-ye*-[*PRO bar-gasht-an*]*ra* kard
police from fugitive -of prisoners request of back-come-inf-OM do-past-3s

'The police requested the fugitive criminals [PRO to return.']. Barjasteh 1983:291

Like Turkish and Hindi/Urdu, Persian also has a Differential Object Marking for direct object. DOM is required for objects whose reference is specific or animate reference (Aissen 2003). Turkish has the marker (*y*)*i* for specific objects, while Persian uses *-ra* for specific and animate objects, with priority for specificity. Otherwise direct objects are unmarked. If the predicate selects a lexical (adpositional) case, this case takes priority and is the same for all kinds of DPs, whether specific or not. It appears that type II predicates are found in both Turkish and Persian.

This fact may explain the preference for Type III marking (0 or *-(yi/raa)* over type I genitive marking on the thematic object in Persian and Turkish, but not in Hindi/Urdu. The prevalence of the specific object marking option over genitive would seem to be an effect of the specificity requirement in Turkish and Persian. Hindi/Urdu has Differential Object Marking as well. The dative postposition *-ko* is preferred over the unmarked nomination for direct objects with specific and animate reference, but Aissen proposes that animacy is the stronger factor. Because specificity is not the stronger factor, it is possible to have specificity-neutral genitive marking in Hindi/Urdu.

Persian is very productive in complex predicate formation. Many authors have remarked that the complex predicates in this language outnumber the simplex predicates (Barjasteh 1987, Karimi 1997, Karimi-Doostan 1997, Karimi 1997). Many simplex verbs have fallen out of use, replaced by newly created complex predicates. A fairly large number of verbs may be used in

aavar ‘bring to memory remember’ (Barjesteh 1988). Though Persian is the source of Perso-Arabic words used in Urdu and Hindi, the complex predicate combinations seems not to have been directly borrowed into Hindi/Urdu. Persian and Hindi/Urdu both show productivity in creating complex predicates, independently of one another.

Complex predicates in Persian show many of the same syntactic properties as complex predicates in Hindi/Urdu. Karimi (1997) argues that even though N + V take a single stress as a unit, the N is not a thematic direct or indirect object of V. The N does not undergo scrambling like an ordinary indefinite object, and infinitives it cannot have the -e conjunction (*ezafe*) which has the function (among others) of marking a genitive relation. Ordinary direct objects can have the *ezafe* construction. Indirect objects with a preposition *be* can be separated from ordinary N objects, but not the PP in PP +V.

Persian treats N+V as predicative with respect to object clitics. Persian has only nominative subjects, and person-number agreement between animate subjects and the verb. The verb may also have pronoun clitics referring to the direct object (Barjesteh 1988, Mahootian 1997). The object clitic may optionally be suffixed to the N part of a complex predicate.

69) a. koomak kard-am -esh
 help do-1s-3s ‘I helped him’. (Mahootian 1997:139)

b. koomak-esh kard-am
 help-3s do-1s ‘I helped him’. (Mahootian 1997:139)

These criteria support the claim that N does not discharge a theta role of V, which follows from the Higginbotham 1985 condition. Interestingly, these arguments for Persian are mostly language specific, following from Persian coding of objects and freedom of scrambling.

Complex predicates in Japanese are made up of Sino-Japanese borrowings, originally verbs, which combine with a single V, *suru* ‘do’ to form a semantic unit. (Terada 1990, Matsumoto 1996, Dubinsky 1997). In addition there are aspectual and other verbs which combine with N, corresponding to raising and control verbs in other languages like Hindi/Urdu, with separate argument structures (Matsumoto 1996).

There are three productive patterns, shown in (70)-(72):

70) hanako-ga suugaku-o benkyoo (*o) sita
 Hanako-nom . math -acc study acc do-past

Hanako studied math.’ Dubinsky 1997: 646)

71) hanako-ga benkyoo -o sita

Hanako-nom . math -gen study acc do-past

Hanako studied math.’ Dubinsky 1997: 646)

Japanese has the three possibilities which correspond broadly to Types I-III in Hindi-Urdu, though a striking difference is that the object marker *-o* may mark either the N within N+V or the thematic object. The N in Japanese may have accusative marking (xxb, xxc) provided no thematic object also has accusative marking

Additional arguments have dative or other case:

