

ON THE CATEGORIAL IDENTITY OF INFINITIVES IN HINDI/URDU

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

Introduction:

Like most Indo-Aryan languages, Hindi/Urdu has finite inflection in independent sentences, expressing tense, mood and person-number agreement. Dependent clauses which are non-finite mark the verb with other kinds of inflection, expressing tense or aspect, but not person agreement. Many non-finite dependent clauses in Hindi/Urdu are marked with the infinitive suffix *-naa*. It has been debated whether the Hindi/Urdu *-naa* clauses are full clauses, equivalent to finite clauses, or whether they are nominalizations or gerunds (for example in Butt 1995). In both possibilities, a clause subject is projected as a constituent. It has also been proposed in Bhatt 2005 that infinitives have the option of being projected without a subject. Wurmbrand 2003 has proposed that many languages have this option, choosing a full clause or a subjectless verb phrase, and different syntactic possibilities are available for the whole clause, depending on whether it consists of two integral clauses, or a matrix clause with a truncated infinitive.

This suffix has some paradoxical properties of being both verbal and nominal in character. On the one hand, it seems to contain tense reference, like the English *to*, other infinitive suffixes like the Latin and Romance and German *-en*, and so on. On the other hand, it is inflected like a masculine singular noun in Hindi, with oblique form and postpositional case. Historically this fact is not surprising, as infinitives in Sanskrit were marked with innovative forms which are basically nominal. Somewhat unusually in Indo-Aryan languages, Hindi/Urdu infinitives have the option of showing agreement within the non-finite clause, which spreads to the main clause. This option has been linked to the option in Wurmbrand and Bhatt of projecting an infinitive without a subject.

In this paper, I will propose that the option of a truncated infinitive is not available in Hindi/Urdu. In the relevant infinitive clauses, there is evidence for the presence of a projected syntactic subject, even if it is not always pronounced.

I. Embedded infinitives may be the subcategorized objects of the main verb, as in (1) or its subject, as in (2). Note that the embedded subject may be null, as in (1), and (3) where it is coreferential to the matrix clause subject or object. This kind of subject is represented by PRO, which is identical in reference to an antecedent in the matrix clause. Overt genitive subjects are possible with subject infinitives (2).

Object clauses:

- 1) laRkooN-nee [PRO yah kitaab paRh-**naa**] caah-aa
 boys-ergative this book.fem.sg read-infinitive.ms. want-pf-masc.sg.
 ‘The boys wanted [PRO to read this book].’

Subject clauses:

- 2) [basooN-kaa Diipoo-see nikal -**naa**] shuruu hu-aa
 bus-pl-gen depot-from come.out inf.m beginning be-pf.m
 ‘The buses began to leave the bus station.’

II. ‘Infinitives’ have noun-like properties, such as obligatory oblique case (3) postpositional markers (4), and optional agreement in number and gender (5).

- 3) [Oblique case]

maaN-nee raadhaa-koo [PRO apnee aap-koo deekh-**nee**] nahiiN diyaa
 mother-erg Radhaa-dat self’s self-dat see-inf-obl not give-pf.ms
 ‘Mother did not allow Radha [PRO to look at self].’

- 4) [Postpositional markers selected by the matrix predicate]

a. maiN-nee [PRO andee khariid-**nee**]-**kii** kooshish kii
 I-erg eggs.m.pl buy-inf.obl-gen attempt.f.s do-pf.f.s
 ‘I tried [PRO to buy eggs].’

b. laRkee-nee [PRO yah kitaab paRh-**nee**] -**see** inkaar kiyaa
 boy-erg this book.f.s read-inf.obl from refusal.m.s do-perf.m.s
 ‘The boy refused [PRO to read this book].’

c. coor-koo [PRO paisee curaa-**nee**]-**par/-kaa** pachtaavaa thaa
 thief-dat money.mpl steal-inf.obl-on/gen regret.ms be.pst.m.s
 ‘The thief regretted [PRO stealing the money].’

- 5) [Optional noun agreement marked on the infinitive and matrix verbs.

