

Feature percolation and agreement in Hindi-Urdu

Alice Davison  
U. of Iowa

ABSTRACT

Agreement in Hindi-Urdu links the verb-inflection complex with an argument in the clause. The choice of agreement is variable, and depends on structural features, such as the presence of a Postposition (which blocks agreement), argument status, and the configurational priority of the NP. Agreement is presented as the percolation of morphological features onto a predicate, reflecting operations on Theta roles proposed by Speas (1990). Agreement is triggered by the left-most predicate on the Theta Grid which has its F- features represented. This proposal accounts for long-distance agreement over clause boundaries, whether or not the agreement trigger is contained within the higher clause at S-structure. It also offers an explanation for differences among N-V predicates, which in some cases single out the non-referential N as the trigger, rather than the object of N-V.

0. An analysis of agreement in Hindi-Urdu has to account for some special features of agreement in this language. Within a clause, the source of agreement features on the verbal complex is not just the subject. Depending on various forced or optional choices of Case-Marking, the verb may agree with the object or a constituent of an argument clause as well. Hence the source of agreement features may be separated by an internal clause boundary from the matrix finite verb which realizes these features. If both subject and object in a simplex clause are possible sources of agreement features, the subject has priority, as only one set of agreement features appear on a verb. I will propose that the coindexing of a verb with the appropriate agreement features under these conditions is quite mechanical. Agreement reflects predicate argument structure in a very general way, and is filtered primarily by postpositional Case marking. The presence of a postposition on an argument effectively blocks agreement.

The theoretical approach I will take involves percolation of features outward from argument expressions, whether referential or not. The representation of argument structure adapts proposals by Speas (1990) for including a theta grid as a property of syntactic projections. The percolation of features is a version of percolation of indices in Kayne (1981) and Sung and Cole (1990). It is therefore not crucially dependent on movement, particularly not on movement to a Specifier position of a functional category, where Specifier-Head agreement transfers agreement features to the head of the projection. (Nothing precludes movement from taking place in a limited domain, as adjunction to either left or right of maximal projections is possible in this language). Features are represented as a single complex which 'spreads' along surface syntactic projections. An earlier sketch of this proposal was given in Davison (1988), which made use of the **Head Feature Convention** and the **Control Agreement Principle** of Generalized Phrase Structure Grammar (Gazdar et al. (1985)). The Head Feature Convention extends to syntactic approaches which form phrases projected from

lexical categories.

I believe this approach, though somewhat speculative, offers some advantages over movement accounts such as Mahajan (1990). It is generalized to include not just the monomorphemic verbs on which Mahajan's account is based, but also the N-V predicates which form a very large percentage of the lexical verbs in this language. In many of these N-V compounds, the N may trigger agreement just as a normal object would, with the exception that the N itself is not independently referential, and hence must be non-specific. In another set of cases not accounted for by Mahajan, left or right adjunction may remove the agreement trigger from a position from which could be raised to a Specifier position, and yet agreement is not affected.

I will show how this proposal works for some examples of the types of sentences described above, but in the process I will note some important problems which this analysis and most other conceivable approaches must confront.

1. In this section, I will give a brief overview of agreement patterns in Hindi/Urdu. The first important condition to note is that the NP with which the verb agrees is always **nominative**, or lacking in a postposition.

a) The verbal agreement in a single clause is triggered by a nominative subject, independently of whether the object is marked with a postposition [1a] or is nominative [1b].<sup>1</sup>

1) Subject agreement:

a. **laRkee** is kitaab-koo paRh rah-**ee hai**  
boys this book dat read prog are  
m.pl. f. sg. mpl.-III-pl.

'The boys are reading this book'.

b. **laRkee** yah kitaab paRh rah-**ee hai**  
boys this book read prog are  
m.pl. f.sg. mpl.-III-pl.

'The boys are reading this book'.

c. **billii** baccee laa-**eegii**  
cat children bring-fut.  
f.sg. m.pl. bring-III/sg/f/

'The cat will bring its kittens'.

d. **billii** baccee laa-**ii**  
cat children bring-pf.  
f.sg. m.pl. f.sg (III-sg)

'The cat brought its kitten' (cf. Khan (1989b))

b) If the subject is marked with a postposition (such as **-koo** 'dative' or **-nee** 'ergative') the verb agrees with a nominative object.

2) Non-subject agreement

a. laRkoo-nee **yah kitaab** paRh-ii  
 boys-erg. this book read-pf  
 m. pl. f.sg f.sg.

'The boys read this book'.

b. baccoo- koo **sarm** aa-ii / aa-tii **hai**  
 children-dat. shame come-pf come-impf.is  
 m.pl. f.sg f.sg f.sg 3p.sg.

'The children were shy/the children are shy'.

c) If both the subject and the object are marked with postpositions, the verb agrees with no other 'nominative' NP trigger, and has the default features **masculine singular third person** (cf. Porizka (1963), Pandharipande and Kachru (1976), Khan (1989) etc.

### 3) Default agreement

laRkoo-nee is-kitaab-koo paRh-aa  
 boys-erg. this book-dat. read-pf.  
 m.pl. f. sg. m.sg-(III-sg)

'The boys read this book'.

Features are realized slightly differently depending on whether the verb is combined with finite or non-finite inflection (see Appendix A). I use the cover term **features** or simple **F** to refer to both types of grammatical agreement features. Note that an inflection complex realizes the F-features of at most one NP.

2. The second important condition to note is that agreement is triggered by arguments, not adjunct modifiers. For example, in [4], the subject and object are both marked by postpositional Case, so the verb has the default F features, as in [3] above:

4) a. laRkii-nee **raat bhar** in ciTThiyoo-koo likh-aa  
 girl- erg. night whole these letters-dat read-pf  
 f.sg f. sg. f.pl. masc-sg.

'The girl wrote these letters for the whole night' (Khan (1989))

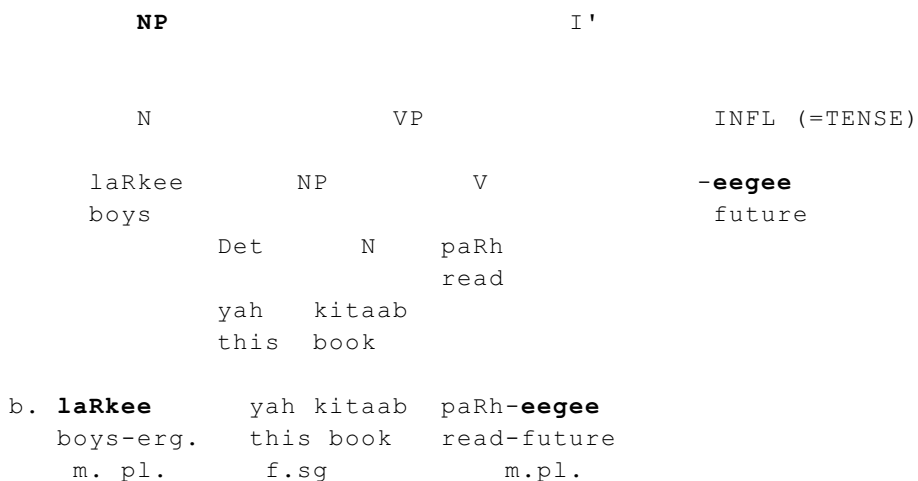
b. pulis-nee coor-koo **jaldii(-see)** pakaR liyaa  
 police-erg thief-dat. haste-with seize- take-perf.  
 f.sg ms.sg f.sg masc.sg.