73) taroo-wa ziroo-ni [yuusyooku-**no teian -o**] sita [Japanese]
Taro-top Ziro-dat dinner -gen proposal-acc did
'Taro proposed dinner to Ziro.' (Dubinsky 1997: 630)

There are restrictions on specific Ns about the case marking possibilities they may have, the differences seem to be correlated with aspectual differences among the complex predicates (Dubinsky 1997). The class of predicates like *benkyoo* 'study' allowing a genitive complement are all activities and accomplishments. The class of Ns like *hooyoo* 'embrace' requires a bare N, and *-o* marking on the thematic object (Dubinsky 1997: 646, 666). These predicates are achievements or states. Another class which rejects *-o* on the N can be used transitively or as unaccusative verbs, including *hunsitu* 'lose' (Dubinsky 1997: 644, 664).

74) a. ziroo-wa kon'yaku-yubiwa-o hunsitu (*o) sita
Ziro-top engagement-ring-acc lose (*acc) do-past

“Jiro lost the engagement-ring.” Dubinsky 1997:645)

b. kon'yaku-yubiwa-ga hunsitu (*o) sita
engagement-ring-nom lose acc do-past

‘The engagement ring got lost.’ (Dubinsky 1997:645)

c. kanseitoo- wa SOS-o jushin (*o) sita
control-tower-top SOS-acc reception did

‘The control tower received an SOS.’ (Terada 1990 109)

It appears that the N plays a much larger role in determining the verbal aspect of the combination in Japanese than in Hindi/Urdu or Persian. Japanese relies on one verb *suru* 'do in complex predicates. Persian, on the other hand, has a wide variety of verbs which differ in aspect and transitivity (see discussion in Karimi 1997; in Karimi-Doostan 1997 there is a

In principle, complex predicate formation is universally available, given Higginbotham's process of argument identification. It permits identification of two compatible argument structures. Identification is forced if some theta positions would be left open otherwise. The consequence I draw from the availability of these universal processes is that it is not *necessary* to create all complex predicates in the lexicon (or from components which have special lexicalized properties, such as the light verbs of Grimshaw and Mester 1988). Nevertheless, there are degrees of complex predicate formation, ranging from combinations which are lexically idiosyncratic to those where argument identification gives no different results from the literal meaning of syntactic V-object combination. For Persian, Barjesteh 1988 and Karimi 1997 note combinations which are not compositional and semantically transparent (*dast andaxt-an* 'hand throw' = 'to mock'). It would be expected, however, that the number of opaque combinations or combinations with special properties would be limited, and not outnumber the transparent combinations formed in the syntax, and deriving their properties from the properties of the component parts.

Summary and conclusions

As we have seen in this last section and the body of the paper, complex predicates are found in many languages, including Hindi/Urdu, the topic of this paper. In HU, complex predicates represent the most productive source of new predicates, superceding the older derivation pattern of denominal verbs. It is evident from the immediately preceding section that languages do not form identical complex predicates using exactly the same nouns and verbs. The morphological realizations are constrained by the overall case-marking patterns of the language, including the conditions for Differential Object Marking. Nevertheless, most of the combinations are compositional and transparently formed from the component N and V. The N portions are event nouns with transitive argument structures and other selectional specifications. The Vs show similarities as well. They express generic causation, change of state and change of (abstract) location, and compared with the Ns, they tend to have fewer and more general selectional properties. Butt 1995 notes that the verbs which combine with main verbs tends to be interpreted abstractly rather than concretely (in reference to actual or metaphorical space for example).

It is not possible to predict in advance which N + V combinations will be found in a given language.. The combinations seem to be 'idiomatic' in some sense, reflected for example by the listing of possible V combinations under the N entry in dictionaries of Hindi/Urdu, perhaps as a help for the non-native-speaker (eg. Bahri 1992, McGregor 1995). But the combinations are also quite compositional, with a small number of exceptions (in Persian for example). One should not confuse 'idiomaticity' with non-compositionality, as Nunberg et al 1994 point out. In fact complex predicates are even more compositional than phrasal idioms in English, as the Ns have independent uses as event nouns.