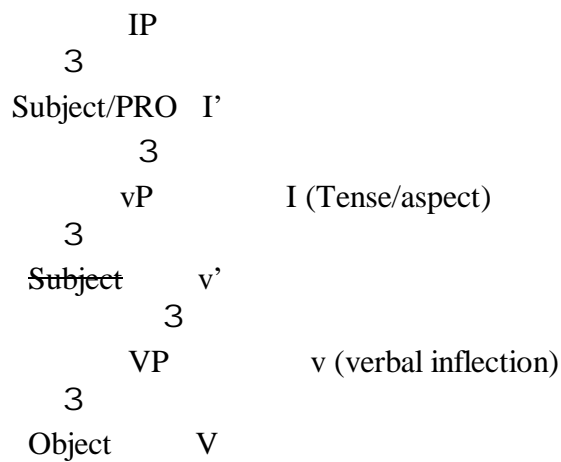
laRkooN-nee [PRO yah kitaab paRh-**nii** /*naa] caah-**ii**
 boys-ergative this book.fem.sg read-infinitive.**fs**/*ms want-pf-**fem**.sg.
 ‘The boys wanted [PRO to read this book].’ (cf. Davison 1991, Butt 1995, Bhatt 2005)

In this paper I will focus on the following issue (Question A) Is there an option which involved a truncated clause, without a projected subject?

(Wurmbrand 2003). Truncated clauses explain the option of agreement within a single clause (Bhatt 2005) A subsidiary question is Question B: Is the *-naa* inflection verbal, like non-finite tense in other languages, or is it nominal, affecting the category of the clausal structure?

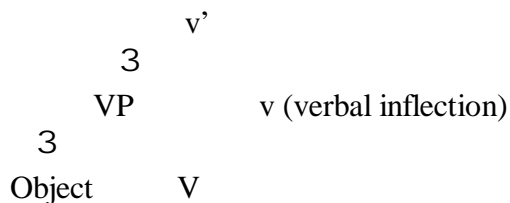
III. Before going to the arguments, I want to define three possible structures and category labels for infinitives. The (a) structure is an ordinary inflectional phrase, the same structure as an independent finite clause. There are projections for tense/aspect, which in the case of the infinitival is filled by the *-naa* affix. The subject is projected, whether it is pronounced or is controlled PRO, a null category. This follows from the requirements of tense inflection to have a specifier, by the Extended Projection Principle; the subject is attracted to the Specifier of IP position.

a. Full clause (Inflection phrase, IP)



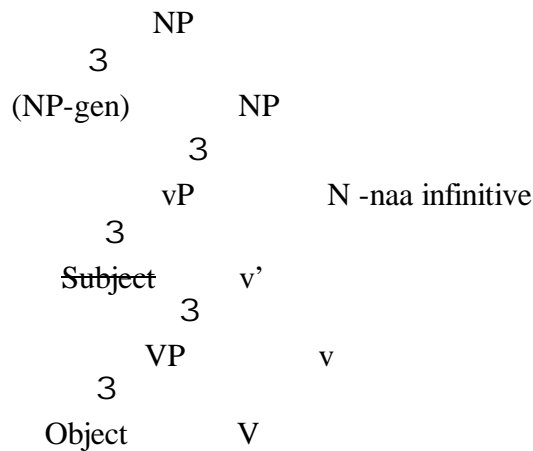
The truncated infinitive is represented as (b). It has some verbal inflection, but no inflectional tense head which attracts a subject. No subject is syntactically projected in this version even within the verbal projection. This structure is an approximation of what is proposed in Wurmbrand 2003, which is not specific about exactly how much clausal structure there is in this constituent.

b. Truncated clause



Finally, the (c) structure represents a gerund version of infinitive clauses, proposed in (Butt 1995). As in the case of gerunds in English, the structure is paradoxical. A verbal projection id

dominated by NP, whose specifier may be projected as a subject with genitive case.
 c. Noun Phrase, Gerund



IV. Evidence for the presence of a projected subject, whether it is null or overt.

The main thrust of this paper is to argue that in Hindi/Urdu, infinitives have the same structure as an independent clause, supporting structure (a), and to a lesser degree, structure (c). There is a syntactically projected subject, whether or not the subject is pronounced or is the controlled null PRO. I argue against the truncated infinitive (b) on the basis of three sets of data from Hindi/Urdu. I summarize these arguments as a preview.