'The police quickly caught the thief'(Mahajan (1990, 88))

Adjuncts, especially adverbials, are not possible agreement triggers, even if, like **raat bhar** in [4], they lack overt postpositions. The agreement trigger must have an argument role of some kind. It is not enough, as Baber Khan (1989a) points out, for a sentence constituent to be nominative and to be in a position to the left of VP. Feminine singular NPs like **raat bhar** in [4a] or **jaldii** 'haste, in a hurry'in [4b] do not trigger agreement when they are used as adverbials.

In [5] and [6] below, phrase structure trees illustrate the surface structure relations between the verbal complex which manifests the agreement, and the NP which triggers agreement. In order to show the structural similarity of [5] and [6] I have abbreviated TENSE and ASPECT as INFL. The structure in [5] is the more usual instance of agreement with a nominative subject in the non-perfective:

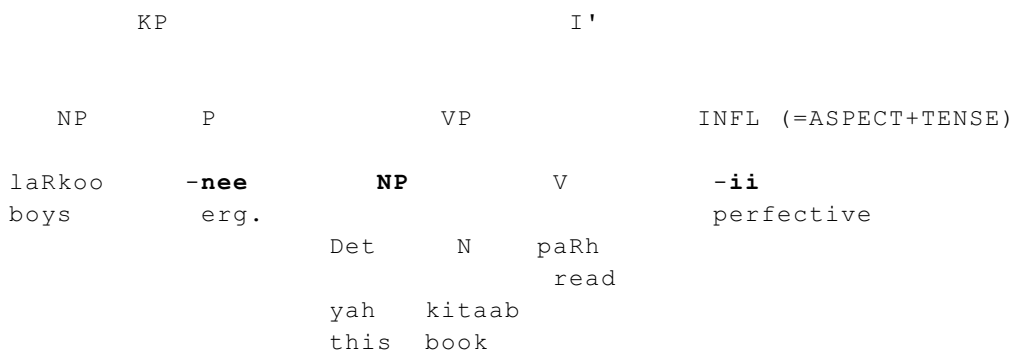
5) a. IP (=TP)



'(The) boys will read this book'.

The surface structure in [6a] is an example of object agreement: the subject NP has postpositional Case (here represented as **KP**, or Kase Phrase, made up of a NP and a postposition used as a Case maker, a phrase which is syntactically equivalent to NP in an argument position, rather than an adjunct or modifier PP).

6) a. IP (=TP)



b. laRkoo-nee    **yah kitaab**    paRh-ii  
 boys-erg.    this book    read-pf  
 m. pl.    f.sg    f.sg.

'The boys    read this book'.

The intuitive generalization that has been made by a number of writers on Hindi is that postpositions on argument NPs block agreement. The subject phrase, the leftmost, highest argument in S-Structure, has priority for controlling agreement unless it has a postposition which prevents morphological coindexing on INFL (Aspect and tense). In this case, the object and then the default set of morphological features take priority.<sup>2</sup>

3. It is important to note that in Hindi/Urdu, verbal agreement involves morphological features, not referential indices. Whether or not INFL has the **morphological** features of the subject, the subject is still a specific (external) argument of the predicate. I will therefore make a distinction between **morphological** features (F) and **referential** features. The morphological features are those which are realized on INFL, and are grammaticized. Their source is the lexical features of common nouns and NP expressions like proper names, pronouns, etc. For example the pronoun **tum** 'you, familiar' has the lexical features **II person, plural**, although its reference is to a single individual. In a sentence, a NP has these morphological features and in addition, a referential index which figures both in the discourse representation and in syntactic coindexing. The F features and the referential index (R) may be composed on a single node, or dissociated.

For example, a NP like **vee kitaabee** 'those books' would have the following representation, showing the association of features with syntactic nodes:

7)		NP(i)	
		(F <sub>i</sub> )	
	DET		N
	(F)		(F)
	vee		kitaab-ee
	pl.	F	plural

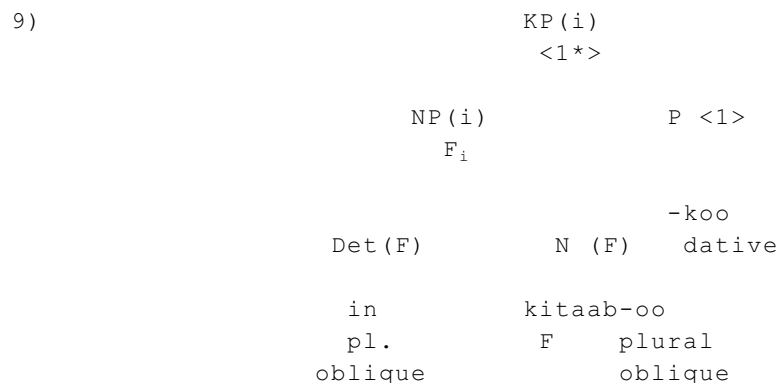
The lexical NP **kitaab** has the morphological features **third person** and **feminine**. This is added to by the plural morpheme **-ee**. The set of F-features **third person, feminine, plural** is propagated to the maximal projection NP, which bears the referential index (i) of the whole phrase. For clarity, I have added a referential subscript to F at NP.

There is a principle which percolates the grammatical gender of N to NP, and the plurality of N to Det and NP. This is the Head Feature Principle (Gazdar et al. (1985), Sung and Cole (1990)), which says:

8) **Head Feature Principle:**

- a. The features of the mother node and the features of the daughter nodes must be the same.
- b. If the features of the daughters conflict, the mother node will have the features of the head node.

Postpositions in Hindi/Urdu have the exceptionless property of being inert for morphological features in their objects. I represent this fact by subtracting the F features from the maximal projection of P, while retaining the R index.



