TARA MOHANAN

WORDHOOD AND LEXICALITY: NOUN INCORPORATION IN HINDI*

This paper examines the syntactic, semantic, morphological, and phonological properties of a certain type of noun-verb sequence in Hindi, and argues that it is an instance of noun incorporation. The sequence must be analysed as a lexical category; yet paradoxically, verb agreement and negation show that this noun is on par with a syntactically independent argument. The paper proposes a solution to this dual behaviour of the noun by recognizing grammatical categories and grammatical functions as belonging to distinct but co-present dimensions of syntactic representation. This factorization of dimensions yields an account of the facts of Hindi Incorporation which are not amenable to analysis in terms of head movement (Baker 1988) or coanalysis (Sadock 1991). When combined with the idea of different notions of wordhood, the dual representation leads to a distinction between categorial word and functional word, central to the untangling of numerous issues surrounding lexicality.

This paper addresses a problem posed by a construction in Hindi that has not been recognized in the literature on Hindi syntax, and which comes under the rubric of Noun Incorporation. Noun Incorporation is the phenomenon of a noun combining with a verb to form a single morphological unit, while still retaining its syntactic status (Mithun 1984; Allen, Gardiner and Frantz 1984; Hopper and Thompson 1984; Sadock 1980, 1985, 1986, 1991; Knecht 1986; Baker 1985, 1988; Di Sciullo and Williams 1987; Rosen 1989; among others). If we assume that the verb and one of its arguments form a lexical unit (N + V) in the Hindi Noun Incorporation Construction (henceforth NIC), various semantic, syntactic, morphological, and phonological characteristics of the construction follow as a natural consequence. However, the verb can agree with that argument when agreement with the subject is not possible. In order to account for verb agreement in NIC,

* The analysis in this paper began as a response to Paul Kiparsky's persistent and incisive questioning. Joan Bresnan and K. P. Mohanan made enormous contributions to the paper during various stages of its growth. Comments from N. S. Prabhu, Carol Georgopoulos, and three anonymous reviewers have resulted in substantial improvements. I thank them all. The paper has also benefited from presentations at the Stanford Linguistics Colloquium in May 1990, and at the panel on Agreement at the 20th Annual Conference on South Asia held at the University of Wisconsin, Madison in November 1991. This study was done in part with support from the Center for the Study of Language and Information, Stanford.

The judgements reported in this paper are largely my own. I am grateful to Akhil Gupta, Purnima Mankekar, Ravi Oswal, Rajeshwari Pandharipande, and Alka Warrier, who have generously offered their own judgements, increasing my confidence in my own judgements. I thank Rajendra Singh for making me aware of the extent to which speaker judgements can differ.

Natural Language and Linguistic Theory **13**: 75–134, 1995.

the noun in the N + V compound must be analysed syntactically as an argument on par with the other arguments of the verb.

This situation raises two problems. First, evidence from one set of facts shows that the incorporation must be lexical, while another set of facts shows that it is not lexical. Second, a verb in Hindi agrees with one of its sisters, but in NIC, the verb agrees with the incorporated noun, a daughter. The intriguing question is, if the controller of verb agreement must be a sister of the verb, how can the verb agree with its daughter?

These puzzles resolve themselves if, in the architecture of grammar, we factor apart the internal organization of the verb's arguments and their grammatical functions on the one hand, and their category structure on the other. The factorization of different types of information into separate dimensions of organization throws open the possibility of different constituency relations at different dimensions of representation. The facts of negation in Hindi support such factorization.

The paper is organized as follows. Section 1 introduces NIC in Hindi in terms of a set of correlations associated with object case marking, and outlines some theoretical assumptions used in the account of the phenomenon. Section 2 argues for the treatment of NIC as a lexical unit $[N + V]_V$. Section 3 examines the facts of verb agreement, and shows that agreement in NIC cannot be lexical. Section 4 discusses some of the alternative analyses in detail. Section 5 deals with negative placement, which provides support for this account of NI. Section 6 shows how the account distinguishes different notions of wordhood in terms of well-motivated dimensions of structure, and yields a promising way of resolving the numerous debates on issues of lexicality.

1. Introduction

This section presents the puzzling asymmetries in Hindi which are analysed as Noun Incorporation, and spells out the theoretical assumptions relevant for the analysis.

1.1. *The Puzzle*

In order to formulate the puzzle, let me first give some well-known facts of object case marking in Hindi. Animate primary objects in Hindi bear the accusative case clitic *-ko*:[1]

[1] The following abbreviations are used in the glosses:

N: Nominative NF: Nonfinite

(1)a. ilaa anil-ko uṭ^{h}aaegii.
Ila-N Anil-A lift-FU
Ila will lift up Anil.

b. *ilaa *anil* uṭ^{h}aaegii.
Anil-N

Inanimate objects have a choice between accusative (ACC) case, and nominative (NOM) case which is morphologically unmarked.[2] If ACC, an inanimate object must be definite.[3]

(2)a. ilaa-ne haar-ko uṭ^{h}aayaa.
Ila-E necklace-A lift-PA
Ila lifted up the necklace.

b. ilaa-ne *haar* uṭ^{h}aayaa.
necklace-N
Ila lifted up the/a necklace.

Thus, an ACC object must be animate or definite. Now consider (3), with an ACC animate object, and (4), where the same object is NOM:

(3) ilaa baccõ-ko k^{h}ojtii rahtii hai.
Ila-N children-A search-HAB PROG be-PR
Ila keeps searching for the/some children.

E:	Ergative	PA:	Past
A:	Accusative	PR:	Present
D:	Dative	FU:	Future
L:	Locative	HAB:	Habitual
I:	Instrumental	PERF:	Perfective
G:	Genitive		
M:	Masculine	SG:	Singular
F:	Feminine	PL:	Plural

The convention adopted in the word glosses is as follows. When a form uniquely expresses one member in a contrasting pair, the member is included in the gloss. Thus, the specification of number is indicated in the gloss if the singular and plural forms (or the NOM and NONNOM case forms) of a noun are distinct on the surface. But the specification is omitted from the gloss if the forms are identical.

[2] I analyse subjects and objects without overt case marking as bearing nominative case. For a defense of this position, see T. Mohanan (1993).

[3] For details, see Srivastava (1969), McGregor (1972), T. Mohanan (1990, 1993) and the references therein.

(4) ilaa bacce kʰojtii rahtii hai.
Ila-N children-N search-HAB PROG be-PR
Ila keeps children-searching (i.e. performing the act of searching for children).

The only visible difference between (3) and (4) is the case marking on the object. Yet, their interpretations as indicated in the sentence glosses are different. Suppose we refer to the interpretation of (3) as "reading A", and that of (4) as "reading B". How do we account for the correlation between ACC/NOM case of animate objects and readings A/B?

To extend the puzzle further, consider the following contrasts. The animate objects in (5)–(9) are ACC in (a), and NOM in (b):

(5)a. ilaa baccõ-ko hamešaa/har jagah kʰojtii
Ila-N children-A always every place search-HAB
rahtii hai.
PROG be-PR
Ila is searching for the/some children all the time/everywhere.

b. *ilaa bacce hamešaa/har jagah kʰojtii
Ila-N children-N always every place search-HAB
rahtii hai.
PROG be-PR

(6)a. ilaa cʰoṭe baccõ-ko kʰojtii rahtii hai.
Ila-N small children-A search-HAB PROG be-PR
Ila keeps searching for the/some small children.

b. *ilaa cʰoṭe bacce kʰojtii rahtii hai.
Ila-N small children-N search-HAB PROG be-PR

(7)a. ilaa kis-ko kʰojtii rahtii hai?
Ila-N who-A search-HAB PROG be-PR
Who does Ila keep searching for?

b. *ilaa kaun kʰojtii rahtii hai?
Ila-N who-N search-HAB PROG be-PR

(8)a. ilaa laṛkõ aur laṛkiyõ-ko kʰojtii rahtii hai.
Ila-N boys- and girls-A search-HAB PROG be-PR
Ila keeps searching for the/some boys and girls.

b. *ilaa laṛke aur laṛkiyãã kʰojtii rahtii hai.
Ila-N boys-N and girls-N search-HAB PROG be-PR

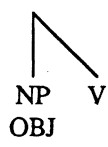
(9)a. ilaa baccõ-ko kʰotii aur kʰojtii rahtii hai.
Ila-N children-A lose-HAB and search-HAB PROG be-PR
Ila keeps losing and searching for the/some children.

b. *ilaa bacce kʰotii aur kʰojtii rahtii hai.
Ila-N children-N lose-HAB and search-HAB PROG be-PR.

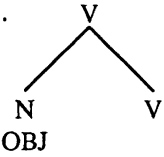
The only visible difference between (a) and (b) in (5)–(9), again, is that the object is ACC in (a) and NOM in (b). Given that (4) is acceptable, the second problem is: when an animate object is nominative, it cannot be (i) separated from the verb ((5b)), (ii) modified ((6b)), (iii) questioned ((7b)), or (iv) conjoined ((8b)). Nor can the verb be conjoined ((9b)).[4]

The goal of this paper is to provide an account for the phenomenon partially illustrated in (4). I will argue that the NOM object in (4) forms a morphological unit with the verb, and has the structure in (10b), while the ACC object and the verb in (3) have the structure in (10a):

(10)a.



b.



1.2. *Theoretical Background*

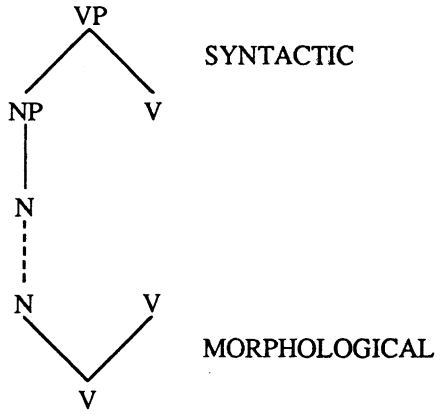
The representation in (10b) contains the information that the noun forms a lexical category with the verb, and is at the same time the object of the verb. (10b) represents the substance of what has been called Noun Incorporation (henceforth NI) in syntactic theory. I use the term NI, following Sadock (1980, 1991), to refer to the phenomenon of a noun stem exhibiting dual behaviour: it is a syntactic argument of a verb, but morphologically part of that verb.[5]

[4] There are dialects of Hindi in which NOM animate objects may be more freely available in such sentences. See, for instance, Verma (1971, p. 104).

[5] That is, NI does not refer to a specific analysis of the phenomenon, e.g., the movement of a lexical head noun into a verb, as in Baker (1985, 1988). For a historical perspective on NI, see Sadock (1991: pp. 78–100).

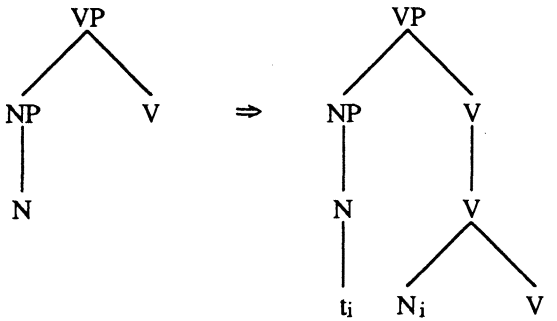
Most analyses of NI, including the one proposed in this paper, agree that there must be some representation in which the verb and the incorporated noun together form a lexical (X^0) category ($[N + V]_{V^0}$). Sadock (1980 and subsequent work) and Baker (1985, 1988) additionally recognise the synactic argument status of the incorporated nominal, which they represent as an NP sister of the verb. The N and the NP are related in Sadock as parallel independent representations:

(11)



In Baker, they are related in a single representation through the movement of the head of the NP into the verb, leaving a trace:

(12)



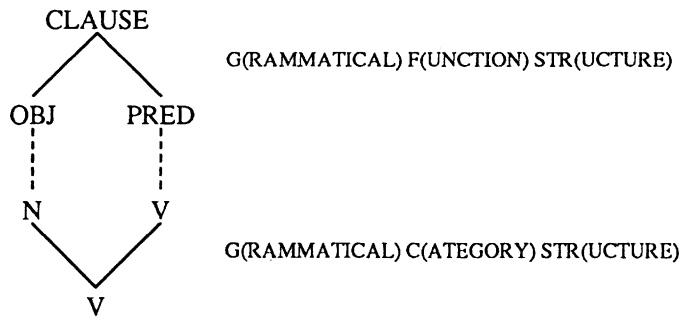
The parallel representations in (11) and the output of head movement in (12) express the same information as in (10b): the noun forms a lexical category with the verb, but is also a syntactic argument of the verb. The dotted line in (11) which connects the two types of information corresponds to the coindexed trace in (12).