Argument (1) is based on what I call the ‘Dative Effect’. Hindi/Urdu rejects embedded infinitive clauses if the subject of the embedded predicate would have dative marking if expressed. This restriction affects subjects, and is explained of the structure is (a) or (c), which projects a subject (PRO). The subject-truncated structure (b) can’t represent the ill-formed subject.

Argument (2) is concerned with reflexive binding, the linking of a reflexive pronoun with a subject antecedent. Reflexives in Hindi/Urdu must have a subject antecedent. Reflexives may be simplex or complex. Complex reflexives may have only a local subject antecedent. This is consistent with the structures (a), (c) which project a subject within the infinitive projection, but not with (b), which has no projected subject.

Argument (3) is based on the possibilities for optional infinitive agreement. The truncated structure (b) is offered as an explanation for this kind of ‘Long distance agreement’. If it is the object complement, then the whole combination is just one clause, allowing agreement throughout. But the ‘spreading’ kind of agreement is found in subject clauses with an overt subject, favoring structures (a) and (c).

V. The first argument is based on the fact that some predicates in Hindi/Urdu require dative case on the subject (Mohan 1994), in (6a) and (8a). Close synonyms require nominative or ergative subjects (6b), (8b).

6) a. [Dative subject]

shyaam-**koo** apnee-aap-par kroodh *aa-yaa*
 Shyam-dat self's-self-on anger.m.s come-pf.ms
 'Shyam got angry with himself.'

b. [Non-dative subject]

shyaam-**nee** apnee bhaaii-par kroodh *ki-yaa*
 Shyam-erg self's-brother-on anger.m.s do-pf.ms
 'Shyam got angry with self's brother.'

In Davison (to appear), I discuss a restriction on controlled PRO subjects. In sentences with an embedded null subject, sentences are robustly ungrammatical if the embedded subject is the subject of a dative-subject predicate (7a). (9a). If the infinitive verb is of the other type, then the sentence is grammatical, (7b), (9b).

7) Controlled infinitive --Matrix subject antecedent

a. * maiN [PRO apnee-aap-par kroodh *aa-naa*] nahiiN caah-tii huuN. [Dative]
 I self's-self-on anger come-inf not want-impf am
 'I don't want [PRO to get angry at self].'

b. shyaam [PRO apnee bhaaii-par kroodh *kar-naa*] nahiiN caah-taa hai [Non dative]
 Shyam self's-brother-on anger.m.s do-inf not want-impf is
 'Shyam does not want to get angry/express anger at self's brother.'

8) a. [Dative subject]

bacchooN-**koo** miThaayiaaN *mil-iiN*
 children-dat sweets receive-pf
 'The children got sweets.'

b. [Non dative subject]

bacchooN-**nee** miThaayiaaN *khaa-iiN*
 children-erg sweets receive-pf
 'The children got sweets.'

9) Controlled infinitive-- Matrix object antecedent

a. *maaN-nee bacchooN-koo [PRO miThaayiaaN *mil-nee*] nahiiN diiN/diyaa
 mother-erg children-dat sweets.f.pl get-inf not gave.-pf
 'Mother did not allow the children [PRO get sweets].'

- b. maaN-nee baccooN-koo [PRO miThaaiyaaN *khaa-nee*] nahiiN diiN/diyaa
 mother-erg children-dat sweets.f.pl eat-inf/*get-inf not gave.-pf
 ‘Mother did not allow the children [PRO eat sweets].’

I conclude that the good and bad controlled infinitives differ only in the verb. In (7a) and (9a), the infinitive verb assigns dative case to its subject. The otherwise identical sentences (7b) and (9b) are grammatical, as the infinitive verb does not assign dative case to its subject. There must therefore be a projected subject, even if it is null PRO. The truncated infinitive (b) has no projected subject, and so it should not be able to express a restriction on the case of the subject.

VI. Simplex and complex reflexives

Reflexives in Hindi/Urdu may be simple or compound, as in (10). The simple reflexive is a form of *apnaa*, which may be compounded with the form *aap*.

- 10) us-nee shiishee-meeN **apnee-koo/apnee aap-koo** dekkh-aa
 3s-erg mirror-in self’s-dat/ self’s self-dat see-pf.ms
 ‘He/she saw self in the mirror.’