The addition of a P to NP(i) introduces the morphological feature **oblique**, added to the feature complex of the NP constituents. By the Head Feature convention, a feature like **oblique** which is imposed on NP by P is spread onto all the daughters. The HFC also filters out the features F<sub>i</sub> on the nodes above NP, so that KP has only the R feature of NP(i). The head of KP is P, so its features (which are unrealizable) must prevail over those of its NP argument.<sup>3</sup>

4. As argument status is crucial to agreement, it is necessary to discuss the representation of the argument structure of predicates. I will adopt the version of argument structure in Speas (1990). Categories such as P and V are represented as having a grid specifying how many arguments they take. For example, P **-koo** in [8] above has the structure in [10a], a **theta grid** showing that it requires one argument.

- 10) a. P: **-koo** <1>  
 b. V: **paRh** <1 2>

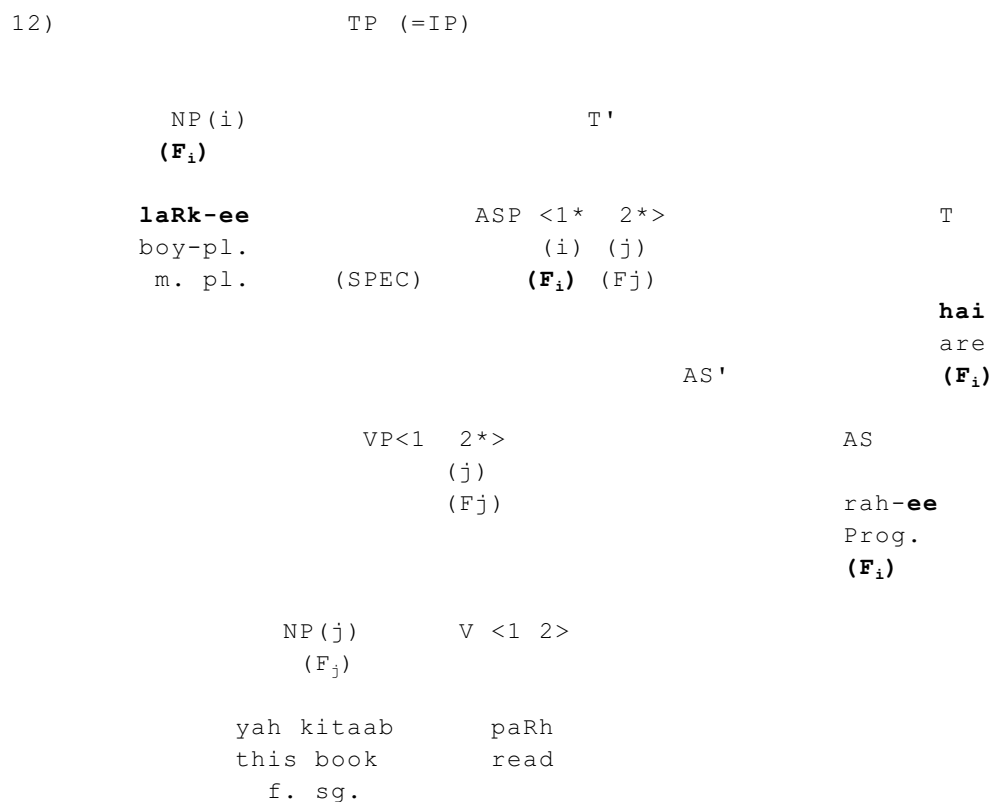
A transitive verb like **paRh** 'read' requires two arguments. The theta grid of the lexical categories is projected up to phrases of which they are the head. If the theta grid is satisfied by the presence of an argument in some projection, the position is starred, as is the case for the KP in [9] and [10a]. A V projection containing the verb **paRh** and its direct object would have the theta grid in [10b] showing that one argument position is discharged, or satisfied, while the other is still open.

- 11) a. P-max: [ N-**koo** ] <1\*>  
 b. V': [ NP **paRh** ] <1 2\* >

The <1 > position is discharged when the V projection is combined with the subject NP. An expression is **saturated** if all the positions in its theta grid are discharged. The grid for **paRh** is saturated within VP if subjects originate there (the Lexical Clause Hypothesis in Speas (1990)) or in IP if subjects originate as external arguments in Specifier of IP.

Below I will note some problems with both analyses, but for the discussion of Case and agreement in this paper I assume that subjects are outside VP in Specifier of I (ASPECT+TENSE). The choice of nominative versus ergative Case depends not just on the properties of the main verb but also on those of ASPECT and any auxiliary or vector verbs.<sup>4</sup>

5. In the sections above, I specified two sets of features, the F morphological features and R indices, which are associated with agreement and coindexing. Phrases whose heads are lexical categories have theta grids associated with them, which contain positions discharged by arguments. In Hindi/Urdu, I will propose that theta discharge and feature percolation are both associated with the theta grid of VP, but are determined independently. Hence they may diverge when P blocks transfer of F features. In this section, I will compare a sentence structure which requires subject agreement with a structurally similar sentence in which the INFL gets the F-features of the object. Theta discharge is identical in both. The structure of [1a] is given as [12]:



'The boys are reading this book'.

The theta grid of the verb **paRh** is undischarged in V. In the next higher projection, the <2\*> is discharged, and both the referential index of NP(j) and its morphological features F<sub>j</sub> are added to the theta grid itself. I make this assumption because the referential index is added to identify **which** syntactic constituent satisfies which argument position, and because I have cross-

referenced the sets of agreement features to referential indices by using subscripts on F. In a higher projection, the index of the subject and its features  $F_i$  are added to the theta grid, which continues to be projected upward to the maximal projections of the functional categories ASPECT and TENSE. The Head Feature Convention will then guarantee that these features are also represented on the heads of these projections.

This sentence has two NPs which are in argument roles, and both sets of F features are attached to the V theta grid. But only one set of agreement features may be realized in the verbal complex in this language (See Appendix B for some other possibilities).<sup>5</sup> Subject priority may be stipulated by saying that the index of the leftmost (discharged) argument in the theta grid selects the features realized on INFL.<sup>6</sup> The principle may be restated on the analogy of a Case-checking rule:

13) Agreement is well-formed if the F on INFL is indexed to the leftmost argument having F features.

We will see that this principle has more general application.

In the next example [14], the agreement is with the object, as in the sentence above [2a].

14)	TP		
	KP <1*> (i)	T'	
	NP(i)	P	ASP <1*2*> (i) (j) ( <b>F<sub>j</sub></b> )
laRkoo- boys-obl m. pl.	-nee erg. (SPEC)		TENSE(null)  AS'
		VP<1 2*> (j) ( <b>F<sub>j</sub></b> )	ASPECT  -ii perfective ( <b>F<sub>j</sub></b> )
	NP(j)	V <1 2>	f. sg. (F <sub>j</sub> )
	<b>yah kitaab</b> this book	paRh- read-	
	'The boys read this book'.		

The presence of the postposition **-nee** 'ergative' on the subject NP prevents its agreement features from being percolated to KP and TENSE/ASPECT. This postposition marks subjects if the following conditions hold:

15) **Conditions on -nee marking of subjects:**

- a. The main verb is lexically specified as an ergative subject verb.
- b. The clause has perfective aspect.
- c. All other auxiliary or vector verbs combined with the main verb are specified as ergative subject verbs.