In contrast to Sadock and Baker, Mithun (1984), Di Sciullo and Williams (1987), and most recently, Rosen (1989) propose that NI is a purely lexical phenomenon, without any syntactic consequences.

My account of incorporation is essentially in the lexicalist tradition of Lexical Functional Grammar and Lexical Phonology and Morphology, in that it claims contra Baker that incorporation in Hindi is part of the lexical module. Within this conception of grammar, the lexical module, where words are formed, is distinct from the phrasal module, where words are concatenated to form larger units.[6] The term *phrasal* in this paper has the same meaning as *postlexical* in Lexical Phonology. It has no X-bar theoretic connotations.

Unlike Rosen and others, and like Sadock and Baker, this analysis also claims that there must be some representation in which the incorporated noun is syntactically independent, at least for languages like Hindi. Like Sadock, it recognizes NI as simultaneously morphological and syntactic. It differs from Sadock in that while the two representations in (11) are those of morphology and syntax, the two relevant representations of my proposal are those of grammatical categories on the one hand and grammatical functions on the other:

(13)



2. Lexicality of Noun Incorporation in Hindi

In this section, I propose a representation of the categorial structure for the Hindi NIC, with syntactic, semantic, phonological and morphological evidence for the analysis of the N + V sequence as a lexical category. I argue, further, that the incorporation is part of the lexical module.

[6] I hesitate to use the term *syntactic* instead of *phrasal*, because of the implication that no word internal information is syntactic.

2.1. *A Partial Analysis: N + V in NI as a Lexical Category*

2.1.1. *Overt Case Marking*

As we saw in (2), an inanimate object in Hindi may be either NOM or ACC. Let us assume that an animate object is assigned ACC case if it is phrasal (10a); it is NOM if and only if it is part of a lexical category (10b).

The inability of the object in structure (10b) to be ACC follows if we assume that case markers in Hindi are clitics that yield phrasal categories, not inflections that yield lexical categories. Evidence for this assumption comes from conjoining. As I will show in section 2.1.5, lexical categories cannot be conjoined in Hindi. Yet, case clitics can be attached to conjoined nominals:

(14)a. madraas aur kalkatte mẽ
Madras and Calcutta L
in Madras and Calcutta.

b. auratõ aur baccõ se
women and children I
from/by women and children.

Therefore, we must conclude that case clitics are attached to phrasal categories.[7] The N in (10b) occurs within a lexical category. Since ACC involves a case clitic, it cannot occur in (10b), unlike NOM.

We saw that (3) with an ACC object has reading A, while (4) with a NOM object has reading B. Given our assumptions above, we can account for this correlation between case marking and meanings by assuming that reading B is licenced by the lexical structure in (10b), and reading A by phrasal structrure in (10a). I will now go on to show how the above analysis correctly predicts the behaviour of inanimate objects in Hindi.

Given the result that structure (10b) does not allow ACC case, and the assumption that reading B is sanctioned only by (10b), it would follow that inanimate objects can have either reading A or reading B if NOM, but only reading A if ACC. This prediction is borne out by the following data:

[7] This is in contrast with a language like Malayalam, where case is indicated by affixes rather than clitics. In a conjoined structure, for example, the affix is attached to every conjunct, rather than to the entire conjoined structure.

(15)a. anil kitaabẽ becegaa.
Anil-N(M) book-N.PL(F) sell-FU.M.SG
A. Anil will sell (the) books.
B. Anil will do book-selling.

b. anil kitaabõ-ko becegaa.
book-A.PL(F)
A. Anil will sell the books.
B. *

(16)a. anil-ne kitaabẽ becĩĩ.
Anil-E(M) book-N.PL(F) sell-PERF.F.PL
A. Anil sold (the) books.
B. Anil did book-selling.

b. anil-ne kitaabõ-ko becaa.
book-A.PL(F)
A. Anil sold the books.
B. *

(17)a. raam-ne lakḍii kaaṭii.
Ram-E(M) wood-N(F) cut-PERF.F.SG
A. Ram cut (the) wood.
B. Ram did wood-cutting.

b. raam-ne lakḍii-ko kaaṭaa.
wood-A(F)
A. Ram cut the wood.
B. *

The objects in (a) are NOM, and the sentences have two interpretations,

reading A and reading B.[8] The objects in (b) are ACC, and the sentences allow only reading A.[9]

2.1.2. *Adjacency*

As shown by (5b), the animate object must be adjacent to the verb if it is NOM. This follows from the assumption that animate objects can be

---

[8] Given below are two more examples of inanimate nouns as NOM objects of sentences:

(i) anil-ne k^{h}aanaa pakaayaa.
Anil-E(M) food-N(M) cook-PERF.M.SG
A. Anil cooked food.
B. Anil did food-cooking.

(ii) raam kapḍe siitaa hai.
Ram-N(M) clothes-N(M) sew-HAB.M.SG be-PRES
A. Ram sews/tailors clothes.
B. Ram does clothes-sewing. (=Ram is a tailor.)

These sentences also have two interpretations. The different interpretations have corresponding differences in word stress and word melody patterns. We will return shortly to these correspondences.

[9] The case restriction on the N in (10) is not on the case feature (NOM/ACC) but on the presence of an overt case clitic. Other than NOM nouns, locative destinations are the only cliticless arguments in Hindi. The locatives in (i) and (ii) have no clitic; those in (iii) and (iv) bear locative case clitics:

(i) raam-ne apne beṭe-ko šahar b^{h}ejaa.
Ram-E self-G son-A city send-PERF
A. Ram sent his son to the city.
B. Ram city-sent his son.

(ii) baccaa skuul gayaa.
child-N school go-PERF
A. The child went to school.
B. The child school-went.

(iii) raam-ne apne beṭe-ko šahar-mẽ rak^{h}aa.
Ram-E self-G son-A city-L(IN) keep-PERF
A. Ram kept his son in the city.
B. *Ram 'city-kept' his son.

(iv) baccaa kursii-par baiṭ^{h}taa hai.
child-N chair-L(on) sit-HAB be-PR
A. The child sits on a chair.
B. *The child 'chair-sits'.

Like (15a) and so on, (i) and (ii) allow reading B: I assume, then, that cliticless locatives can be part of (10b). In contrast, (iii) and (iv), in which the locatives bear case clitics, do not allow reading B. I will not elaborate further on locatives that can be part of (10b).

NOM only in structure (10b). Since the noun and the verb form a single lexical category in (10b), the noun cannot be separated from the verb.

This analysis predicts that reading B will not be available to inanimate objects if they are not adjacent to the verb. This prediction is borne out by the following data:

(18)a. anil kitaabẽ hamešaa becegaa.
Anil-N(M) book-N.PL(F) always sell-FU.M.SG
A. Anil will always sell books.
B. *Anil will always do book-selling.

Like other South Asian languages, Hindi is a 'free word order' language in that the predicate and its dependents in a clause are free with respect to one another in their relative order. Thus, (18b, c) are acceptable versions of (15a):[10]

(18)b. kitaabẽ anil becegaa.
book-N.PL(F) Anil-N(M) sell-FU.M
A. Anil will sell the books.
B. *Anil will do book-selling.

c. anil kitaabẽ baazaar-mẽ becegaa.
Anil-N(M) book-N.PL(F) market-L sell-FU.M
A. Anil will sell the books in the market.
B. *Anil will do book-selling in the market.

As our analysis predicts, (18b) and (18c) disallow reading B. As predicted, the result of scrambling the animate NOM object away from the verb in (10b) is ungrammatical.

2.1.3. *Modifier Stranding*

Since the left member in structure (10b) is an N, not an NP, it follows that the phrasal unit [modifier + N] cannot participate in structure (10b). This explains why animate objects cannot take modifiers if they are NOM (6b). This analysis predicts that inanimate objects cannot have reading B

[10] Entities or sequences that are not in their canonical order exhibit certain definiteness effects in Hindi. The definiteness of the object in (18) follows from this condition, which I will not discuss further in this paper.

if they contain modifiers. This prediction is confirmed by the following data:

(19)a. anil-ne puraanii kitaabẽ becĩĩ.
Anil-E(M) old-F book-N.PL(F) sell-PERF.F
A. Anil sold a/the old books.
B. *Anil did the selling of old books.

b. anil-ne acc^{h}aa k^{h}aanaa pakaayaa.
Anil-E good food-N(M) cook-PERF.M
A. Anil cooked good food.
B. *Anil did the cooking of good food.[11]

c. raam baccõ-ke kapḍe siitaa hai.
Ram-N(M) children's clothes-N(M) sew-HAB.M be-PRES
A. Ram sews/tailors children's clothes.
B. *Ram does children's clothes-sewing.

The non-cooccurrence of object modification and reading B further leads to the assumption that there is no representation in which the N in (10b) is associated with an NP.[12]

[11] For some speakers, a sentence like (i) below, similar to (i) in note 9, allows reading B:

(i) anil acc^{h}aa k^{h}aanaa pakaataa hai.
Anil-N good food-N(M) cook-HAB.M be-PRES
A. Anil cooks good food.
B. Anil cooks well.

The reason for this ambiguity despite modification is that the modifier *acc^{h}aa* itself is ambiguous between the adjective "good" and the adverb "well". Co-occurring with reading B is the latter meaning, modifying the entire verb complex and not just the noun.

[12] An anonymous reviewer has expressed scepticism about this claim, and points out that even though [[green car] driver] in English is unacceptable in ordinary situations, it becomes acceptable in a situation where ([green car] is nameworthy, for example, under a gas-rationing scheme where one's permission to fill up depended on the colour of one's vehicle. However, rather than assuming that the NP *green car* can form a compound with another noun if the phrase is nameworthy, I suggest that *green car* can become a compound noun in English if nameworthy (like *blackbird*). If Hindi allowed compound nouns of the form [A N], the analysis in this paper would predict noun incorporation constructions of the form [[A N] V]$_{V}$. However, I have not been able to find any [A N] compounds in Hindi.

The issue of nameworthiness is discussed in greater detail in section 2.2.

2.1.4. *Wh-words and Pronouns*

Given the widely accepted assumption that *wh-* words and pronouns are phrasal categories (maximal projections), they cannot participate in structure (10b).[13] This explains why the animate NOM object cannot be questioned (7b). It also predicts that an inanimate object that is either a *wh-* form or a pronoun cannot have reading B. This prediction is borne out by (20a) and (20b):

(20)a. anil kyaa becegaa?
Anil-N(M) what sell-FU.M.SG
A. What will Anil sell?
B. *What-selling will Anil do?

b. anil yah becegaa.
Anil-N(M) this sell-FU.M.SG
A. Anil will sell this.
B. *Anil will do this-selling.

2.1.5. *Conjoining*

Consider the examples in (21), which involve conjoining:

(21)a. laal kitaabẽ aur laal kursiyãã
red book-PL.N and red chair-PL.N
red books and red chairs

b. laal kitaabẽ aur kursiyãã
red book-PL.N and chair-PL.N
red books and chairs (i) chairs and red books
(ii) *red books and red chairs.

For the interpretation in (i), the structure of (21b) is $_{\bar{N}}[_{\bar{N}}$[red books] and $_{\bar{N}}$[chairs]]. For the interpretation in (ii), it must be $_{\bar{N}}[_{A}$[red] $_{N}[_{N}$[books]

[13] Under a DP analysis (Abney 1987), a *wh-* word or pronoun is not necessarily a maximal projection. However, it is still not a lexical category, as required by (10b).

and $_N$[chairs]]]].[14] The absence of this reading is accounted for if we assume that only phrasal categories can be conjoined in Hindi.[15]

Given this independently motivated assumption, it follows that neither the N nor the V in (10b) can be conjoined. This explains why (8b) and (9b) are ungrammatical. Since the relevant expression in each example involves conjoining, that expression cannot be a lexical category. However, only a lexical category can occur in (10b). Hence, (8b) and (9b) cannot be analysed in terms of (10b). However, an animate object can be NOM only if it is part of the structure in (10b). Hence, a structure with a conjoined animate NOM object is ungrammatical.

---

[14] The data below illustrates this further:

(i) bahut baḍii kitaabẽ
a lot large book-PL.N
(a) very large books
(b) many large books

(ii) bahut bʰaarii kitaabẽ
a lot heavy book-PL.N
(a) very heavy books
(b) many heavy books

(iii) bahut baḍii aur bʰaarii kitaabẽ
a lot large and heavy book-PL.N
(a) [[very large] and [heavy]] books
(b) many [large and heavy] books
(c) *very large and very heavy books

In (i) and (ii), *bahut* may function either as an intensifier modifying the adjective as in (a), or as a quantifier modifying the noun as in (b). In (iii), *bahut* may modify either the noun as in (a), or the first adjective of the conjoined structure as in (b). However, the unacceptability of interpretation (c) shows that it cannot modify the conjoined structure itself.