I have proposed in this paper that it is possible to give a syntactic account of complex predicates

This condition says that a phrase (such as the event N) which has even one open argument position may not discharge an argument position of another head, such as V. To do so would result in open argument positions at LF, a violation of Full Interpretation (Chomsky 1986, 1995). This result can be avoided during the course of syntactic derivation if (compatible) argument structures are identified, so that arguments of N are linked with argument positions of V, forming a single argument structure which may merge with other phrases, saturating the remaining argument positions.

This model of derivation can be applied to Hindi/Urdu Ns and Vs, with three different outcomes dictated by the case assignment properties of N and its categorial specification. One outcome merges N with one of its arguments, which receives structural genitive case. The second outcome depends on N being specified for a lexical case. In both instances N merges with an argument, and then undergoes identification with the argument structure of V. Identification takes place immediately for a small number of Ns which lack full N specification and hence cannot assign genitive case (cf. Dubinsky 1997). The result I have sought from this explanation of the derivation are very similar to what was derived by a LFG account by Mohanan 1994, 1997, though the explanation given is of a different nature.

Event Ns may undergo identification with a variety of verbs, with fewer argument positions or with more. In the first instance the external argument of N may optionally be projected with lexical case, in the latter, an argument is added by V. These patterns of argument structure allow complex predicates to mirror the simplex verbs in this language which are related in the transitive and causative diatheses.

Argument identification requires that the innermost arguments be matched, then successively matches higher argument without skipped positions or crossed relations, and without violating the Higginbotham condition. Ditransitive Vs may be fully matched with ditransitive nouns, or by matching the external argument of N with an internal argument of V in a 'causative' relation. Some ditransitive verbs such as *dee-naa* 'give' undergo reanalysis as plain transitives when they would otherwise not be compatible with N.

A closer look at argument structure may be found by examining the descriptions of N and V in terms of event structures proposed by Grimshaw 1990, Pustejovsky 1995. Arguments and selectional properties (qualia) are linked to subevents of a predicate. Sub-events are specified as states or processes, and the relation between subevents involves precedence or overlap, or both. The first subevent may be the head of the predicate, or the second, determining which arguments are obligatorily projected with structural case. The V component of a complex predicate is a verb of causation or change, specified for which sub event is the head. V takes on the argument specifications of N after argument identification, as well as the qualia, and imposes its own head specification, as the sub-events of N are not specified for head.. While I have not explored all possible combinations of N and V, it is clear that the specifications in Pustejovsky's system could be exploited to define compatible and incompatible combinations.

simplex verbs with the same meaning. Hindi/Urdu curiously does have simplex and complex predicates coexisting with similar but not identical meanings, differing in the possibility of a telic meaning. This fact raises questions not answered here about exactly how verbs and their event structures are specified for telicity.

The explanation offered here allows for the derivation of complex predicates in the syntax without the prior existence of special 'light' verbs without a specified argument structure, though they would retain event structure and case marking. It also offers some means of characterizing possible well-formed combinations of N and V.

References

- Bahl, K.C. (1974) *Studies in the semantic structure of Hindi*. Vol. I. Delhi: Motilal Banarsidass.
- Bahl, K.C. (1979) *Studies in the semantic structure of Hindi*, Vol. 2. Delhi: Manohar.
- Bahri, Hardev (1992) *Learner's Hindi Dictionary*. Delhi: Rajpal and Sons.
- Butt, Miriam (1995) *The structure of complex predicates in Urdu*. Stanford: CSLI Publications
- Butt, Miriam and Gillian Ramchand (2001) *Complex aspectual structure in Hindi-Urdu*. In B. Jensen et 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 (1986) *Knowledge of language*. New York: Praeger
- Chomsky, Noam. (1995) *The Minimalist Program*. MIT Press.
- Davison, Alice (2000) *Lexical pronouns and anaphors in Hindi/Urdu*. In B. Lust, K. Wali, J. Gair and K.V. Subbarao (eds.) *Lexical anaphors and pronouns in selected South Asian languages*. Berlin: Mouton deGruyter, 397-471
- Davison, Alice (to appear) *Structural case, lexical case and the verbal projection*. In Veneeta Dayal and Anoop Mahajan (eds) *Clause structures in South Asian languages*. Dordrecht: Kluwer Academic Publications.
- DiSciullo, Anna-Maria and Sarah Thomas Rosen (1990) *Light and semi-light verb construction*. In K. Dziwirek, P. Farrell and E. Mejias-Bikandi (eds) *Grammatical relations: a cross-theoretical perspective*. Stanford: CSLI, 109-127.
- Dubinsky, Stanley (1997) *Syntactic underspecification and light-verb phenomena* *Linguistics* 35, 627-677.
- Folli, Rafaella (2001) *Constructing telicity in English and Italian*. Wolfson College, Oxford University unpublished dissertation.
- Gambhir, Vijay (1993) *Complex verb phrase: a diachronic and synchronic view*. In M. Verma (ed) *The complex predicates in South Asian languages*. Delhi: Manohar Publications, 77-96.
- Grimshaw, Jane (1990) *Argument structure*. Cambridge: MIT Press.
- Grimshaw, Jane and Armin Mester (1988) *Light verbs and 2 marking*. *Linguistic Inquiry* 19,