Condition [15a] may hold for the main verb, within its immediate projection, but if condition [15b] and [16c] are not met in the constituents outside of VP, then the subject is nominative.<sup>7</sup>

For this reason, I represent the phrase marker on which agreement is determined as [14], showing the subject NP in Spec/TP, and marked as ergative. The theta grid bears an index for the subject KP, but this constituent bears no F-features reflecting the NP features, as **-nee**, the head of KP, imposes its (inert) features on the whole projection. The theta grid has only the F-features of the object, and these are the ones realized on Inflection.

I will conclude the demonstration of theta discharge and feature percolation by looking at two more instances of the interaction of agreement with surface syntactic structure. The first involves agreement across non-finite clause boundaries, showing how percolation can be quite extensive within a finite clause. The second involves N-V compound predicates, and shows again the divergence of agreement and theta discharge caused by the presence of a postposition.

The dative-subject construction in the sentence in [16] has an infinitive complement, whose subject is controlled PRO. Agreement with the matrix subject is blocked by the dative postposition **-koo**, so the agreement is optionally determined by the infinitive object **saaiikal** 'bicycle'. (PRO, as an ungoverned, non-lexical element, has no grammatical F-features.)<sup>8</sup>

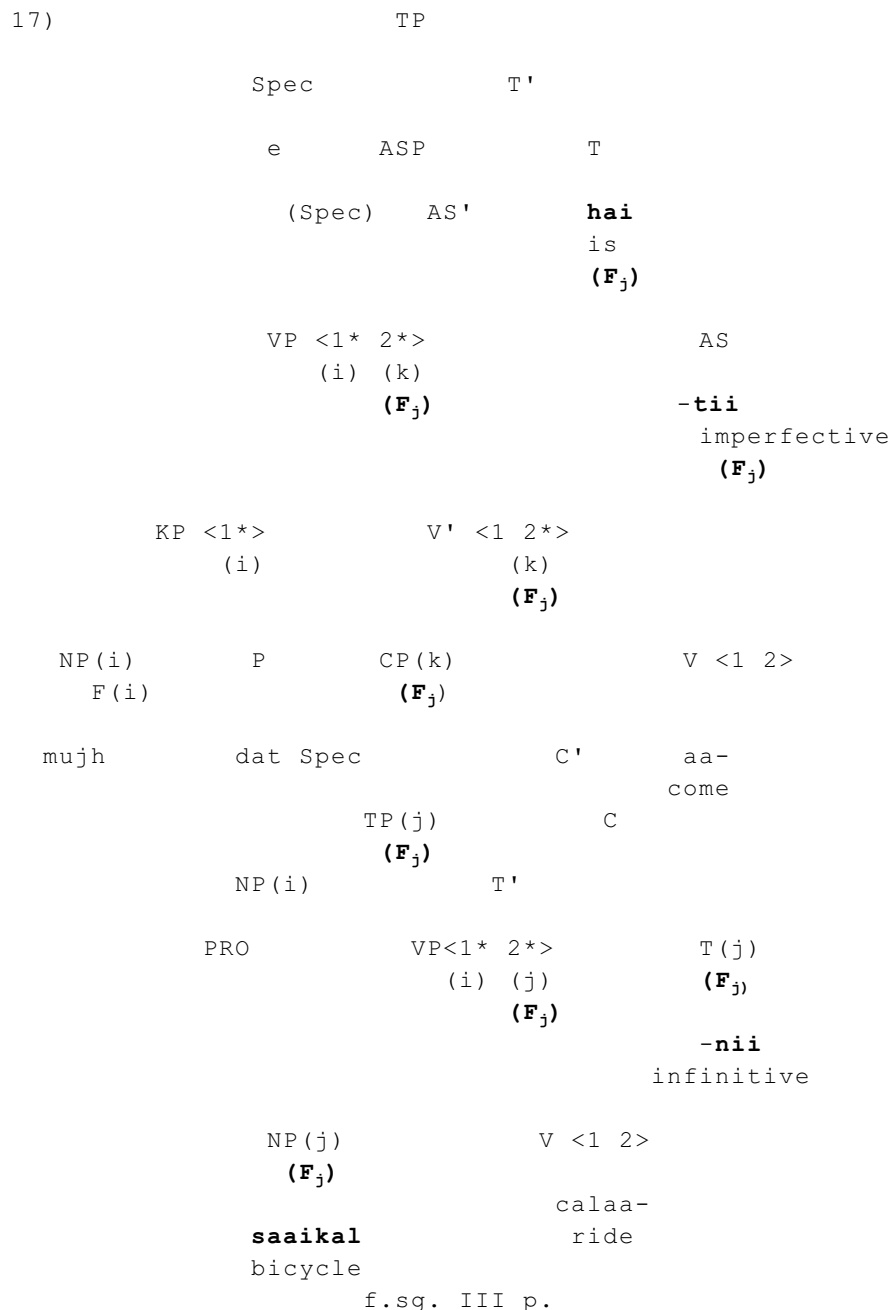
16) mujhee(i) [PRO(i) **saaiikal** calaa-nii] aa-tii hai  
 I-dat. bicycle ride-inf come-impf is  
 I-sg. f.sg. f.sg. f.sg. III sg.

'I know [(how) PRO to ride a bicycle]'

Agreement within the infinitive is realized on the infinitive suffix **-naa**, in TENSE. The infinitive constituent has nominal properties, like **-ing** gerunds in English, but it is also clause-like in having internal agreement on the same principles as main clauses. Since the subject is unavailable (having only an index determined by the matrix subject), the verb **calaa-nii** 'to ride' agrees with **saaiikal** 'bicycle'. The internal agreement features in TENSE are projected to TP (and CP if we assume that infinitives are fully clausal, but nothing hinges on this). The infinitive clause has F-features, and these are attached to the matrix theta grid along with the index of the clause as the argument of the matrix verb. Since the matrix subject has no F-features, because of the postposition **-koo**, the only argument on the matrix theta grid with F-features is the clausal complement.

The phrase structure of sentence [16] is given in [17] below. We see that the F-features of **saaiikal** 'bicycle' are realized on the complement inflection and on the matrix tense and aspect. The complement clause takes on the referential index and F-features of its head, namely Inflection,<sup>9</sup> though the theta grid specific to the verb **calaa** 'ride' is not projected beyond TP/CP, nor is the referential index (j) of **saaiikal** 'bicycle'. Instead, the infinitive has

its own index (k) which discharges the 2 position of the matrix theta grid. Infinitives are nominals normally having only the default F-values. But by the Head Feature Convention, they take on the F-features of inflection.