[15] That this is true for verbal units is illustrated by the following contrast:

(i)a. baccaa boltaa gayaa aur rotaa gayaa.
child-N speak-HAB go-PERF and cry-HAB goPERF
The child kept talking and kept weeping.

b. baccaa boltaa aur rotaa gayaa.
The child kept talking and weeping.

(ii)a. baccaa bol rahaa tʰaa aur ro rahaa tʰaa.
child-N speak stay-PERF be-PA and cry stay-PERF be-PA
The child was talking and was weeping.

b. *baccaa bol aur ro rahaa tʰaa.

I claim that the ungrammaticality of (iib) is the result of conjoining two V's. Arguing for this position calls for a detailed discussion of verbal morphology and the auxiliary system, which is beyond the scope of this paper.

This analysis also predicts that if the verb or its inanimate object involve conjoining, reading B will be ruled out. The prediction is confirmed by (22a, b):

(22)a. anil haat^{h}ii aur g^{h}oḍe bectaa hai.
Anil-N elephants-N and horses-N sell-HAB be-PR

A. Anil sells elephants and horses.
B. *Anil does elephant- and horse-selling.

b. anil g^{h}oḍe k^{h}ariidtaa aur bectaa hai.
Anil-N horses-N buy-HAB and sell-HAB be-PR

A. Anil buys and sells horses.
B. *Anil does horse-buying and -selling.

2.1.6. *Gapping*

Consider the sentences below that involve coordination:

(23)a. anil g^{h}oḍe k^{h}ariidtaa hai aur raam g^{h}oḍe
Anil-N horses-N buy-HAB be-PR and Ram-N horses-N
bectaa hai.
sell-HAB be-PR

A. Anil buys horses and Ram sells horses.
B. Anil does horse-buying and Ram does horse-selling.

b. anil g^{h}oḍe$_{i}$ k^{h}ariidtaa hai aur raam —$_{i}$ bectaa
Anil-N horses-N buy-HAB be-PR and Ram-N sell-HAB
hai.
be-PR

A. Anil buys and Ram sells horses.
B. *Anil does horse-buying and Ram does — -selling.

c. anil g^{h}oḍe —$_{i}$ aur raam haat^{h}ii bectaa$_{i}$ hai.
Anil-N horses-N and Ram-N elephants-N sell-HAB be-PR

A. Anil buys horses and Ram elephants.
B. *Anil does horse- — and Ram elephant-selling.

Each of the conjuncts in (23a) may independently have reading B. However, reading B is unavailable in (23b) and (23c). In other words, if

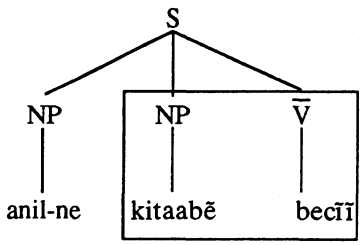
either the object or the verb is gapped in a coordination construction, the sentence disallows reading B.

This unavailability of reading B follows from (10b) if we assume the *Lexical Integrity Hypothesis* which "prohibits syntactic reorderings into or out of lexical categories" (Bresnan 1982b, 54).[16] This condition prevents gapping into a lexical category, since gapping is a phrasal phenomenon. In addition, the condition prevents part of a lexical category from also being the direct daughter of a phrasal category (through movement or coanalysis), a result that is supported by the absence of modifier stranding described in section 2.1.3.

2.1.7. *Summary: The Category Structure of Hindi NI*

In sum, the asymmetries reviewed so far are accounted for by assuming that Hindi has a construction in which a noun and a verb together form a single lexical category. When a noun-verb sequence in a sentence such as (16a) has reading A, it has the syntactic structure given in (24); when it has reading B, it has the syntactic structure in (25):[17]

(24)

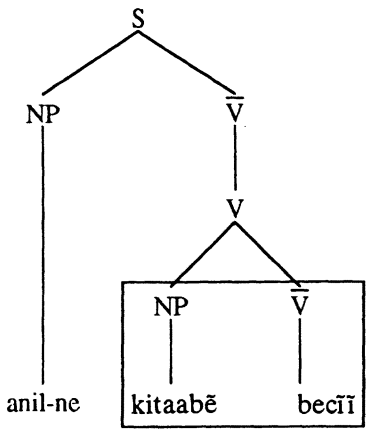


(Anil sold books.)

[16] Observe that the Lexical Integrity Principle is stated specifically on category information, and will not extend, for instance, to grammatical function information. The significance of this point will become clear later in the paper. Also see Simpson (1983, 1991, pp. 44–45) for the idea within LFG that the principle holds only on constituent structure, not on functional structure, and Bresnan and Mchombo (1993) for further support of the Lexical Integrity Principle.

[17] The V̄ in these structures represents a phrasal category, whether or not it is a maximal projection. The V̄ node could as well be labelled as VP without any empirical difference. However, I avoid using the label VP, because a VP is characteristically viewed as a node that contains the verb and at least all its non-subject complements. I assume that in Hindi non-subject complements of a verb, including object NP's, are not dominated by V̄. I will not defend this assumption here because it is not directly relevant to this paper.

(25)



(Anil did book-selling.)

At the outset of the paper, I said that NI refers to the phenomenon of a noun stem (a) being part of a morphological unit and yet (b) having syntactic independence. In the preceding discussion, I gave syntactic evidence for (a). Although I have not demonstrated the syntactic independence of the noun yet, I claim that the structure in (25) is one of NI. In what follows, I will call this construction the NI construction (or NIC) in Hindi. NI in most languages discussed in the literature seem to involve an overtly visible difference between the NI construction and the construction without an incorporated nominal. In Hindi, an NI construction has an external appearance identical to that of its unincorporated counterpart.

2.2. *Semantic and Pragmatic Correlates of NI*

Let us first try to sharpen the distinction that I have appealed to so far between readings A and B. Consider the truth conditions accompanying the two readings of (26a):

(26)a. mohan c^{h}uṭṭiyõ-me vækyum kliinar bectaa t^{h}aa.
Mohan-N holidays-in *vacuum cleaner-N sell-HAB be-PA*

A. Mohan was selling vacuum cleaners during the holidays.
B. Mohan was doing vacuum cleaner-selling during the holidays.

b. usne do mahine-me ek b^{h}ii vækyum kliinar nahĩĩ
he-ERG two month-in *one even vacuum cleaner-N not*
becii.
sell-PA
He didn't sell even one vacuum cleaner in two months.

(26b) and reading A of (26a) cannot both be true at the same time. However, (26b) and reading B of (26a) are not logically contradictory: they can be simultaneously true.[18] This contrast justifies distinguishing the two readings.

Recall our assumption that reading B is sanctioned by structure (10b). In addition to having different truth conditions, reading B requires that the N in the N + V sequence be generic or 'non-referential': it can refer only to the class of entities denoted by the noun, not to the individual members of the class.[19] This is illustrated by the following examples:

(27) anil-ne tiin saal kapḍe dʰoe.
Anil-E three year-N clothes-PL.N wash-PERF

A. Anil washed clothes for three years.
B. Anil did clothes-washing for three years.

(28)a. kapḍe bahut gande ho gaye tʰe.
clothes-N-PL a lot dirty-PL become go-PERF-PL be-PA-PL

The clothes had become very dirty.

b. anil-ne tiin gʰaṇṭe kapḍe dʰoe.
Anil-E three hour-N-PL clothes-N-PL wash-PERF

A. Anil washed clothes for three hours.
B. *Anil did clothes-washing for three hours.

(27) allows both readings. In (28), however, because of prior mention in (a), the object in (b) is definite, and reading B is impossible. We would then predict that an animate NOM object with prior mention in the discourse will be ungrammatical. This is borne out by (29):

(29)a. bacce sab kʰo gaye tʰe.
child-N-PL all lose go-PERF-PL be-PA-PL

The children had all got lost.

---

[18] I thank an NLLT reviewer for pointing my attention to this possible difference.

[19] Genericity and non-referentiality have sometimes been seen as a characteristic of incorporated nominals. However, languages with more widespread, and sometimes obligatory, incorporation show that an incorporated noun is not necessarily accompanied by genericity or nonspecificity (e.g., Mohawk (Mithun 1984), Nahuatl (Merlan 1976), Southern Tiwa (Allen, Gardiner, and Frantz 1984) and so on). In Eskimo (Sadock 1980), an incorporated noun may be specific, or even definite, and may introduce a new topic.

b. *ilaa-ne tiin gʰaṇṭe bacce kʰoje.
Ila-E three hour-N-PL child-N-PL search-PERF

In the context of (29a), the object in (29b) is definite. Therefore, reading B is unavailable. Consequently, structure (10b) is unavailable. NOM case on an animate phrasal object is ungrammatical.

Finally, the activity or process referred to by NI must be 'salient' or 'nameworthy' in what Hale and Keyser (1991, p. 13) call the 'cultural encyclopedia' of the language users.[20] Consider the contrast between A and B in (30a–c):[21]

(30)	*A*	*B*
a.	gʰaas kaaṭnaa / becnaa	dekʰnaa
	grass cutting selling	*seeing*
	A. cutting/selling grass	seeing grass
	B. grass-cutting/-selling	# grass-seeing
b.	kapḍe becnaa / dʰonaa / siinaa	pʰaaḍnaa/ pahanna
	clothes selling washing sewing	*tearing wearing*
	A. selling/washing/sewing clothes	tearing/wearing clothes
	B. clothes-selling/-washing/-sewing	# clothes-tearing/-wearing
c.	kitaab likʰnaa / paḍʰnaa/ becnaa	denaa / uṭʰaanaa
	book writing reading selling	*giving lifting*
	A. writing/reading/selling books	giving/lifting books
	B. book-writing/-reading/-selling	# book-giving/-lifting

Take *gʰaas kaaṭnaa* "cutting grass"/"grass-cutting" and *gʰaas dekʰnaa* "seeing grass"/ #"grass-seeing" in (30a). In most societies, cutting grass is a salient activity in the conceptual system, while seeing grass is not. Hence the former, but not the latter, can be an NIC in Hindi. If we adopt the suggestion that nameworthiness is a property of lexical items (Hale and Keyser 1991, p. 13), the fact that the activity referred to by the Hindi NIC must be salient or nameworthy in the conceptual system will follow from the assumption that the N + V in NI is concatenated in the lexical module.

[20] Bauer (1983, pp. 86–7) refers to this property as the "nameability" condition on lexical items.

2.3. *Phonological and Morphological Evidence: NI in the Lexical Module*

I now turn to phonological and morphological evidence that further supports the representation of NI as a lexical category (10b), and the claim that NI is formed in the lexical module, rather than in the phrasal module.

2.3.1. *Phonology*

Consider the following scenario. A primary school teacher says to a class: *A frog eats . . .*, pauses for an answer from the class, and completes the sentence: *. . .flies*. In other words, the teacher puts a planned meaningful pause within the sentence. A similar pause, often used for dramatic effects as well, can occur in the following sentence: *A frog is a . . .fly eater*. However, a pause between *fly* and *eater* in *A frog is a fly . . . eater* is unacceptable: it almost makes the sentence uninterpretable. Given that *fly eater* is a compound noun, the above contrast in the effects of pauses can be accounted for if we assume that planned meaningful pauses cannot occur inside a lexical category.

In an object-verb sequence in a sentence in Hindi, a pause occurring between the object and the verb blocks the reading B. This would follow from (10b) together with the assumption that planned meaningful pauses cannot occur inside a lexical category.[22]

Another piece of evidence for the N + V in NI to be treated as a lexical category comes from the facts of stress and word melody. Sub(ordinate) compounds of the form [head + modifier] in Hindi have a single primary stress and word melody, characteristic of single words, as opposed to co(ordinate) compounds of the form [head + head + · · ·], which have as many primary stresses and word melodies as they have stems, characteristic of word sequences. Consider the asymmetry between the compounds in (31a) and (31b), where (31a) is a subcompound and (31b) is a cocompound:

(31)

a. jáanakiinandan
Janaki's son

b. jáanakiinándan
Janaki and Nandan

What (31) illustrates is a systematic correspondence between morphological construction types (subcompound vs. cocompound) and their

[21] I am grateful to an anonymous reviewer for bringing these examples to my attention.