- Higginbotham, James (1999) Accomplishments. Unpublished paper
- Higginbotham, James (2000) On events in linguistic semantics. In James Higginbotham, Fabio Pianesi and Achille Varzi (eds.) Speaking of events. Oxford University Press, 49-80.
- Hook, Peter (1973) The compound verb in Hindi. Ann Arbor: University of Michigan Center for South and Southeast Asian Studies.
- Hook, Peter (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 2 -theory. In. W. Wilkins (ed.) *Syntax and semantics* 21, Thematic relations. San Diego: Academic Press, 91-111.
- Kachru, Yamuna (1982) Conjunct verb sin Hindi and Persian. In P.J. Mistry (ed) *South Asia Review* 6.3.
- Karimi, Simin (1997) Persian complex verbs: idiomatic or compositional? Lexicology 3.1 PP 273-318 .
- Karimi-Doostan, Gholamhossein. (1997) Light verb constructions in Persian and Kurdish. University of Essex unpublished dissertation.
- Mahootian, Sharzad (1997) Persian. London: Routledge
- Matsumoto, Yo (1996) Complex predicates in Japanese. Stanford: CSLI
- McGregor, R.S. (1995) Oxford Hindi-English dictionary. Oxford: Oxford University Press.
- Mohanan, Tara (1994) Arguments in Hindi. Stanford: CSLI Publications.
- Mohanan, Tara (1995) Woodhood and lexicality: Noun incorporation in Hindi. *Natural Language and Linguistic Theory* 13.1, 75-134.
- Nespital, Hellmut (1997) Dictionary of Hindi verbs. Allahabad: Lokbharati Publishers.
- Nunberg, Geoffrey, Ivan Sag and Thomas Wasow (1994) Idioms. *Language* 70: 491-538.
- Pustejovsky, James (1995) The generative lexicon. MIT Press.
- Rosen, Carol Thomas (1990) Argument structure and complex predicates. New York: Garland.
- Speas, Margaret.(1990) Phrase structure and the lexicon. Dordrecht: Kluwer Academic Publications.
- Terada, Michiko(1990) Incorporation and argument structure in Japanese. University of Massachusetts dissertation.
- Verma, Manindra K (1993) Complex predicates and light verb in Hindi. .In M. Verma (ed) *The complex predicates in South Asian languages*. Delhi: Manohar Publications, 197-215..

Appendix

- 1) The Ns which allow Type III include the following .
 - 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'

- 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'
- 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'
- t. talaash (f) search

NOTES

1. The N *yaad* combines with a genitive marked theme 'memory about N'. The ability to remember is expressed with a different word, *yaaddaasht*.
2. Another criterion for the non-literal interpretation is due to Dubinsky (1997: 647). In Japanese a N + V compound can refer to multiple events as well as a single event if the thematic object is quantified.

i) Tanaka-wa san-nin -to aiseki-o sita
 Tanaka-top 3-person with table-sharing did

- a) 'Tanaka shared a table on one occasion with 3 people.'
- b) 'Tanaka (over the course of several occasions) shared a table with three people.'

Complex predicates in Hindi/Urdu of all three types have this range of meanings:

ii) kal maiN-nee tiin ciizooN-kaa anubhav ki-yaa
 yesterday I-erg 3 things-gen experience do-pf

'Yesterday, I experienced 3 things (a) simultaneously (b) separately over the course of time.'