The non-finite complement in this sentence is able to be postposed to the right (or left), like any clause constituent:

18) mujhee zaruur e<sub>k</sub> aa- **tii hai** [PRO **saaikal** calaa-**nii**]<sub>k</sub>  
 I-dat. surely come-impf is bicycle ride-inf

'I certainly know [(how) PRO to ride a bicycle]'

Whatever differences there are in syntactic structure between [16] and [18], they do not affect agreement over a clause boundary. I will return to the point when comparing percolation along phrase structure paths, proposed here, with raising of NP to Specifier positions, as in Mahajan (1990).

6. I conclude with the last set of examples, N-V compounds which involve merger of the theta grids of N and V, and optional assignment of a genitive postposition to the object of N. The theta roles of N-V are discharged in exactly the same way as with V, but object agreement is different from the examples given above. N-V compounds in Hindi are made up a generic verb like **kar-naa** 'to do' combined with either an abstract noun like **madad** 'help' or a concrete noun like **Teelifoon** 'telephone' used nonreferentially.

19) a. mai-nee un- **kii madad kii** (Genitive P is obligatory)  
I-erg. 3pl. of help do-perf.  
Ip. sg f. f. sg. f.sg

'I helped them'

b. mai-nee un-**koo Teelifoon kiyaa** (-**koo** is obligatory)  
I-erg 3pl.-dat. telephone do-perf.  
m. sg. m.sg.

'I made a telephone call to them/I telephoned them'.

The N of N-V is the source of agreement features, whether it is an abstract noun or a nonreferential concrete noun. Modifiers of the N actually modify the whole action: a **buraa Teelifoon** 'bad telephone' is a crank call, not a malfunctioning instrument.

An interesting subclass of N-V involves an **optional** genitive on the object, as illustrated by the pair of sentences in [20] from Hook (1979). The main verb is **isteemaal kar-naa** 'to make use':<sup>10</sup>

20) a. isteemaal (m) kar-naa 'to use'

bhaiyaa-nee **apnii taaqat** isteemaal **kii**  
brother-erg. self's strength use do-pf.  
m.sg. f.sg m.sg f.sg.

'Brother used his strength' (Hook (1979:158))

b. diidii-nee [apnii aqal- **kaa**] **isteemaal kiyaa**  
elder-sister-erg. self's wits of use do-pf.  
f.sg. f.sg. m.sg m.sg m.sg

'Sister used her wits'. (Ibid)

In [20a], the N-V combination **isteemaal kar-naa** 'do use' functions as a single unit, and does not include the object **apnii taaqat**. This sentence has object agreement just like [2a]. In [20b], the syntactic object is linked by a structural genitive postposition to the nominalization **isteemaal** 'use', which has

its own theta grid < 1 2 >. This whole NP is the syntactic object of **kar-naa** 'to do', and is the source of a different set of agreement features.

This difference can be represented as a subcase of the set of conditions already proposed. The only new mechanism which is invoked is **theta merger**, which is another possibility discussed in Speas (1990) for the combination of two constituents. Theta merger is compared with theta discharge in [21]:

21) a. **Theta discharge:**

NP <1\* > P <1> : PP <1\*>

b. **Theta merger:**

N < 1 2 > V < 1 2 > : N-V < 1 2 >

The combination of P with an open argument position with a saturated expression NP allows the P's argument position to be satisfied or discharged, as in [21a]. The combination of two expressions with open positions, as in [21b] precludes discharge, but allows the two grids to merge as one. (If V is a one-place verb like **hoo-naa** 'to be', a passive-like lexical derivation takes place, and the 'inner' argument is retained in the new predicate.)

The structures of these sentences in [22] below allows us to see that theta merger applies to different parts of the corresponding phrase structure trees. Only the VP is shown, as the rest of the structures are the same as sentences like [2a].

22) a. bhaiyaa-nee **apnii taaqat** isteemaal **kii**  
 brother-erg. self's strength use do-pf.  
 m.sg. f.sg. m.sg. f.sg.

'Brother used his strength' (Hook (1979:158))

b. VP < 1 2\* > (Theta discharge)  
 (j)  
 (F<sub>j</sub>)

(F<sub>j</sub>) NP(j) V < 1 2 > (Theta merger)

NP	N	N(k) < 1 2 >	V < 1 2 >
		(F <sub>k</sub> )	
apnii	<b>taaqat</b>		kar-
self's	strength	isteemaal	do
f.	f.sg.	use, m.sg	

c. diidii-nee [apnii aqal- **kaa**] **isteemaal kiyaa** elder-  
 sister-erg. self's wits of use do-pf.  
 f.sg. f.sg. m.sg. m.sg. m.sg.

'Sister used her wits'. (Ibid)

d.		VP < 1 2* > (Theta merger)	
		(j)	
		<b>(F<sub>k</sub>)</b>	
	NP <1 2* >		(Theta discharge)
	(j)		
		<b>(F<sub>k</sub>)</b>	
	KP <1*>	N (k) <1 2 >	V <1 2>
	(j)	<b>(F<sub>k</sub>)</b>	
NP(j)	P <1>	<b>isteemaal</b>	kar
(F <sub>j</sub> )		use	do
		m.sg	
apnii aqal	kaa		
self's wits	of		

In [22d], the KP discharges an argument position in the theta grid of **isteemaal**. The latter is also a N with grammatical F-features. The P in KP filters out the F-features of its object **apnii aqal** 'self's wits', by the Head Feature Convention. But the head N, **isteemaal**, imposes its F-features on NP, again by the Head Feature Convention. In [22b], however, **isteemaal** with no theta positions discharged merges with **kar**, and since the result is a V in category, it carries along no F-features derived from N.<sup>11</sup>

Theta merger and theta discharge reflect the 'basic' constituent structures shown in the trees above. But as phrase order in Hindi/Urdu is not completely fixed, other possible orders are possible without affecting theta discharge and agreement. The object clause in [18] can be right-adjoined, without destroying the coindexing between its internal object and the matrix verb. The KP object of N-V compounds like **madad kar-naa** 'to help' in [23a] can also be permuted away from the N-V predicate, which is itself right adjoined, as in [23]. As both versions are syntactically well-formed, the composition of theta and agreement grids cannot crucially depend on adjacency.

23) a. mai [PRO [[in loogoo kii] madad] kar-naa caah-taa huu  
I these people of help do-inf want-impf am

'I want [PRO to help these people]'

b. [in loogoo-kii] mai zaruur e caah-taa huu [e madad karnaa]  
these people of I surely want-impf am help do-inf.

'I certainly want to help THESE PEOPLE'

Rather, the crucial syntactic relations for theta merger, theta discharge and agreement percolation seem to be preserved via traces left by adjunction to the right or left. The implications of such facts are discussed in the next section.