[22] Planned meaningful pauses may occur between a nominal and its case clitic. This is consistent with the prohibition of pauses within a lexical item. Recall that a case clitic in Hindi is attached phrasally; this accounted for the impossibility of case marking on an incorporated nominal, and the possibility of case marking a conjoined nominal.

phonological properties (stress and word melody). This correspondence is expressed in the level ordered conception of lexical phonology and morphology by associating the two constructions with sequentially ordered strata or levels and defining the domain of the phonological patterning in terms of these levels. In a theory of phonology and morphology that does not admit sequential modularity, the correspondence is expressed by associating the different construction types with different representational units (e.g. prosodic stem, prosodic word) and defining the domain of the phonological patterning in terms of these units. The choice between the two strategies is not crucial for my purposes.[23]

In the level ordered conception, the pattern in (31) is accounted for by assuming that (i) subcompounding and cocompounding constitute different lexical strata, (ii) the former is an input to the latter, and (iii) stress and word melody assignments take place at the end of the subcompounding stratum.[24] Under these assumptions, the subcompound in (31a) is assigned a single stress and word melody, whereas each stem in the cocompound in (31b) is assigned an independent stress and word melody.

The N + V in NI behaves exactly like a subcompound in that it has a single stress and word melody, as in (32a):

(32)a. gʰóḍe becnaa
horse-PL sell-NF
to horse-sell

b. gʰóḍe bécnaa
horse-PL sell-NF
to sell horses

c. safed gʰóḍe bécnaa
white horse-PL sell-NF
to sell white horses

[23] For a detailed discussion of the two strategies, and arguments against sequentiality, see K. P. Mohanan (in press).

[24] This is similar to the pattern observed in Malayalam (K. P. Mohanan 1982a).

d.
gʰóɖõ-ko bécnaa
horse-PL.A sell-NF
to sell the horses

The contrast in stress and word melody in (32a) and (32b–d) follows if the N and V in (32a) combine in the same lexical submodule as the subcompound in (31a). In a non-sequential approach, the corresponding hypothesis would be that subcompounds and the N + V in (32a) are of the same morphological construction type.

Yet another piece of evidence that NI in Hindi takes place in the lexical module comes from an optional phonological process of stem-final high vowel shortening. A stem-final high vowel that is long when word-final can optionally be short when not word-final, as shown by (33b):

(33)a. jáanakii ~ *jáanaki
Janaki

b. jáanakiinandan ~ jáanakinandan
Janaki's son

c. jáanakiinándan ~ *jáanakinándan
Janaki and Nandan

d. jáanakii-ne ~ *jáanaki-ne
Janaki-ERG

The stem-final *ii* in *jaanakii* may optionally be short in the subcompound in (33b). However, it cannot be short word-finally ((33a)), or within a cocompound ((33c)). The facts in (33) would follow if the domain of vowel shortening is the subcompounding submodule, like stress and word melody assignment. (33d) shows that the stem-final vowel cannot be shortened when followed by a case clitic. Recall that the case clitic is attached in the phrasal module. Hence, when shortening takes place, the long vowel is final, and cannot shorten.

If the domain of vowel shortening is a subcompound, and the N + V in NI is like a subcompound with respect to stress and word melody, we

predict that vowel shortening should be possible in the N + V in NI. This prediction is borne out by the facts in (34):

(34)a. (i) miṭ^{h}aa**ii** banaaii. (ii) miṭ^{h}aa**i** banaaii
sweets-N make-PERF *sweets-N make-PERF*
made sweets.

b. (i) sabz**ii** becii. (ii) sabz**i** becii
vegetables-N sell-PERF *vegetables-N sell-PERF*
sold vegetables.

This vowel shortening is not found in all speakers. However, for those who can shorten the vowel, the process is systematic: when the vowel is short, the noun and verb must have a single primary stress and word melody, and must exhibit the syntactic properties discussed in section 2.1. These facts follow if vowel shortening applies within the same lexical submodule or morphological construction type as the subcompound.

Now, one might entertain the idea that the domain of stress, word melody, and vowel shortening can be specified in terms of the notion *phonological word* à la Sproat (1986), without appealing to the lexical module or the corresponding morphological type. In order to evaluate this proposal, it is first necessary to distinguish two meanings of the term 'phonological word'. In the sense of Liberman and Prince (1977), it refers to a metrical structure created by putting together a number of feet. That is, phonological word is the *output* of syllabification, foot tree formation, and word tree formation. We refer to this concept as METRICAL WORD (Inkelas 1989, Cole 1990). For Sproat, the term phonological word refers to a unit that acts as a domain in which phonological rules can apply, including metrical rules like syllabification, foot tree formation, and word tree formation. We refer to this concept as PROSODIC WORD. Thus, the prosodic word defines the domain on which a metrical word is constructed; in other words, it is the label for a morphological construction type. The term 'phonological word' has been used in the literature indiscriminately to refer to either metrical word or prosodic word.

Given this distinction between a phonological structure (metrical word) and a domain in which the structure is built (prosodic word), the domain of stress cannot be the metrical word, because metrical words are the result of stress assignment. The metrical word cannot be the domain of vowel shortening either. A noun with a case clitic forms a single metrical word; together, they take a single stress and word melody: *dahii-mẽ* "in the yogurt"; *sabzii-se* "with the vegetable". Yet, the vowel in these nouns

cannot be short: **dahi-mẽ*; **sabzi-se*. Therefore, the domain of vowel shortening cannot be captured in terms of the metrical word.[25]

The alternative is to state stress assignment, word melody, and vowel shortening in terms of the prosodic word. If we assume that prosodic words are the output of a particular sequentially ordered lexical submodule (Inkelas 1989), our arguments that NI is in the lexical module remain intact. If we do not assume sequential modularity, as in Sproat's (1986) approach, prosodic wordhood is a way of representing a morphological construction type. Thus, whether in terms of sequential modularity or morphological construction types, noun incorporation in Hindi must be analysed as belonging in the lexical module.

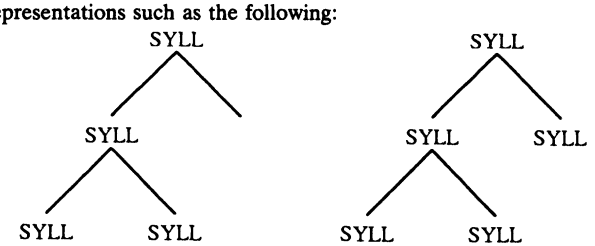
2.3.2. *Morphology*

Morphological evidence also points to the conclusion that NI in Hindi is part of the lexical module: the derivational suffix *-vaalaa*, which may be thought of as an agentive-occupational marker, can be attached to the N + V sequence as illustrated in (35):

(35) g^{h}oṛe becne vaalaa
horse-PL sell-NF
horse-seller

The sequence N + V-*vaalaa* in (35) takes a single primary stress and word melody, and can undergo stem-final high vowel shortening. The object nominal in (35) cannot be modified or conjoined, or take a case clitic. Furthermore, it must be adjacent to the verb. These properties of the N + V-*vaalaa* construction can be explained if we analyse the N + V as

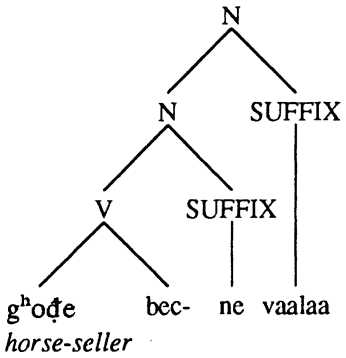
[25] An anonymous reviewer suggests that these facts can be captured in terms of the notion metrical word if we represent, e.g. *dahii-mẽ* as a metrical word containing another metrical word. ([$_{w}$[$_{w}$*dahii*]*mẽ*]). This proposal, however, violates the Strict Layer Hypothesis in Prosodic and Metrical phonology which disallows the recursion of prosodic metrical units. If we allowed metrical units to contain metrical units of the same level, there is nothing that prevents representations such as the following:



Given that such freedom is widely acknowledged to be undesirable, I take it that the reviewer's proposal is unimplementable.

an NIC. If so, NIC is an input to derivational morphology. I suggest that the structure of (35) is as in (36):

(36)



Accepting the widely held assumption that derivational morphology is part of the lexicon, we must also assume that NI is part of the lexical module.[26]

The attachment of *-vaalaa*, though fairly productive, is not uncon-

[26] The derivational suffix *-vaalaa* must be distinguished from the *-vaalaa* used as a relative clause marker in (i), and as the marker of immediate future in (ii) (Verma 1971):

(i) moṭe pannõ aur šaandaar tasviirõ vaalii vah kitaab . . .
thick pages and grand pictures that book
That book, which has thick pages and grand pictures . . .

(ii) raam kaanfrens-ke liye kæliforniya jaane vaalaa hai.
Ram conference-for California go about to is
Ram is going soon/about to go to California for a conference.

The form *-vaalaa* also appears in an N + *vaalaa* construction, where it is an occupation marker, illustrated in (iii):

(iii)a. gʰoḍaa gʰoḍe-vaalaa
horse a horse keeper/seller (M)

b. sabzii sabzii-vaalii
vegetable vegetable vender/seller (F)

c. miṭʰaaii miṭʰaaii-vaalaa
sweets sweet-seller/marker (M)

d. ṭʰelaa ṭʰele-vaalaa
cart cart pusher/owner (M)

The *-vaalaa* in (iii) is clearly derivational. The *-vaalaa* that appears in (35) shares with the *-vaalaa* in (iii) all the properties we have discussed so far as being characteristic of lexical elements, such as the genericity requirement, prohibition against case marking, vowel shortening, and so on. In this respect, the *-vaalaa* in (35) and (iii) sharply contrasts with those in (i) and (ii). Therefore, (35) cannot be treated as a relative clause.

strained. Thus, it is not difficult to coin novel forms like *caawal k^{h}aane vaalaa* "rice-eater" and *lakḍii kaaṭne vaalaa* "wood-cutter", from *caawal k^{h}aanaa* "rice-eating" and *lakḍii kaaṭnaa* "wood-cutting", respectively. But coinages like *gaaḍii d^{h}one vaalaa* "car washer" and *g^{h}oḍe k^{h}ariidne vaalaa* "horse buyer" are extremely awkward if not impossible, even though the NI's *gaaḍii d^{h}onad* "car-washing", and *g^{h}oḍe k^{h}ariidnaa* "horse-buying" are quite acceptable. Irregularities such as these support the lexical status of NI in Hindi.

3. Verb Agreement and NI

Verb agreement in Hindi exhibits an alternation between subject agreement and object agreement. A verb agrees with its object if and only if the object is NOM and the subject is not NOM.[27] In this section, I will outline the facts of verb agreement in Hindi, demonstrate that agreement must take place in the phrasal module, and then explore verb agreement in the NI construction, where the verb can agree with a lexically incorporated noun. Juxtaposing the analysis of NI as lexical and that of verb agreement as phrasal (i.e., postlexical) forces us to re-examine and tease apart different issues surrounding lexicality.

3.1. *Verb Agreement: A Phasal Phenomenon*

A verb in Hindi can agree only with a NOM argument. In (37) below, the verb agrees in number, gender, and person with its NOM grammatical subject (SUBJ).

(37)a. ravii baalak-ko uṭ^{h}aaegaa.
*Ravi-N(M) boy-A(M) lift-FU.**M.SG***
Ravi will lift up the boy.

b. niinaa baalak-ko uṭ^{h}aaegii.
*Nina-N(F) boy-A(M) lift-FU.**F.SG***
Nina will lift up the boy.

In (38), the SUBJ is non-nominative (NONNOM), and the primary object is NOM. The verb agrees with this OBJ:

[27] In Hindi, subjects may bear non-nominative (ergative/dative/instrumental/locative/genitive) case, depending on meaning. Objects, as we have seen, are either ACC or NOM.

(38)a. {ravii-ne / *Ravi-E(M)* ; niinaa-ne / *Nina-E(F)*} roṭii *bread-N(F)* k^{h}aayii. *eat-PERF.F.SG*

Ravi/Nina ate bread.

b. {ravii-ne / *Ravi-E(M)* ; niinaa-ne / *Nina-E(F)*} kelaa *banana-N(M)* k^{h}aayaa. *eat-PERF.M.SG*

Ravi/Nina ate a banana.

In (39), the primary object is also NONNOM. The verb is in the default or non-agreeing form: it bears the third person singular masculine inflection:

(39) {ravii-ne / *Ravi-E(M)* ; niinaa-ne / *Nina-E(F)*} {baalak-ko / *boy-A(M)* ; baalikaa-ko / *girl-A(F)*} uṭ^{h}aayaa. *lift-PERF.M.SG*

Ravi/Nina lifted up the boy/girl.

If the SUBJ and the OBJ are both NOM, then the verb agrees with the SUBJ, as in (40):

(40)a. ravii *Ravi-N(M)* {roṭii / *bread-N(F)* ; kelaa / *banana-M(M)*} k^{h}aaegaa. *eat-FU.M.SG*

Ravi will eat bread/banana.

b. niinaa *Nina-N(F)* {roṭii / *bread-N(F)* ; kelaa / *banana-N(M)*} k^{h}aaegii. *eat-FU.F.SG*

Nina will eat bread/banana.