3. What I am calling the literal interpretation retains two independent argument structures, corresponding to the *heavy* interpretation of Grimshaw and Mester 1988. The *light* interpretation corresponds to the result of argument structure merger in my analysis; I am not assuming specially derived light verbs devoid of thematic content.
4. An example might be *hamlaa kar-naa* 'attack-do, make an attack on NP attack NP', assuming that the N selects a locative object with the postposition *-par* 'on'.
5. Only the local subject antecedent identifies *apnee* 'self's', though there are two c-commanding

‘One child snatched *self*’s toys from another child.’

6. 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. The Ns which assign only genitive case may have another lexical representation which blocks the assignment of genitive case, based on a suggested by Dubinsky 1997 for Japanese. Dubinsky proposes that N loses its categorial distinction from an adjective. Adjectives may also have argument structures and assign case, but not genitive or lexical case. In the Type III cases, theta identification is forced to apply early because the A cannot merge with and case-mark its thematic object. Hindi/Urdu has a large number of A+V complex predicates which assign structural object case, such as *mazbuur kar-naa* ‘make forced, force, insist upon’ (ia). This case is structural, as it disappears in the passive (iib):

i) a. unhooN-nee DaakTar-*kaa* [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 ki-yaa ga-yaa hai
3s-gen life agin rocks -on dash-against-inf-gen-for forced do-pf go-pf is

‘Her life again was forced to dash against the rocks.’ Bahl 1979:37

7. Adjectives assign oblique case (i) or dative (ii):

i) gadhee-jaisii aawaaz ‘voice like a donkey’ (ii) [jaa-nee]-koo taiyyaar ‘ready to go’
donkey-obl-like voice go-inf -dat ready

8. Jayaseelan 1988 gives a very clear account of the problems posed by these constructions for the Projection Principle. His solution is for N to postpone combination with its locative object, and to project its argument structure to V’ after N discharges the object theta role of V. This proposal has some similarities with my proposal, in that theta identification can be early or late. Jayaseelan’s account is somewhat different, in treating N as a thematic argument of V, even with open theta positions which would violate Higginbotham’s condition.

The V’ then assigns/licenses dative case to the remaining internal argument. This analysis depends on the properties V, in particular one which means ‘give’. In fact combinations with ‘give’ may well have an analysis like this, but it does not carry over to the combinations discussed here, where verbs such as ‘make’ and ‘come to’ have to accommodate a thematic object as well as a syntactic object in one syntactic position, rather than the two object positions allowed by ‘give’..

9. Fassi Fehri uses the evidence of Raising to Object within the masdar phrase to show that case

10. The early or late option for theta identification also account for the two forms of ‘argument transfer’ in Grimshaw and Mester 1988. If N combines with its thematic argument, theta identification follows and it ‘late; this is partial argument transfer. Complete argument transfer is the same as my ‘early’ theta identification.

11. These verbs might be called ‘intransitive’ because they pattern morphologically with monovalent verbs. But semantically they are mostly bivalent, with implied or expressed goal arguments.

12. Recall that the ergative case is required when the sentence is perfective and finite (and has no nominative-subject auxiliaries). Otherwise the subject is nominative in finite clauses. There is a small number of exceptions to the preservation of ergative case (Bailey 1963, Porizka 1963). The experiencer verbs *dikhaaii dee-naa* ‘sight give; see, be visible to; and *sunaaii dee-naa* ‘hearing give; hear, be audible to’ have a dative experiencer and a nominative theme.

13. These lexical cases are the same ones used for the agent of a detransitive simplex verb (43) and the experiencer of a simple verb such as *dikh-naa* ‘be visible to, glimpse’.

i) mujh-*see* yah peeR kaT nahiiN ga-yaa kaT-naa ‘be cut’ kaaT-naa ‘cut’
I-from this tree be-cut not go-pf

‘I couldn’t cut down this tree.’

ii) us-koo naaNw dikh -tii hai dikh-naa ‘be seen’ deekh-naa ‘see’
3s-dat boat be-seen-impf is

‘The boat is visible to him, he sees the boat.’

14. A somewhat idiosyncratic combination involves Ns which all describe serious or unpleasant experiences with the verb *khaa-naa* ‘eat’:

i) *.dhookaa khaa-naa* ‘eat deception, be deceived’,
gam khaa-naa ‘eat sorrow, suffer’
Thookar khaa-naa ‘eat stumbing, stumble’
maar khaa-naa ‘eat a beating, get beaten’,
GoTaa khaa-naa ‘eat a dive, get a ducking, nearly drown’ (Bailey 1962, 59).