7 In this last section, I want to contrast this proposal, in which features and theta grids are percolated along syntactic projections, with a proposal which moves the NP trigger of agreement (Mahajan (1989, 1990)). Movement of NP to the Specifier position of the functional categories AGR<sub>s</sub> and AGR<sub>o</sub> has two functions

in Hindi syntax. It provides a source of nominative Case, as in other languages where subjects regularly get Case from finite inflection. It also creates a local relation, of Specifier and Head, for agreement coindexing. Movement to Specifier position from within VP is motivated by the optional or obligatory failure of the verb to assign a structural Case. I will discuss two problems with the 'NP raising' account of agreement raised by the data in this paper. The first involves long-distance agreement, as in [16-18], and [23]. Agreement holds even when the agreement trigger is right (or left) adjoined. The second involves the claim that agreement of the object with AGR<sub>o</sub> necessarily involves raising of the object outside of VP, and that this structure accounts for specificity of the object, as well as adverbial readings. I believe these two sets of problems are serious obstacles to an account based on NP movement.

In sentences such as [16] and [18], the object of a complement infinitive triggers both local and matrix agreement:<sup>12</sup>

24) raam-nee [PRO **rootTii** khaa-**nii**] caah-**ii**  
 Ram-erg. bread eat-inf. want-perf.  
 m.sg. f.sg. f.sg. f.sg.

'Ram wanted [PRO to eat bread]' (Mahajan (1989, 235))

Mahajan accounts for the agreement by a (string-vacuous) movement to Spec AGR<sub>o</sub> in both the complement infinitive, and the matrix finite clause. The result is a structure [25] with **rootTii** in the main clause, triggering matrix agreement on **caahii** 'wanted', while its trace triggers agreement on **khaa-nii** 'to eat'.

25) raam-nee [**rootTii** [PRO **e** [ **e** khaa-**nii**]]] caah-**ii**  
 Ram-erg. bread eat-inf. want-perf.  
 m.sg. f.sg. f.sg. f.sg.

(Mahajan (1989, 237))

This structure precludes the movement of the object and the complement verb as a unit, triggering agreement. They are no longer part of the same syntactic phrase in [25]. Yet long-distance agreement is found in sentences like [18] above, in which the object and verb are found to the right of the main verb. The complement clause can be adjoined to IP on the left or right, like other nominal constituents, with various stylistic consequences. This movement to what we assume is an A' position, does not affect agreement, or other coindexations such as anaphor-antecedent relations. Rather, the trace of the internal clause bears the relevant agreement features, realized on the matrix verb.

For the NP movement account to work, we would have to assume reconstruction, perhaps at LF, and a second opportunity for agreement to be checked, after S-Structure. The F and R percolation account treats constituents as having complex indices, made up of syntactic features such as +N, -V, morphological features making up the F index, and a referential index which both cross-references syntactic constituents, and relates constituents to the discourse representation. I will assume, then, that traces consist of a complex of features, include F and R features, which percolate within constituents. Traces are coindexed with constituents having the matching set of features, and features are realized in accordance with the restrictions outlined above in [8], [13] and [15].

The second set of problems involves N-V compounds, a major set of lexical predicates in Hindi/Urdu. The N is incorporated as a part of a predicate with a 'light' verb like **kar-** 'do' or **hoo-** 'become'. (See also T. Mohanan (to appear).) I have proposed above that Speas' theta merger is involved in relating N-V to the arguments of N. N is not specific in any sense, in referring to a particular individual or mass quantum. Yet it triggers agreement, as in [19] and [20]/[22]b.

Syntactically, **madad** 'help', **Teelifoon** 'telephone' or **isteemal** 'use' is an object of **kar** 'do', and if the placement of P Case permits, the agreement in INFL treats this N as an object. The N in some instances can have modifiers, and thus be syntactically indistinguishable from a referential NP:

26) pro [meerii is-mee baRii **madad**] kii  
       my this-in great help do-perf  
       f sg. f.sg. f.sg. f.sg.

'(They) helped me **greatly in this (matter)**' (Hook 1979, 107)

Agreement does not make this NP specific, nor does the absence of agreement, as in the non-perfective, make the NP non-specific.

These facts do not coincide with the NP raising analysis, if specificity is structure-dependent in the way suggested by Mahajan (1989). I have proposed that N-V compounds are for the purposes of agreement just like other object-verb structures, and agreement is possible independently of specific reference. The possibilities of agreement are defined by the same set of principles which cover agreement with referential arguments as subjects or objects. The main factor is the presence of P as Case. Stating the principles for the placement of P Case is a far from trivial matter, but once the P Case markers are in place, agreement follows automatically.

8. In this paper, I have proposed that there is a single pattern of agreement of inflectional elements with argument NPs in Hindi-Urdu. Agreement within a clause holds between inflection and at most one nominative argument. This pattern involves the combination of argument structure and morphological feature percolation. Argument structures are represented as positions in the theta grid which are discharged by coindexation of the grid positions with the NPs in the clause which bear the relevant grammatical role (Speas (1990)) Coindexation places the R-index or referential index of the argument on the theta grid, and the F or morphological features as well if they are features of the whole phrase which is an argument.

Maximal projections bear the F-features of the head, by the Head Feature Convention (Gazdar et al (1985), Sung and Cole (1990)). Hence NPs bear the features of N, such as grammatical gender and plurality.<sup>13</sup> Postpositions are used to mark all but nominative Case. These Case postpositions are heads of KP, and their (inert) F-features prevail. Hence P blocks F-feature percolation but not R-index percolation onto the theta grid. Arguments are always represented in the thetagrid when they discharge a position, but they may fail to project F-features. The annotated theta grid is projected upward to functional phrases whose heads are ASPECT and TENSE, the morphemes on which agreement of the prevailing argument is realized. This is the left-most argument with F-features in the Theta grid of the predicate, combined with the INFL projections which manifest agreement features.

Agreement is possible over non-finite clause boundaries, in default of an available argument with F-features in the matrix. PRO also is unavailable, since it has no lexical features. An argument within an infinitival argument clause may index its own theta grid, which in turn projects the F-features to the infinitive clause itself. This clause has its own index, as an argument of the matrix clause, but takes on the internal F- features of its inflection. The matrix theta grid in turn projects this to matrix inflection. Dislocation of the internal clause, with the agreement trigger inside it, does not disrupt agreement. Rather the relevant features are preserved as part of the complex set of features in the trace of the dislocated constituent. This agreement possibility argues against a NP raising analysis proposed by Mahajan (1990).

N-V compounds involve merger of the theta grids of the nominal and the 'light' verb. The N may trigger agreement, though it is non-referential and nonspecific. In this case, theta discharge by the object takes place before theta merger. Some N-V compounds, however, allow theta merger before theta discharge, with different placement of P Case, and a different agreement pattern. Overall, the same agreement pattern holds for subjects, objects and N-V, regardless of specificity and regardless of whether the agreement trigger is syntactically local to inflection.