The generalization about verb agreement in Hindi recognized widely in the literature (e.g. Kachru, Kachru, and Bhatia 1976, p. 86) is as follows. If the SUBJ is NOM, the verb agrees with it. If the SUBJ is NONNOM, and the OBJ is NOM, the verb agrees with the OBJ. If the OBJ is also NONNOM, the verb is in the neutral form, namely, masculine third person singular. In short, NOM case is a necessary condition for agreement. The principle governing agreement may be stated as follows:[28]

(41) *Verb Agreement in Hindi*
The verb agrees with its highest argument associated with NOM case.

Thus, verb agreement in Hindi depends upon the NOM case on the argument. NOM case is determined in the phrasal module. Therefore, verb agreement is determined in the phrasal module, and not on the basis of lexical information alone.[29]

One could imagine a situation in some language where the case of the arguments of a verb is entirely predictable from the lexical entry of the verb, and hence, the case requirement for agreement can be satisfied lexically. But in Hindi this is untenable. First, for the same verb, the subject may be either NOM or NONNOM, depending on the presence of a modal:

(42)a. ravii baiṭʰ gayaa.
Ravi-N(M) sit-NF go-PERF.M.SG
Ravi sat down.

b. niinaa baiṭʰ gaii.
Nina-N(F) sit-NF go-PERF.F.SG
Nina sat down.

[28] The principle in (41) makes reference to the "highest argument": that is, it makes reference to prominence relations of a particular kind, namely, those among the arguments of a predicate. The principle could equally well be stated in terms of prominence relations of another kind, namely, those among grammatical functions: *The verb agrees with its highest grammatical function associated with NOM case.* The choice between these two interpretations of "highest" does not in any way affect the point of this paper.

[29] Despite significant differences, analyses available in the literature (for instance, Gair and Wali 1989; Mahajan 1989) agree on the assumption that Hindi verb agreement is not a lexical phenomenon.

c. {ravii-ko / *Ravi-N(M)* ; niinaa-ko / *Nina-N(F)*} baiṭʰ jaanaa paḍaa
sit-NF go-NF modal (obligation)-PERF.M

Ravi/Nina had to sit down.

In (42a) and (42b), the SUBJ is NOM, and the verb agrees with it. But when the modal of obligation *paḍaa* is present ((42c)), the SUBJ bears dative case, and the verb fails to agree with it.[30] Given that agreement is sensitive to the presence of NOM case, and that the modal can assign dative case to the SUBJ, it follows that agreement with the SUBJ is not lexical. Since agreement with the SUBJ pre-empts agreement with the OBJ, it follows that OBJ agreement is not lexical either.

Secondly, as stated earlier, an OBJ is NOM if inanimate, and ACC if animate. Now, a verb may require that its object be animate or inanimate. Or it may make no such demand: its object may be either animate or inanimate. So the animacy of the OBJ nominal, and hence its case, can be ascertained only after the OBJ and the verb have been put together. In sum, we must conclude that verb agreement in Hindi takes place in the phrasal module.

3.2. *Agreement with the Incorporated Object*

We have seen compelling evidence that in Hindi NIC, the verb and one of its arguments form a single unit in the lexical module. We have also seen that agreement between the verb and its argument cannot be lexical. Let us now look at the facts of agreement in NI:

(43)a. anil kitaabẽ becegaa. (=(15a))
Anil-N(M) book-N.PL(F) sell-FU.M.SG

A. Anil will sell books.
B. Anil will do book-selling.

---

[30] The modal does not form a word with the verb: it can be attached to a coordinate verb structure:

(i) bahut baar uṭʰnaa aur baiṭʰnaa paḍaa.
a lot times rise-NF and sit-NF modal(obligation)-M.PERF
Had to get up and sit down many times.

Given our assumption that lexical categories cannot be conjoined, the coordinate structure *uṭʰnaa aur baiṭʰnaa* must be a $\bar{V}$, not a V. Therefore, *paḍaa* is attached to a $\bar{V}$.

b. anil-ne kitaabẽ becĩĩ. (=(16a))
Anil-E(M) book-N.PL(F) sell-PERF.F.PL
A. Anil sold books.
B. Anil did book-selling.

In (43a), the verb agrees with the masculine singular SUBJ. In (43b), the verb agrees with the feminine plural OBJ. This holds even when (43b) has reading B, and is therefore an NIC.

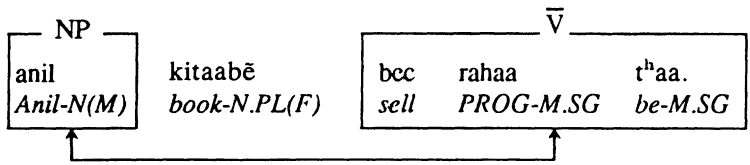
Now, verb agreement is a relation between the entire verbal complex and one of the arguments; it is not always manifested on the main verb. In (44), the agreeing unit is not the main verb, but the whole verbal complex: agreement is marked on the progressive and the *be* form:

(44)a. anil kitaabẽ bec rahaa t^{h}aa.
Anil-N(M) book-N.PL(F) sell PROG-M.SG be-M.SG
Anil was selling books.

b. ila kitaabẽ bec rahii t^{h}ii.
Ila-N(F) book-N.PL(F) sell PROG-F.SG be-F.SG
Ila was selling books.

The agreement relation in (44a) is diagrammatically given in (45):

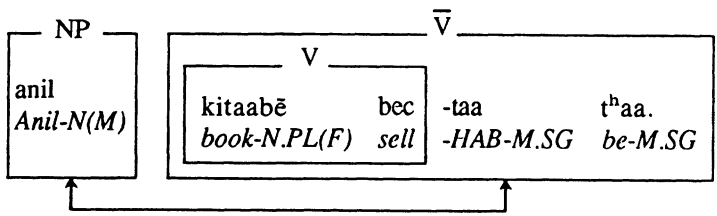
(45)



"Anil was selling books."

Under our analysis of NI, the incorporated noun is part of the agreeing verb complex:

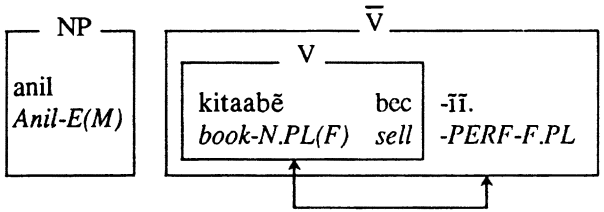
(46)



"Anil used to do book-selling."

As shown by (43b), the agreeing verbal complex can agree with the OBJ even when the OBJ is incorporated, that is, when it is internal to the verbal complex. The agreement relation in (43b) is given in (47):

(47)



"Anil did book-selling."

This brings us directly to the two puzzles posed by the NI construction. First, incorporation is lexical, and agreement is not; how can a verb agree with an incorporated argument? Second, if verb agreement is between the verbal complex and one of its sisters, how can the verbal complex agree with an element internal to it? In what follows, we will see that solutions to these problems demand the separation of two dimensions of syntactic information, those of grammatical functions and grammatical categories.

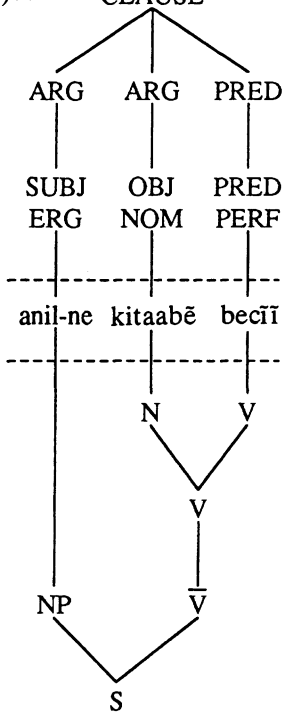
3.3. *A Formal Analysis: Factorization of Information*

The restrictions on the NI construction involving nominal modification, case clitics, adjacency, conjoining, gapping, pauses, stress and word melody, stem-final high vowel shortening, and *vaalaa* attachment motivate the analysis of N + V in Hindi NI as a lexical category formed in the lexical module. At the heart of our explanation for these restrictions is the category structure in (10b). The principle of agreement in (41), on the other hand, makes no reference to category structure. Agreement in Hindi is sensitive to the argument status of the nominal, and to NOM case. The syntactic independence of the argument is not captured by the category information in (10b), but by the grammatical-function information under it.

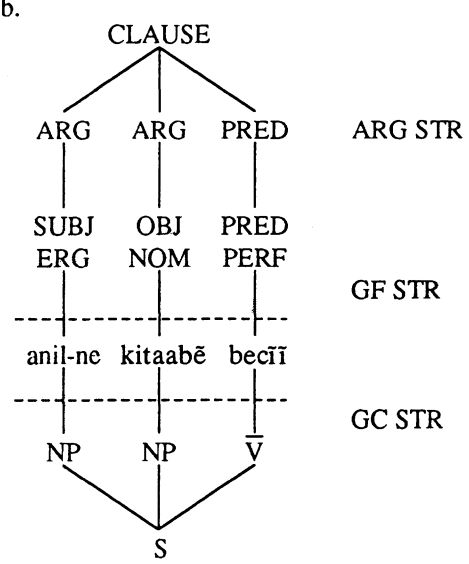
Without stating it explicitly, what we have done in (10b) is to separate category information from the predicate-argument and grammatical-function information required for agreement. Suppose we represent these different types of information along different dimensions. This separation is familiar in the literature, and has been argued for in Bresnan (1982a)

within Lexical-Functional Grammar.[31] The conditions for agreement can then be satisfied along the dimension of predicate-argument and grammatical-function information, represented in (48) below as ARG STR and GF STR above the sentence.[32] Two representations for (43b) are given below. (48a) is the representation for (43b) as an NIC and (48b) is its unincorporated counterpart:

(48)a.



b.



Even though (48a) and (48b) have two distinct category structures, represented as GC STR, corresponding to the two readings of (43b), the representation of ARG STR and GF STR for both readings is identical. Observe that this representation satisfies the requirements for agreement. Although incorporated on the dimension of GC STR in (48a), the N is an independent ARG in ARG STR, and an OBJ associated with NOM

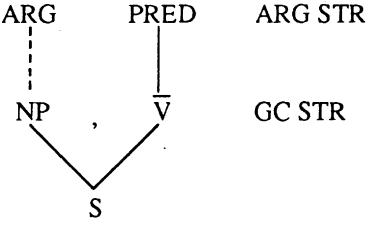
[31] The separation of information is also found in works within GB (Riemsdijk and Williams 1981, Chomsky 1981, Hale 1983, K. P. Mohanan 1984, Zubizaretta 1987). However, grammatical function information in these works is also represented in terms of category labels.

[32] Following the framework in T. Mohanan (1990), I assume the separation of argument structure information and grammatical function information in the representations in (48). I will not defend the separation here because it is not directly relevant for the purposes of this paper.

case in GF STR. It is therefore a legitimate controller of agreement. In short, the solution to the puzzles posed by NI in Hindi lies in an organization of grammar which allows non-identity of the mother-daughter relations with respect to different dimensions of syntactic information.

Given any parallel representations, it is necessary to specify the constraints that hold on the relation between them. First consider the interpretation of *anil-ne* as the SUBJ and *kitaabē* as the OBJ of the verb *bec* in (48b). I assume that this association is governed by (49). The convention of solid and dotted association lines notate the premises and consequences, respectively, of an *if-then* conditional.

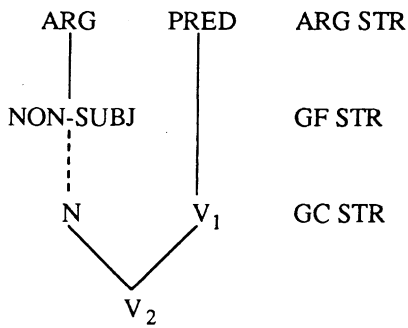
(49) Optional



(49) is interpreted as: "If the GC STR is $[NP, \bar{V}]_S$, and $\bar{V}$ is associated in ARG STR with a PRED that has an ARG, then associate the NP with the ARG. [Optional]" The ARG can bear any of the GFs of the PRED. Thus, (49) allows any NP dominated by S to be associated with an argument of the predicate.[33]

The interpretation of the GF of the N in NIC calls for an additional principle, given in (50):

(50)



(50) is interpreted as: "If the GC STR is $[N\ V_1]_{V_2}$, and V_1 is associated in ARG STR with a PRED that has a NON-SUBJ ARG, then associate

[33] Within this conception, the grammar also requires principles that state the relationship between grammatical functions on the one hand, and case, word order, nominal meanings (e.g. animacy), and the like on the other.