Both types of reflexive have only subject antecedents, as noted by many authors including Gurtu 1990. If there are two possible antecedents, the reflexive is coindexed only with the subject, not the other argument, as in (11).

- 11) **maaN** baccce-koo kaisee **apnee-see/apnee aap-see** alag kar sakeegii?
 mother child-dat how self’s-from/self’s self-from separate be.able-fut.fs
 ‘How can the mother separate the child from self? (self = Mother/*child)
 (Davison 2000)

Subbarao 1984 and Gurtu 1990 show that simple reflexives may have both a local antecedent and a long distance antecedent, as in (12):

- 12) maaN-nee shyaam-koo [PRO **apnee-koo/apnee aap-koo** gumnaam patr bhee]nee]see
 mother-erg Shyam-dat self’s-dat / self’s self-dat anonymous letters send-inf-from
 manaa kiyaa
 forbidden do-pf
 ‘Mother forbade Shyam[PRO to send self anonymous letters].’ (Davison 2000)

The simple form **apnee-koo** is ambiguous between the local antecedent PRO, or Shyam, and the matrix subject ‘mother’, constituting the long-distance reading). The complex form **apnee-aap-koo** is unambiguous ; it is coindexed only with the local reading PRO or Shyam, not the matrix subject ‘Mother’.

Let us look at the ditransitive verb *dee-naa* in the sense of ‘to allow’, as in (13)

13) *dee-naa* ‘allow’

maaN-nee raadhaa-koo [*PRO apnee aap-koo* deekh-nee] nahiiN diyaa
mother-erg Radhaa-dat self’s self-dat see-inf-obl not give-pf.ms

‘Mother did not allow Radha [*PRO* to look at self].’ [‘self’ = Radha not Mother].’

This verb has a matrix indirect object which is coindexed with the embedded subject *PRO*, assuming that the embedded infinitive is a full clause with a projected subject. The indirect object *raadhaa-koo* can’t be the antecedent of *apnee aap-koo*, as it is not a subject. The matrix subject *maaN-nee* can’t be the antecedent of the compound *apnee aap-koo*, as it is not local. So there is only one possible local subject to be the antecedent of the complex reflexive, namely the null projected subject *PRO*. The truncated infinitive (b) does not provide the local subject in (14).

14) maaN-nee raadhaa-koo [*vP apnee aap-koo* deekh-nee] nahiiN diyaa
mother-erg Radhaa-dat self’s self-dat see-inf-obl not give-pf.ms

‘Mother did not allow Radha [*PRO* to look at self].’ [‘self’ = Radha not Mother].’

This analysis in (14) would predict that *apnee aap-koo* would be ambiguous in its antecedent, contrary to the facts.

VII. Argument 3: ‘Long-distance’ agreement.

Infinitives allow long distance agreement, involving both the embedded infinitive verb and the matrix verb (15). The embedded clause object *yah kitaab* is feminine singular. It can trigger feminine agreement on the infinitive and on the matrix clause. This is an option, contrasting with the default masculine singular forms in (16).

15) Long-distance Agreement (f.g.)

laRkooN-nee [PRO **yah kitaab** paRh-**nii** /*naa] caah-**ii**
boys-ergative this book.fem.sg read-infinitive.fs/*ms want-pf-fem.sg.
‘The boys wanted [*PRO* to read this book].’

16) Default Agreement (masculine sg. p.)

laRkooN-nee [PRO yah kitaab paRh-**naa**] caah-**aa**
boys-ergative this book.fem.sg read-infinitive.ms. want-pf-masc.sg.
‘The boys wanted [*PRO* to read this book].’

Wurmbrand 2003 proposes that many languages have the option of a full infinitive object clause (structure a) or a truncated infinitive (structure b)). Bhatt 2005 makes use of this option to explain (15), as a single clause, with agreement throughout. His analysis says that (15) has only a truncated infinitive, without a projected subject (17).

17)) Long-distance Agreement (f.g.)

laRkooN-nee [vP yah kitaab paRh-**nii** /*naa] caah-**ii**
boys-ergative this book.fem.sg read-infinitive.fs/*ms want-pf-fem.sg.
'The boys wanted [vP to read this book].'