#### Acknowledgements

The research reported here was supported in part by the National Science Foundation, under grant RII-88-00534, and by the University of Iowa. I am grateful for discussion of issues in this paper with Wayne Harbert, Peter Hook, Peggy Speas, Manindra Verma, James Yoon, and the participants in the 1989 Cornell seminar on syntactic problems in South Asian Languages, Jim Gair, Barbara Lust, L. Sumangala and Jaqui Toribio. I also thank the audiences at the 1990 Delhi conference on anaphora and the 1991 Agreement panel at the University of Wisconsin South Asia Conference for their helpful suggestions. The data discussed in the paper, where not otherwise indicated, is based on judgements from the following, who I gratefully thank for their generous help: Pramod Dvivedi, Aravind Gupta, Vijay Gupte, Sanjay Kaul, and M.K. Verma, in the U.S. and Sunita Garg, Minhajudin, Anamika Sharma and Chhaya Mathur in India.

#### APPENDIX

##### Part A: Agreement features in Hindi-Urdu V -ASP -TENSE

27)	AGR1: (=NON-FINITE/ASPECT)	AGR2 (=TENSE)
	number singular/plural	singular/plural
	gender masculine/feminine	person: I, II, III
	(Case direct/oblique)	---

AGR2 is realized only tensed verbs

AGR1 is realized on non-finite forms, aspectual participles and adjectives.

##### Part B: Agreement systems with multiple agreement positions:

28) [Kurmali] The verb has two positions, V(INFL)-AGR-AGR:

AGR1 : Subject  
AGR2 : Indirect object > direct object > direct object  
genitive > subject genitive (Mahto (1989))

29) [Marathi] The verb has two positions, V-ASP-agr-AGR

tu kavitaa vaac- l- i -s  
you poem read pf. agr AGR  
2p f.sg fsg 2p

'You read the poem' (Gair and Wali (1988))

30) [Magahi] The verb has V- TNS-AGR1-AGR2-AGR3 (Verma (1990))

AGR1 encodes subject (person)  
AGR2 encodes object (person, honorific)  
AGR3 expresses 2nd person honorific (addressee)

1. Case except for Nominative is expressed as a Postposition. Postpositional Case imposes oblique morphological case on NP, and forms a projection KP, distinct from Postpositional phrases, which are not arguments and whose heads have lexical, non-selected meaning. Nominative Case refers to a bare NP in the 'direct' or citation lexical form.

2. The agreement trigger must have an argument relation to a predicate, since, as Khan (1989)a, b argues, there are many non-postpositional NPs which could trigger agreement within a clause and do not do so, as in [4]. His conclusion is that the agreement trigger cannot be adequately characterized as the left-most NP in a string of phrases. The phrase structure of the string is relevant, as NP triggers may be within internal clauses, and so is the argument status of the agreement trigger. Argument status may be represented in terms of phrase structure configuration, which is assumed in this paper, or as primitive grammatical relations, such as Subject and Direct Object. The choice between these alternatives is a matter of controversy for a language like Hindi-Urdu, which has relatively free phrase order.

3. All postpositions in Hindi-Urdu block agreement. It is interesting to note that one postposition, the genitive, is itself inflected like an adjective, for number, gender and direct/oblique case. For this reason, I suggest that postpositions in this language all have their own morphological features, which are imposed on PP or KP by the Head Feature Convention. For other languages like Marwari and Nepali, postpositions may vary in having independent, possibly inert features, to allow for agreement with NP-P under very restricted conditions. The conditions for agreement with an ergative NP in Nepali are complex and partly dependent on the type of referent of the NP as well as on aspect (Wallace (1985)).

Hook and Sahai (this volume) discuss agreement between adverbials or their specifiers, and subjects or objects. This agreement pattern is found primarily in Marathi, Gujerati and Marwari, but is present to a limited degree in Hindi.

On the basis of the information available so far, it seems that the adjuncts which reflect the option of agreeing with an argument are similar to genitive postpositions, which reflect the features of the phrase of which they are predicated. Adverbials may, however, retain their own inherent features. For example, a Noun used as a postposition may keep its nominal gender, or allow its specifier to agree with a subject or object.

4. The conditions on the assignment of **-nee** are given in [15] below. A possible alternative preserves the idea that subjects along with other verbal arguments originate in VP. Theta discharge takes place in VP, subjects are raised to Spec/TP position, where their Case realization is determined. F-features are expressed on the theta grid in accordance whether subjects have Nominative or Ergative Case.

5. A complete survey of concord systems in Indic languages is given in Masica (1991, 259ff). In addition to the 'Eastern Hindi' languages mentioned in the Appendix, languages without postpositional Case, there are others like Marathi which do mark Case very much like Hindi/Urdu, and have multiple agreement markers.

6. The priority of the subject is difficult to state without stipulation, independently of the syntactic theory being assumed. The priority of the argument in the theta grid reflects the phrase structure priority of NPs which c-command other NP arguments (Speas (1989), Grimshaw (1989)). Davison (1988), appeals to an interpretation of the Control Agreement Principle of Gazdar et al. (1985). Agreement reflects predicate argument structure, and is registered in a bottom-to-top fashion. Structurally 'higher' arguments propagate their features which overwrite features propagated from 'lower' arguments. Both this and the proposal made here are based on phrase structure configurations which encode argument relations. The string of arguments in the theta grid includes only arguments of a specific predicate, not all the NPs in the sentence.

7. For example, the verb **lee-naa** 'take' normally has a **-nee subject**. **If combined with an intransitive auxiliary baiTh (lit. 'sit', used to express something done in a sudden or annoying manner)**, then the subject of the compound verb is nominative and the verbal complex shows subject and not object agreement in the perfective.

i) ... **vee**      laataa-kii baat      lee    baiTh-**ee thee**  
       3ppl.    Lata    -of matter    take sit-perf. were  
       m.pl                    f.sg. f.sg                    m.pl. m.pl.

'He (hon.) brought up the question of Lata'  
 (Montaut (1991, 39))

8. The dative experiencer NP in [16] presents additional problems in representing subjects. Nominative or ergative subjects discharge a theta position in VP before receiving Case, or in IP after the choice of Nominative or **-nee** is determined. Dative experiencers receive Case in VP, since **-koo** is determined lexically, and is not affected by the choice of aspect or auxiliaries. So theta discharge within VP is unproblematic. Nevertheless, dative experiencers have properties unique to VP external subjects, such as control of PRO and antecedence of reflexives. This property is not represented

in [17], though we might postulate that dative subjects raise to Spec/IP independently of the agreement relations shown here (cf. Davison (1988b)).