N with the ARG." (50) allows an N dominated by V to be interpreted as a nonsubject argument of the predicate.

Within the formalism of LFG, these principles will be stated as the annotations of phrase structure rules which state the mapping between c-structure and f-structure. For example, (49) corresponds to the annotation $\uparrow GF = \downarrow$ on the NP in a PS rule $S \rightarrow NP^*, \bar{V}$. (50) corresponds to the annotation $\uparrow NON\text{-}SUBJ = \downarrow$ on the N in a PS rule $V \rightarrow N, V$. (49) states the effect of 'scrambling', that is, the independence of grammatical functions from word order, which is stated in a movement theory in terms of optional movement. Within the formalism of head movement (Baker 1988), (50) corresponds to the head movement of an N from an NP into a V. As in the theories mentioned above, (49) and (50) are derived from a universal theory that expresses the invariance and constrained valiability in the relation between arguments and their grammatical functions on the one hand, and their grammatical categories on the other.[34]

3.4. *Absence of Doubling in Hindi NI*

The English compound *stage-manage* has the same GC STR as postulated for Hindi NI: $[N\ V_1]_{V_2}$. Nevertheless, as shown by the possibility of independent phrasal objects, e.g., *the show* in *John stage-managed the show*, the N in the $[N\ V_1]$ cannot be a syntactic argument of the verb. The grammar of NI that we have sketched so far predicts that unlike what happens in the English *stage-manage*, an incorporated object in Hindi cannot co-occur with a phrasal object, under the assumption that a clause cannot contain two instances of the same grammatical function (function-argument biuniqueness within LFG). This prediction is borne out by (51c):

(51)a. anil jaanvar becegaa.
Anil-N(M) animal-N sell-FU.M.SG
A. Anil will sell animals.
B. Anil will do animal-selling.

b. g^hoḍõ-ko anil becegaa.
horse-PL-A Anil-N(M) sell-FU.M.SG
Anil will sell the horses.

c. *g^hoḍõ-ko anil jaanvar becegaa.
horse-PL-A Anil-N(M) animal-PL sell-FU.M.SG
Anil will do animal-selling of the horses.

[34] See, for instance, Alsina (1993).

(51c) is ungrammatical because its GF structure contains two PR.OBJ's, animals and horses.

Rosen (1989) points out that natural languages fall into two types with respect to modifier stranding as well as doubling in the NI construction. Languages like Caddo and Mohawk allow an incorporated noun to be modified by a 'stranded' word or phrase outside the verb complex (Sadock 1985, pp. 402, 407; Baker 1988, p. 93; among others). They also allow doubling. In contrast, Polynesian and Micronesian languages disallow such modification and doubling (Rosen 1989, pp. 311–2). Clearly, Hindi belongs to the latter class with respect to both these properties. The status of this typological distinction and its theoretical consequences will be taken up in the next section.

4. Modularity and Representation

What we have seen so far is that Hindi has an N + V sequence which has the crucial characteristics of NI: morphologically, it behaves like a single word, but the N within this word is a syntactic argument. This section examines how this duality is formally expressed in other accounts of NI (Sadock 1980, 1985, 1986, 1991; Baker 1985, 1988), and articulates the substance that the analysis in this paper shares with their accounts. It also shows in what ways my proposals differ from those of Sadock and Baker.

In the analysis of NI in Hindi, we are faced with issues of both representation and modularity. The modularity issue can be stated as: is Hindi NI part of the lexical module or the phrasal (=post lexical) module? As far as the issue of representation is concerned, we have two questions. First, does Hindi NI need to distinguish between the representation of information relevant for the structure within words (lexical) and that relevant for the structure across words (phrasal)? Second, does the representation of Hindi NI need to distinguish between grammatical categories and grammatical functions? We turn now to these issues.

4.1. *The Lexicality of NI*

To repeat, most accounts of NI acknowledge that the N + V sequence in an NI construction exhibits lexical properties. This raises the question

what we mean by the term 'lexical'.[35] In order to clarify the issues of lexicality, it is essential that we separate at least two different conceptions: (i) 'lexical" as referring to a MODULE in which representations are formed, and (ii) 'lexical' as referring to a UNIT OF REPRESENTATION. The claim that NI in Hindi is lexical can be interpreted as either (52a) or (52b), or both:

(52)a. NI is part of the lexical module;
b. The N + V sequence in NI is a lexical category.

The claim in (52b) seems uncontroversial: all accounts of NI accept that the lexical properties of NI require it to be represented as a lexical category at some dimension of representation or stage in the derivation (see (11), (12), and (13)). One of the claims of this paper is that, in addition, the N + V sequence must be concatenated in the lexical module ((52a)). In other words, the N + V sequence in the Hindi NIC is lexical in both senses: it is a lexical category, and is formed in the lexical module. The claim in (52a) is implicit in Sadock's analysis of NI: the level of representation in which NI is a lexical category, called the morphological representation, as distinct from the syntactic representation, encapsulates the autonomy of morphology required by (52a).

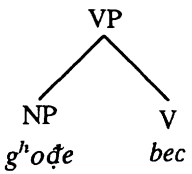
The statement in (52a) becomes a logical consequence of the statement in (52b) if we assume that lexical categories cannot be created in the postlexical module, and phrasal categories cannot be created in the lexical module. Baker accepts (52b), but rejects (52a) by claiming that representations of phrasal categories may constitute the input to the formation of lexical categories through head movement. Our evidence for (52a) comes from the analysis of the phonological and morphological properties of NI: it undergoes lexical phonological processes such as stress and word melody assignment and final high vowel shortening. It is also an input to derivational affixation, widely acknowledged to be a characteristic of the lexical module. It is unclear how Baker's analysis of NI can account for these facts.

[35] During the early days of generative grammar, lexical was that which is idiosyncratic or unpredictable in a language. The lexicon was therefore viewed as the list of morphemes, specifying all their unpredictable properties. With Chomsky (1970), the lexicon came to include not only idiosyncratic information, but also certain kinds of predictable information such as the regularities of derivational morphology and compounding. Over the years, the lexicon has developed into a rich and organized module of the grammar whose output captures the notion 'word' (e.g. Lieber 1980, Bresnan 1982a, K. P. Mohanan 1982, Kiparsky 1982).

4.2. *The Syntacticity of NI*

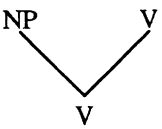
Sadock's lexical (=morphological) representation of *g^{h}oḍe bec* "horse-sell" would be as in (53b), and the representation of the syntactic objecthood of the incorporated noun as in (53a):

(53)a.



SYNTACTIC

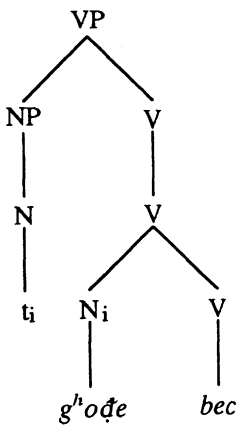
b.



LEXICAL (=MORPHOLOGICAL)

(53a) and (53b) are parallel representations: they do not bear an input–output relation. Baker derives the representation in (53b) from (53a) through head movement:

(54)



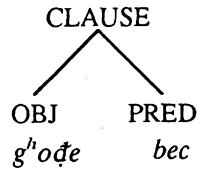
When (53) and (54) are juxtaposed, it becomes obvious that the coindexed trace within a single dimension of representation in (54) performs the same task as the linking of entities in two dimensions of representation in (53): both express the idea that the nominal inside the verb is at the same time the object of the verb.

Like Sadock and Baker, the analysis in this paper recognizes that the incorporated noun has syntactic reflexes. In particular, it functions as the syntactic object of a verb. The syntactic reflexes of NI are expressed in

the representation of its argument and GF structure, where the incorporated object ((48a)) is an argument on par with a phrasal object ((48b)).

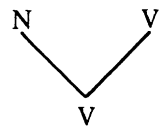
In order to factor out the notational differences from the substantive differences between the analyses, we must bear in mind that both Sadock and Baker represent grammatical functions in terms of category labels and dominance relations. The representations in (55a, b) are the counterparts of (53a, b) in my analysis:[36]

(55)a.



GF STR

b.



GC STR

The above discussion reveals that the formal representations in (53), (54), and (55) are roughly equivalent in recognizing the dual nature of NI: they agree that NI is simultaneously 'lexical' (it constitutes a lexical category) and 'syntactic' (it consists of a predicate and its object). Thus, the facts of adjacency, prohibition against conjoining and against the occurrence of case clitics, *wh*-words, and pronouns within NIC, are equally well accounted for by the structure $[N\ V]_V$ in the three representations. Verb agreement and the prohibition against doubling are accounted for by the structure $[NP\ V]_{VP}$ in (53) and (54), corresponding to [. . . OBJ PRED]$_{CLAUSE}$ in (55).

A significant difference then is that while (55) factors out the representation of grammatical functions from that of grammatical categories along two dimensions of *syntactic* representation, (53) and (54) conflate them into a single representation. Empirical evidence for the former position comes from the analysis of modifier stranding and gapping in Hindi.

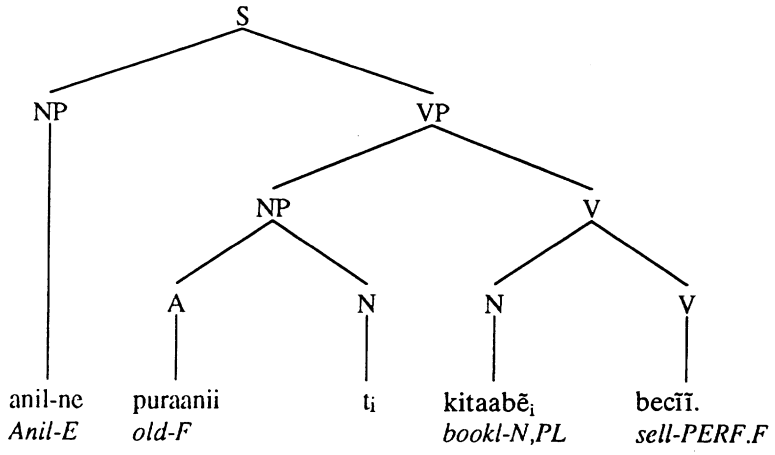
[36] The tree notation in (55a) is a notational variant of the bracket notation of functional structure in LFG:

$$\begin{bmatrix} \text{OBJ} & [g^h o\dot{d}e] \\ \text{PRED} & [bec] \end{bmatrix}$$

4.2.1. *Modifier Stranding*

As noted earlier, an incorporated nominal cannot be modified in Hindi. One of Baker's motivations for analysing incorporation as head movement in syntax is that in a number of languages, for instance, Mohawk, Onondaga, and Southern Tiwa, the incorporated noun can be modified by a word or phrase that "remains morphologically outside the verb complex" (Baker 1988, p. 93). An extension of the analysis to Hindi would allow incorrect representations containing stranded modifiers such as in (56). As far as I know, there is no constraint in Baker's analysis that accounts for the inability of the N to leave behind a stranded modifier:

(56)

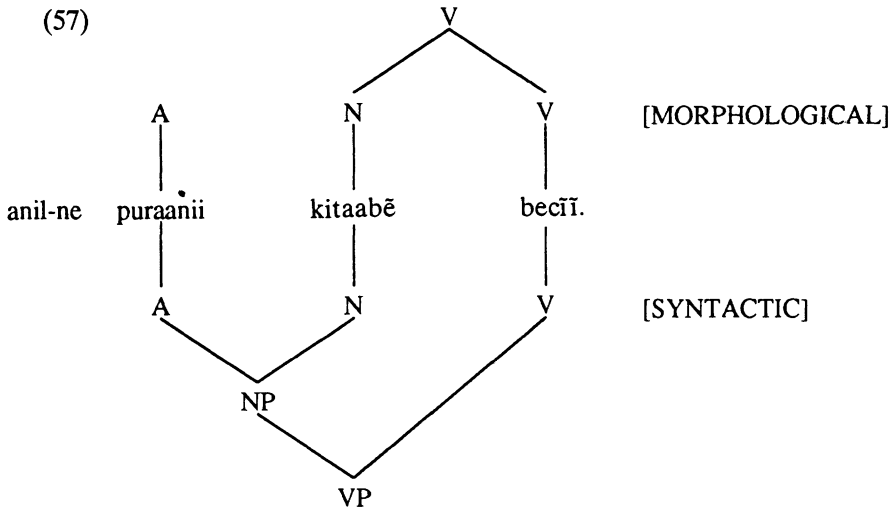


*Anil did old book-selling.

Thus, a head movement analysis must either assume that NI in Hindi is not the result of head movement, or stipulate that NI in Hindi prohibits modifier stranding.