(i) us-nee patthar-see Thookar khaa-yaa ‘He/she stumbled on a stone.’ (Porizka 1963L 439)
3s-erg stone-from stumble eat-pf

The external argument of *khaa-naa* is identified with an experiencer or theme associated with the N. For example:

. These combinations have a passive meaning ‘be beaten, be deceived’ though the syntax is active

15. In sum, the following combinations with ditransitive verbs are well-formed by Higginbotham’s constraint (i)-(iii), while (iv-vi) are illformed::

i) V: 1, 2, 3 (ii) V: 1, 2, 3 (iii) V: 1, 2, 3
N: 1, 2, 3 N: 1, 2 N: 1

iv) V: 1, 2, 3 (v) V: 1, 2, 3 (vi) V: 1, 2, 3
N 1 - 2 N 2 1 N 1,2

16. English has many complex predicates in which both N and V are ditransitive (Jackendoff 1971, Cattell 1984):

- i) N. *promise* X promises Y (to) Z
- ii) N. *blame* X blames Y for Z, X blames Z on X.

Other expressions are more problematic: *give NP a kiss = kiss*. It appears that *kiss* has only a result interpretation, not an event interpretation:

iii) *John’s kiss of Mary.

17. For example:

i) us-nee hameeN bharoosaa dilaa-yaa [ki ham yah kaam kar sak-tee haiN]
3s-erg 1pl-dat confidence give-caus-pf that 1pl this work do be-able-impf be-pres-3p

‘He gave us confidence [that we can do this job].’

18. Pustejovsky 1995: 70ff resolves complex events into two sub-events, which have two dimensions of variation, precedence and headedness.. Events may be ordered strictly, or overlap wholly or partially, representing the varying properties of lexical verbs. The head of the whole event may be the first or second of the sub-events, or both (yielding ditransitive and symmetric predicates). Unheaded events allow several interpretations, such as causative and unaccusative inchoative meanings. This representation captures event structures of many different aspectual types when combined with argument structures and the qualia specifications of the sub events. I adopt this view because of its generality and wide coverage.

An alternative which gives a more syntactically structured account of bounded events uses three sub-events (Hale and Keyser 1993, Folli 2000, Butt and Ramchand 2003) the initiating event, a change and a resulting state, which are strictly order, and have a common argument as object of

19. This explains the interpretation of (i) focussed on the internal argument, and the ill-formedness of (ii) versus (iii)

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' CHECK

iii) meerii raam-kee baaree-meeN yaad

my Ram-gen about memory 'My memory of Ram' CHECK

The internal argument of *yaad* is specified as a proposition. A DP satisfying this position is coerced as a proposition. See Pustejovsky 1995: _____ for discussion of coercion and type shift.

20. The predicate then combines with the subject, which saturates the (x) argument position. Existential closure at the sentential level, satisfying a requirement of semantic completeness. (Chung and Ladusaw 2003:10).

21. See Dayal 2003, Farkas and De Swart 2003, Chung and Ladusaw 2003 for somewhat different proposals for combining the N as a predicate linked with the argument structure of the verb.

Hindi/Urdu has some N + V combinations which are not explained by argument identification. They consist of the verb *maar-naa* 'beat, kill' with a N referring to some manner of striking, such as *thappaR* 'slap' or instrument *goolii* 'bullet', 'to shoot', *juutaa* 'shoe'. The results mean 'slap, hit/kill with a bullet', 'beat with a shoe.' These Ns are bare indefinites, and it is plausible for the instruments that they do not have argument structures of their own. The N is also not an argument of 'beat', and instead it is an adverbial modifier of the event argument. The result of combining *juutaa maar-naa* 'shoe-beat' would be (i)

i) 8y8x8e {beat' (y) (x) (e) with-shoe'(e)}

22. This construction involves non-local phifeature checking (Bhatt 2003) or a series of local relations, if the matrix verb agrees with features on the embedded infinitives.

23. Verbs used in Persian for complex predicates include (Karimi 1997, Karimi-Doostan 1997):

kar-dan 'do'

dad-an 'give'

dasht-an 'have'

raft-an 'go'

amad-an 'come'

bord-an 'carry'

keshid-an 'draw, pull'

avad-an 'bring'