PRO subjects are represented in [17] in Specifier of TENSE/ASPECT. The conditions for the licensing of PRO include non-finite TENSE in the projection outside VP. PRO cannot remain VP-internal, where subjects are governed and Case-marked. Experiencer NPs cannot be controlled PRO subjects of verbs which mark their lexical experiencer subjects as dative.

9. This percolation of theta grid properties beyond VP and into TENSE/ASPECT projections may be a consequence of V or VP raising. It might be somewhat less artificially stated in terms of a Complex Category Formation which combines properties of both VP and TENSE/ASPECT (cf. Bayer and Kornfilt (1991)).

10. Other similar N-V compounds include **anubhav kar-naa** 'to experience' (Hook (1979, 159)).

11. According to Hook (1979, 159), if the N of N-V compounds is modified--forming a projection of N, it is obligatory to place a genitive postposition on the object of N. For example:

i) coor-nee apnee dimaag-**kaa** saahii isteemal/prayoog nahii kiyaa  
thief-erg self's mind-of true use /use not do-pf

'The thief made bad use of his intelligence'.

Hence theta discharge takes place before theta merger.

12. Agreement within a non-finite clause is optional, or perhaps only optionally realized. Sentences with long-distance agreement sometimes have special pragmatic implicatures, as in examples in Hook (1979, 30). There is a great deal of individual and possibly regional or stylistic variation in acceptability of the various combinations of agreement in the matrix and complement. It probably is premature to state absolute conditions on what agreement combinations are possible.

Agreement can cross a clause boundary without realization on the embedded verb:

i) us-nee mujhee [PRO baas-waalii kursiya khariid]-**nee dii**  
3psg-erg I-dat. bamboo chairs buy inf give-pf  
f.pl. m.obl. f.pl.

'She allowed me [PRO to buy bamboo chairs]' (based on Hook (1979, 29))

The oblique inflection on the embedded verb **khariid** 'buy' does not realize agreement locally with the feminine plural object, but it does not block its percolation to the higher clause.

13. Interesting confirmation of the role of the head comes from conjoined NPs with non-human referents in Urdu (Naim et al. (1975: Vol. I, 222ff)). The F-features of the right-most conjunct prevail on a following predicate. As in other projections in a head final language, the right-most constituent is the syntactic head. This choice of head is quite general, for all NPs which

trigger agreement:

i) Subject NP

[caae aur **kaafii**] xatm hoo ga-**ii**  
tea and coffee end be go-pf  
m.sg f.sg f.sg.

'[Tea and coffee] are finished'. (Naim et al (1975:223))

ii) Object NP

mai-nee [eek qalam aur **doo kitaabee** khariid-**ii**  
1sg-erg. one pen and two books buy pf  
m.sg. f.pl. f. pl.

'I bought one pen and two books.' (Ibid)

iii) Predicate N-V combination

us-koo [kuch hairat aur **kuch afsoos**] hu-**aa**  
3psg-dat. some amazement and some sorrow be-pf  
f.sg. m.sg m.sg

'He/she felt [some surprise and some sorrow]' (Ibid)

Peter Hook and Tahsin Siddiqi (p.c.) report that the possessive, which precedes the N head, agrees with the left-most conjoined NP:

iv) mai-nee [us-kaa haar aur cuuRii] pulis-kee hawaalee kii/\*kiyaa **I-erg**  
**3ps-of necklace and bangle police-of custody did**  
m.sg. f.sg f.sg./\*m.sg

I turned [her necklace and bangle(s)] over to the police.

The verb on the right agrees with the adjacent verb. If the whole conjoined phrase is adjoined to the right of the verb, the adjunction structure itself is no longer a head-final one. Hook (p.c.) reports that agreement is triggered by the left conjunct:

v) mai-nee pulis kee haweelee kiyaa/\*kii [us-kaa haar aur cuuRii]  
I-erg. police-of custody did 3psg-of necklace & bangle  
m.sg/\*fsg. m.sg f.sg

I turned [her necklace and bangle(s)] over to the police.

The issues raised by the data in (i) - (v) clearly deserve further investigation.

## References

- Bayer, Josef and Kornfilt, Jaklin (1991) Against scrambling as an instance of move-alpha. Paper presented at the Tilburg conference on scrambling.
- Davison, Alice (1988a) Constituent structure and the realization of agreement features. CLS Agreement Parasession, 41-53.
- Davison, Alice (1988b) Case as a motivation for move-alpha. Cornell Working Papers in Linguistics, no. 8
- Davison, Alice (1990) Long-distance coindexing in Hindi-Urdu. Paper presented at the 1990 SALA meeting.
- Gair, James and Kashi Wali (1988) On distinguishing AGR from agr: evidence from South Asia. CLS Agreement Parasession, 87-104.
- Gazdar, Gerald, Ewan Klein, Geoffrey Pullum and Ivan Sag (1985) Generalized Phrase Structure Grammar. Cambridge: Harvard University Press.
- Hook, Peter (1979) Hindi structures: Intermediate level. University of Michigan: Center for South and Southeast Asian Studies.
- Khan, Baber S. (1989a) A note of disagreement on verb agreement in Hindi-Urdu. Linguistics 27, 71-87.
- Khan, Baber S. (1989b) The Urdu case system. U. of Illinois dissertation.
- Mahajan, Anoop (1989) Agreement and agreement phrases. M.I.T. Working Papers in Linguistics, Volume 10, 217-52.
- Mahajan, Anoop (1990) The A/A-bar distinction and movement theory. M.I.T. dissertation.
- Mahto, Panchanan (1989) On the nature of empty pronominals. CIEFL dissertation.
- Masica, Colin P. (1991) The Indo-Aryan languages. Cambridge University Press.
- Mohanan, Tara W (to appear) Object incorporation in Hindi. Natural Language and Linguistic Theory.
- Montaut, Annie (1991) Aspects, voix et diatheses en hindi moderne. Paris: Societe pour l'information grammaticale.
- Naim, C.M. et al. (1975) Introductory Urdu. University of Chicago: Committee on Southern Asian Studies.

Pandharipande, Rajeshwari and Yamuna Kachru (1976) Relational grammar, ergativity and Hindi-Urdu. Studies in the Linguistic Sciences 6.1, pp. 82-99.

Porizka, Vincenc (1963) Hindi language course. Prague.

Speas, Margaret (1990) Phrase structure in natural language. Dordrecht: Kluwer Academic Publishers.

Sung, Li-May and Peter Cole (1990) Feature percolation. U. of Delaware, Ms.

Verma, Manindra K. (1989) Light verbs in Hindi-Urdu. Paper presented at the 1989 SALA meeting.

Verma, Manindra K. (1990) Object agreement and the addressee component in Magahi. Paper presented at the 1990 SALA meeting.

Wallace, William (1985) Subjects and subjecthood in Nepali. University of Illinois dissertation.