The analysis of the incorporated N as being associated with an NP via coanalysis (following Sadock 1985, 1991) faces a similar problem. Given the basic premises of Sadock's theory, it is difficult to see how the representation of NI as in (57) can be ruled out for Hindi:

(57)



Given what we have said about Hindi NI, neither (56) nor (57) are legitimate structures: The noun *kitaabẽ* in the N + V sequence is part of a lexical category; its modifier *puraani* is outside the lexical category. By the Lexical Integrity Hypothesis (see section 2.1.6), therefore, the internal structure of ***kitaabẽ becĩĩ*** is opaque to phrasal modification.[37] Within our analysis, there is no representation in which the N in NIC is associated with an NP.

4.2.2. *Gapping*

Recall that neither the N nor the V in NIC can be gapped in Hindi ((8b), (9b)). Given our analysis of the N + V sequence as a lexical category, the failure to gap follows from the Lexical Integrity Hypothesis. Since the Lexical Integrity Hypothesis is stated on grammatical categories, and does not hold on grammatical functions, syntactic phenomena that are stated on grammatical functions, like verb agreement in Hindi, are irrelevant for the principle. Given a syntactic movement analysis or the co-analysis of morphology and syntax, where grammatical functions are represented in terms of categories, the failure to gap remains unexplained.

One may think that the explanation lies in the N and the V being lexical (X^0) categories. However, as shown by the example in (58) below, there is no constraint against the gapping of a lexical category in Hindi.

---

[37] Sadock (1980) observes that modifier stranding and anaphoric reference in Greenlandic Eskimo violate the Lexical Integrity Hypothesis. For a response, see Simpson (1991, pp. 226–237).

(58) anil-ne laal kaar becii aur raam-ne niilii k^{h}ariidii.
Anil-E red car-N sell-PA and Ram-E blue buy-PA
Anil sold a red car and Ram bought a blue —.

Hence the failure of the N and V in NI to gap cannot follow from these constituents being lexical categories.

In sum, verb agreement and doubling show that the N in the Hindi NIC corresponds to a syntactic argument. Yet, NIC does not allow modifier stranding or gapping. This peculiar combination of syntactic and lexical properties makes Hindi NI significantly different from the instances of NI I have seen in the literature.

4.3. *A Purely Lexical Analysis*

Rosen (1989) claims that NI is purely lexical. She distinguishes two types of NI, which she calls Compound NI and Classifier NI, which differ in the following ways:

(59)

	Compound NI	Classifier NI
(a) the verb's valency is reduced	√	x
(b) NP OBJ's can co-occur with NI	x	√
(c) stranded modifiers are allowed	x	√

Rosen claims that doubling (59b) and stranding (59c) are found only in languages that allow null heads in the syntactic structure, independently of NI. The properties of classifer NI follow from the combination of the possibility of syntactic null heads on the one hand, and purely lexical NIs on the other. The apparent objecthood of the incorporated noun is the result of the incorporation of a non-argument co-occurring with a null head semantically related to it.

Rosen's claim, then, is that the NI construction does not in itself have any syntactic representation associated with it: the apparent syntactic properties which have been attributed to the incorporated noun are the effects of the independent existence of null heads in the language. Thus, she rejects the substance of the representations in both (53a) and (55a).

Rosen's theory predicts a correlation between the possibility of doubling (59b) and stranded modifiers (59c), both being consequences of the possibility of null heads. This correlation has been questioned by Sadock (1991), who shows that languages like Southern Tiwa and West Green-

landic allow stranded modifiers, but not doubling, and that Rosen's proposed solutions to this problem are untenable.

Hindi NI is also incompatible with Rosen's typology. Hindi does not allow null heads. Nor does NI in Hindi allow modifier stranding ((59c)) or doubling ((59b)). Therefore, Hindi NI should be compound NI. Yet, there is no reduction of valency in Hindi NI in that NI does not make a transitive clause intransitive: the incorporated nominal behaves as an independent ARG with respect to verb agreement. Thus, the predicted correlation between (59a) on the one hand, and (59b, c) on the other, is incorrect.

In the light of these facts, Rosen's typology of NI in (59) must be revised as (60):

(60)

	A	B	C	D
a, N is an argument	x	√	√	√
b. Modifier stranding allowed	x	x	√	√
c. Doubling allowed	x	x	x	√

Of the four classes in (60), type A is purely lexical (Rosen's compound NI). But by the definition of NI we have adopted, this is not NI, but simply compounding. Type D has the most characteristics normally associated with syntactic constructions (Rosen's classifier NI). Southern Tiwa and West Greenlandic fall into type C, and Hindi into type B. In short, the claim that NI is purely lexical, and that the incorporated noun has no syntactic status, must be abandoned.

The central claims that I advance in this paper are that an analysis of NIC in Hindi requires syntactic theory to recognize GF representation and GC representation as separate though interacting syntactic representations, and that lexical categories are the output of the lexical module. Now, the construction that has been called NI obviously allows certain typological differences across languages. The analysis of NI in terms of (55) captures the absence of modifier stranding, doubling, and gapping in a language such as Hindi. Clearly, this analysis cannot be directly extended to languages which do allow modifier stranding and/or doubling. A universal theory of NI should predict the range of typological variation. However, such a theory is not within the scope of this paper.

5. Negation and NI

Our analysis of NI hinges on the dual structure that separates category information from grammatical function information ((55)), contra Sadock ((53)) and Baker ((54)). We found evidence for this separation in modifier stranding and gapping. In this section, I will show that the requirements for sentential negation in Hindi lend additional support to this dual structure. Sentential negation in Hindi simultaneously requires in category structure that the negative word be attached to the left of the verb, and in grammatical function structure that it be adjacent to the predicate.

5.1. *Phrasal and Sentential Negation*

Consider the facts of the placement and scope of the negative word *nahĩĩ* "no/not" in (61a, b):

(61)a. raam-ne ilaa-ko nahĩĩ dek^{h}aa t^{h}aa.
Ram-E Ila-A not see-PERF be-PA
Ram had not seen Ila.

b. niinaa ravii-ko kitaabẽ nahĩĩ b^{h}ej rahii t^{h}ii.
Nina-N Ravi-D book-PL.N not send PROG be-PA
Nina was not sending the books to Ravi.

The sentences in (61) are interpreted as instances of sentential negation oniy if the tonic, or the nuclear stress of the sentence, is placed on the word *nahĩĩ*. If the tonic falls on the phrase immediately before the word *nahĩĩ*, the result is the negation of that phrase, as illustrated below. The tonic is indicated by underlining.

(62)a. niinaa ravii-kjo **kitaabẽ** nahĩĩ b^{h}ej rahii t^{h}ii.
Nina-N Ravi-D book-PL.N not send PROG be-PA
Nina was sending not the **books** (but something else) to Ravi.

b. niinaa kitaabẽ **ravii-ko** nahĩĩ b^{h}ej rahii t^{h}ii.
Nina-N book-PL.N Ravi-D not send PROG be-PA
Nina was sending the books not to **Ravi** (but someone else).

c. ravii-ko kitaabẽ **niinaa** nahĩĩ b^{h}ej rahii t^{h}ii.
Ravi-D book-PL-N Nina-N not send PROG be-PA
Not **Nina** (but someone else) was sending (the) books to Ravi.

Phrasal negation as in (62) involves focussing the negated phrase: the focussed phrase in (62a) is "books", and the sentence presupposes that Nina sent something to Ravi, as is characteristic of the negation of a focus. Similarly, a *wh*-phrase is allowed in a sentence containing phrasal negation only if the negated phrase is not distinct from the *wh*-phrase, again a characteristic widely acknowledged as a property of focus.

Contrary to what one might conjecture on the basis of (62), the negative in phrasal negation is not necessarily adjacent to the verb, as shown by (63a, b):

(63)a. ?**niinaa-ne** nahĩĩ ravii-ko kitaabẽ bʰejĩĩ.
Nina-E not Ravi-D book-PL-N send-PERF
Not Nina sent the books to Ravi.

b. ?niinaa-ne **ravii-ko** nahĩĩ kitaabẽ bʰejĩĩ.
Nina-E Ravi-D not book-PL-N send-PERF
Not Nina sent the books to Ravi.

Although the examples in (63a, b) are a little odd, the oddness disappears if the right context is provided, as in (64a, b), where the negated phrase is a contrastive structure:

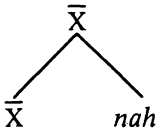
(64)a. **niinaa-ne** nahĩĩ ilaa-ne ravii-ko kitaabẽ bʰejĩĩ.
Nina-E not Ila-E Ravi-D book-PL-N send-PERF
Not Nina but Ila sent the books to Ravi.

b. niinaa-ne **ravii-ko** nahĩĩ mohan-ko kitaabẽ bʰejĩĩ.
Nina-E Ravi-D not Mohan-D book-PL-N send-PERF
Nina sent the books not to Ravi but to Mohan.

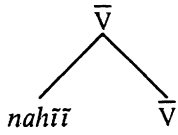
None of the sentences in (63) and (64) can be interpreted as sentential negation. In sentential negation, the negative is attached to the left of the verb, while in phrasal negation, it is attached to the right of the negated phrase. In order to account for this, we assume the following structures:[38]

[38] The structures in (65) assume for the sake of concreteness that the sister of *nahĩĩ* is an X-bar category. Nothing in the analysis crucially hinges on this assumption: the sister of *nahĩĩ* may be an X-zero category.

(65)a. *Phrasal negation*



b. *Sentential negation*



(65a) predicts that if the negated expression is the verb rather than the entire clause, the negative should occur after the verb. Thus, in contrast to the sentences in (61) which cannot be interpreted as instances of negation of the verb, those in (66) can be interpreted only as negation of the verb:

(66)a. ?raam-ne ilaa-ko **dek^{k}aa** nahĩĩ.
Ram-E Ila-A see-PERF not
Ram didn't **see** Ila.

b. ?niinaa-ne ravii-ko kitaabẽ **b^{h}ejĩĩ** nahĩĩ.
Nina-E Ravi-D book-PL-N send-PERF not
Nina didn't **send** the books to Ravi.

Like the sentences in (63), those in (66) are somewhat odd, because they require special contexts. Once the contexts are provided, they are perfectly acceptable:[39]

(67)a. raam-ne ilaa-ko **dek^{h}aa** nahĩĩ, bas yaad kiyaa.
Ram-E Ila-A see-PERF not enough memory-N do-PERF
Ram didn't **see** Ila, only remembered her.

---

[39] In a contrastive structure, phrasal negation can also be achieved by placing the tonic on the negated phrase and attaching the negative to the left of the verb, as in (i):

(i) **niinaa-ne** ravii-ko kitaabẽ nahĩĩ b^{h}ejĩĩ.
Nina-E Ravi-D book-PL-N not send-PERF
Nina didn't send the books to Ravi (someone else did).

In (i), the position of the negative indicates sentential negation, but the specific intonation narrows the scope of the NEG to the phrase that bears the nuclear sentence stress. It is interesting to note that if the negated phrase is indicated by the tonic, the NEG cannot occur to the right of the verb:

(ii) ***niinaa-ne** ravii-ko kitaabẽ b^{h}ejĩĩ nahĩĩ.
Nina-E Ravi-D book-PL-N send-PERF not

In (ii), the scope of NEG must be the verb because of the position of NEG, but it must be the underlined NP because of the tonic. The result is an uninterpretable sentence.

b. niinaa-ne ravii-ko kitaabẽ **bʰejĩĩ** nahĩĩ, kʰud
Nina-E Ravi-D book-PL-N send-PERF not self
jaa-kar dĩĩ.
go-NF give-PERF
Nina didn't **send** the books to Ravi, but took them to him herself.

The structures in (65) also correctly predict that *nahĩĩ* can be sentence initial if and only if it is immediately followed by the verb, and its interpretation is that of sentential negation:

(68)a. nahĩĩ bʰejĩĩ niinaa-ne ravii-ko kitaabẽ.
not send-PERF Nina-E Ravi-D book-PL.N
Nina did not send the books to Ravi.

b. *nahĩĩ niinaa-ne kitaabẽ ravii-ko bʰejĩĩ.
not Nina-E book-PL.N Ravi-D send-PERF

c. *nahĩĩ ravii-ko kitaabẽ niinaa-ne bʰejĩĩ.
not Ravi-D book-PL-N Nina-E send-PERF

In (68a), the NEG is attached to the left of the verb, and satisfies the structure for sentential negation ((65b)). In (68b, c), however, the NEG has nothing to its left to satisfy (65a), and the element to its right is not the verb, so it cannot satisfy (65b). The sentences are therefore ungrammatical.