(17) constitutes a single large clause, with obligatory agreement with the embedded object, as there are no other nominative NPs. Truncated infinitives lacking subjects occur in object position, and their lack of a subject is what ultimately is what allows Long-Distance Agreement

This proposal assumes that Long-distance Agreement is possible only with object infinitives lacking subjects. In fact this turns out not to be the case. There is an option for subject infinitives with subjects. Normally the subjects of infinitives have genitive case, which blocks agreement with the embedded and local verbs, as in (18). There is another option if the embedded clause subject has inanimate reference. In that case, the subject may be nominative. Without a postposition to block agreement, there is spreading agreement, as in (19)

18) Default agreement, genitive subject:

[basooN-kaa Diipoo-see nikal -**naa**] shuruu hu-aa
bus-f.pl-gen depot-from come.out inf.m beginning be-pf.m
'The buses began to leave the bus station.'

19) Long-distance agreement, nominative subject:

[baseeN Diipoo-see nikal -**nii**] shuruu hu-ii
bus-f.pl.nom depot-from come.out inf.f beginning be-pf.f
'The buses began to leave the bus station.' (K.V. Subbarao, p.c.)

The Bhatt 2005 analysis of long-distance or spreading agreement depends crucially there being a truncated infinitive. Nevertheless, long-distance agreement *is* possible with a full infinitive subject clause. So the presence of a subjectless truncated infinitive is not a necessary condition for Long Distance Agreement.

VIII. IP clauses or NPs? Is there a real difference?

I have argued that structure (b) does not provide a satisfactory account of properties of infinitives in Hindi/Urdu. Can we distinguish between structures (a) and (c), or really do they say the same thing, phrased differently, as in (20).

20) a. Structure (c) says the head of the infinitive/gerund is N, which has tense properties.

b. Structure (a) says that the head of the infinitive is Tense/Aspect, with nominal properties

Structure (c) is a relic of early ideas about phrases structure which take category as primary. We can rephrase (c) in more modern terms, using clusters of formal features to characterize the head

of a clause. These uninterpretable features must be valued by other constituents in a local domain. This idea is due to a proposal from 1988 by C-T. James Huang, which says that Inflection has category features like lexical items. Gerund/inflection has the [N] feature, as in (21), predicting genitive or nominative case on the specifier, and the possibility of agreement. Finite tense has the [V] feature, as in (22), with similar properties except for the specific case on the specifier, and the kind of agreement features which are realized.

21) INFL[N] = [Tense, Case, (Agreement)] features, Case is realized as Genitive or Nominative.

22) INFL[V] = [Tense, Case, Agreement] features, Case is realized as Nominative or Ergative.

IX. To sum up briefly, I have proposed three structures for infinitives in Hindi/Urdu, which really boil down to two: a full clausal projection with a subject position and tense/aspect inflection, and a truncated infinitive without a projected subject. I have offered arguments that the truncated infinitive fails to account for the Dative effect, the properties of complex reflexives, and for Long-Distance Agreement in subject infinitives.

References

- Bhatt, Rajesh K. (2005) Long-distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory* 23, 757-807.
- Butt, M. (1995) The structure of complex predicates in Urdu. Stanford: Center for the Study of Language and Information
- Cole, Peter, Gabriella Hermon and C.T. James Huang (eds.)(2000) Long distance anaphors. Irvine:Academic Press.
- Davison, Alice (1991) Feature percolation and agreement in Hindi/Urdu. Paper presented at the South Asian conference, Madison Wisconsin.
- Davison, Alice (2000) Long distance anaphors in Hindi. In Cole et al.
- Davison, Alice (to appear) A case restriction on control: implications for movement. *Journal of South Asian Linguistics*.
- Gurtu, Madhu (1990) *Anaphoric relations in Hindi and English*. Delhi
- Mohanan, Tara W (1994) *Argument structure in Hindi*. Stanford: CSLI Publications.
- Subbarao, K.V. (1984) *NP complementation in Hindi*. Delhi: Academic Publications.
- Wurmbrand, S. (2003) *Infinitives: Restructuring and clause structures*. Berlin: W. deGruyter