5.2. *The Absence of Sentential Negation in NI*

Given the requirement for sentential negation that NEG must immediately precede the verb ((65b)), and the GC STR structure $[N\ V]_V$ for NIC ((55)), we predict that *nahĩĩ* cannot occur between the N and V in NIC. We will see below that this prediction is correct.

Since the N and V in NIC form a verbal unit, one would expect sentential negation in NIC with *nahĩĩ* to the left of the N + V. As it happens, however, this is not possible. The net result is that NIC does not permit sentential negation. I will show that this gap is explained by the independently required condition that the NEG must be adjacent to the PRED. That is, Hindi requires two conditions for sentential negation: in terms of categories, the NEG must immediately precede the verb; in terms of grammatical functions, it must be adjacent to the predicate. The need to

distinguish VERB and PREDICATE in this manner justifies the separation of the two dimensions of structure in (55).

5.2.1. *Sentential Negation and NI*

Consider (69), which is (15a) with *nahĩĩ* before the verb:

(69) anil kitaabẽ nahĩĩ becegaa.
Anil-N(M) book-N-PL(M) not sell-FU.M
(i) Anil will not sell (the) books.
(ii) *Anil will not do book-selling.

Recall that (15a) had two interpretations, one of them being reading B, associated with NI. Reading B is not available to (69). This is correctly predicted by our assumptions about the structures of NIC ($[\mathrm{N\ V}]_{\mathrm{V}}$ in (55)) and of sentential negation ($[nahĩĩ\ \bar{\mathrm{V}}]_{\mathrm{V}}$ in (65b)): the latter cannot occur inside the former: $*[\mathrm{N}[[nahĩĩ\ \bar{\mathrm{V}}]_{\bar{\mathrm{V}}}]_{\mathrm{V}}]_{\mathrm{V}}$.

Our analysis of Hindi NI also correctly predicts that if *nahĩĩ* occurs between an animate NOM object and the verb, the result should be ungrammatical ((70b)): if an animate object is NOM, it must be part of NIC, as in (70c), but *nahĩĩ* cannot occur between the N and V in NIC:

(70)a. ilaa baccõ-ko nahĩĩ samhaaltii hai.
Ila-N children-A not take care of-HAB be-PR
Ila doesn't take care of the children.

b. *ilaa bacce nahĩĩ samhaaltii hai.
Ila-N children-N not take care of-HAB be-PR

c. ilaa bacce samhaalti hai.
Ila-N children-N take care of-HAB be-PR
Ila babysits (as a job).

Now, given the NI structure $[\mathrm{N\ V}]_{\mathrm{V}}$, it should be possible to attach *nahĩĩ* to the left of the structure, as in (71) below, and get sentential negation with reading B. However, the only interpretation available to (71) is that of negation of the subject to the left of *nahĩĩ*:

(71) ?anil nahĩĩ kitaabẽ becegaa.
Anil-N(M) not book-N-PL(M) sell-FU.M
(i) Not Anil (but someone else) will do book-selling.
(ii) *Anil will not do book-selling.

In order to explain why (71) cannot be an instance of sentential negation, we must take a detour to look at the facts of *nahĩ* in complex predicates in Hindi.

5.2.2. *Negation in Complex Predicates*

Hindi has a construction, illustrated in (72), in which a noun combines with a verb to form a complex predicate:[40]

(72)a. anil-ne ilaa-ko **yaad** **kiyaa**.
Anil-E(M) Ila-A(F) memory-N(F) do-PERF.M
Anil remembered Ila.

b. anil-ne ilaa-kii **prašamsaa** **kii**.
Anil-E(M) Ila-G(F) praise-N(F) do-PERF.F
Anil praised Ila.

The noun *yaad* "memory" in (72a), and *prašamsaa* "praise (N)" in (72b), form complex predicates (henceforth CP) with the verb *kar* "do". They are CPs because (i) they function as a single predicate semantically, and (ii) the burden of determining the number of arguments in the clause, and their case marking, is not borne by the verb alone, but is shared by the noun.

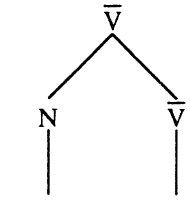
Both NI and the CPs in (72) involve a noun + verb sequence that forms a single unit. However, they differ in fundamental ways: (i) unlike the noun in NI, the noun in the CP is semantically predicative, and can contribute to the argument structure of the clause; and (ii) NI is formed in the lexical module, and CP is formed in the phrasal module. I will briefly discuss the relevant aspects of the ARG STR and GC STR of CPs below, but will not defend these structures in this paper.[41]

Evidence points to the CP in Hindi having the GC STR illustrated in (73):

[40] In the literature on complex predicates, the noun is often called the "host", and the verb the "light verb" (see Cattell 1984).

[41] A detailed analysis of the CPs in Hindi, and a defense of the structure of CPs assumed in this paper, are given in T. Mohanan (1990, chapter 8).

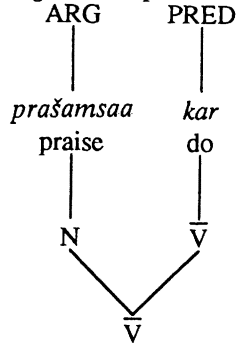
(73)



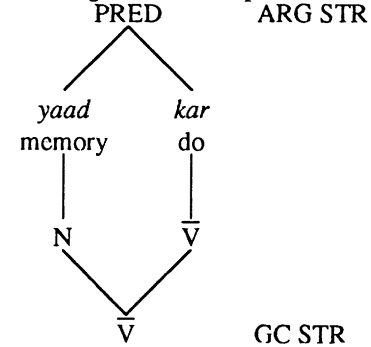
(= remember)

Now, the nominal hosts in the CP construction fall into two classes with respect to verb agreement: those the verb can agree with (e.g., (72b)), and those the verb cannot agree with (e.g., (72a)). We can formally distinguish the two kinds of N + V CPs in terms of the argument status of the N within the CP (T. Mohanan 1990). The N is an argument in (74a); the verb may agree with it. It is not an argument in (74b); the verb cannot agree with it.[42]

(74)a. *Agreement possible*[43]



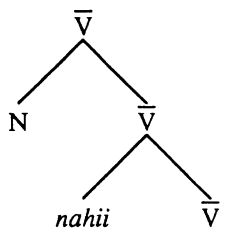
b. *Agreement not possible*

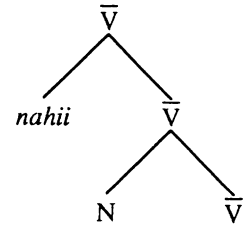


Now to turn to the facts of *nahĩĩ* placement and sentential negation in these structures. Given the structure for sentential negation in (65b), and the GC STR of CPs in (73), it should be possible to place *nahĩĩ* between the noun and verb, as in (75a), or to the left of the entire CP, as in (75b):

[42] In precisely those CPs where the verb can agree with its N host, the argument providcd by the host bears genitive, instrumental or locative case, and a non-subject non-object function. In contrast, in those CPs where the verb cannot agree with its host, the argument contributed by the host bears either ACC or NOM case, depending on whether it is animate or not, indicating that it is a primary object. The assumption that the nominal host in (74a) is an argument, while that in (74b) is not, correctly predicts this set of correlations. The N host that is an argument is the primary object of the clause ((74a)), so there cannot be another primary object. A nominal that is not an argument cannot be an object ((74b)), and the object function can be occupied by another argument.

[43] At ARG STR the predicative noun *prašamsaa* in (73a) is an ARG; at the level of semantic structure it is an independent predicate with its own semantic dependents.

(75)a. 

b. 

The structures in (75a) and (75b) are instantiated by (76a) and (76b) respectively:

(76)a. anil-ne ilaa-ko **yaad** nahĩĩ **kiyaa**.
Anil-E(m) Ila-A(F) memory-N(F) not do-PERF.M
Anil didn't remember Ila.

b. anil-ne ilaa-ko nahĩĩ **yaad** **kiyaa**.
Anil-E(M) Ila-A(F) not memory-N(F) do-PERF.M
Anil didn't remember Ila.

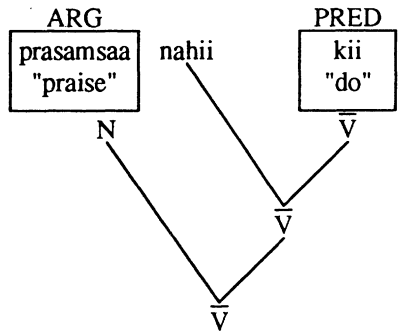
The N in the CP in (76a, b) is one that the verb cannot agree with. According to our analysis, it is not an argument of the verb ((74b)). What is interesting is that with an N that can control agreement ((74a)), *nahĩĩ* cannot be placed to left of the N for sentential negation ((77b)):

(77)a. anil-ne ilaa-kii **prašamsaa** nahĩĩ **kii**.
Anil-E(M) Ila-G(F) praise-N(F) not do-PERF.F
Anil didn't praise Ila.

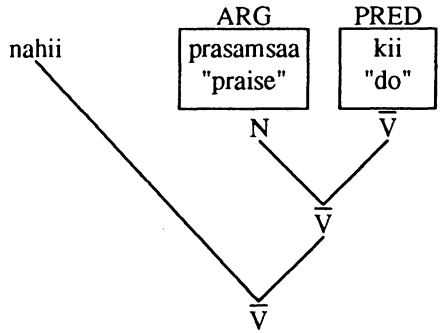
b. ?anil-ne ilaa-kii nahĩĩ **prašamsaa** **kii**.
Anil-E(M) Ila-G(F) not praise-N(F) do-PERF.F
Anil praise not Ila (but someone else).
*Anil didn't praise Ila.

We can account for the contrast between (76b) and (77b) by requiring the NEG to be adjacent to the predicate in sentential negation. Consider the relevant aspects of the structure of CPs in (76a, b) and (77a, b), given in (78a–d):

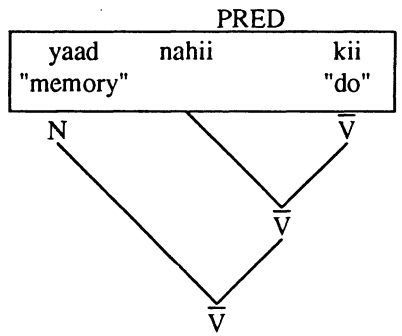
(78)a.


ARG
prasamsaa
"praise"
nahii
PRED
kii
"do"
N
V
V
V
ARG STR
GC STR

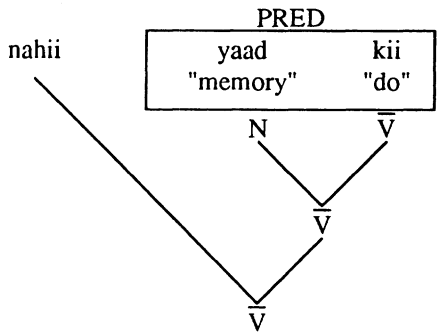
b. *


nahii
ARG
prasamsaa
"praise"
PRED
kii
"do"
N
V
V
V
ARG STR
GC STR

c.


PRED
yaad
"memory"
nahii
kii
"do"
N
V
V
V
ARG STR
GC STR

d.


nahii
PRED
yaad
"memory"
kii
"do"
N
V
V
V
ARG STR
GC STR

In each of the representations in (78), *nahĩĩ* is attached to the left of the V, as required by (65b). However, in the structure that corresponds to the ungrammatical sentence ((78b)), *nahĩĩ* is not adjacent to the PRED. In (78a) and (78d), it is clearly adjacent to the PRED. In (78c), it is not non-adjacent, and therefore, the adjacency condition is not violated.

The two conditions for sentential negation can then be stated as in (79):

(79) In sentential negation:
(a) NEG is attached to the left of the verb in GC STR (=(65b));
(b) NEG must be adjacent to the PRED in ARG STR/GF STR.

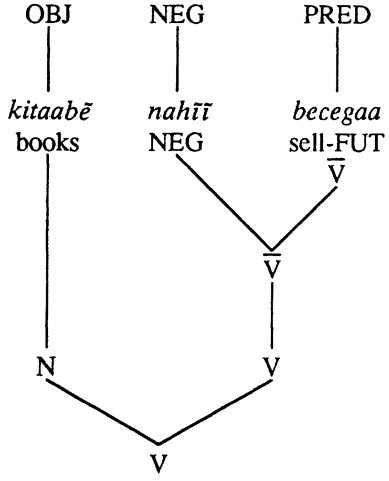
(79a) is required to account for the ungrammaticality of (70b), and (79b) to account for the ungrammaticality of (77b). If both (79a) and (79b) are necessary in the grammar of Hindi, then the dual representations in (55) and (78) are also necessary.

5.2.3. *Absence of Sentential Negation in NI: An Explanation*

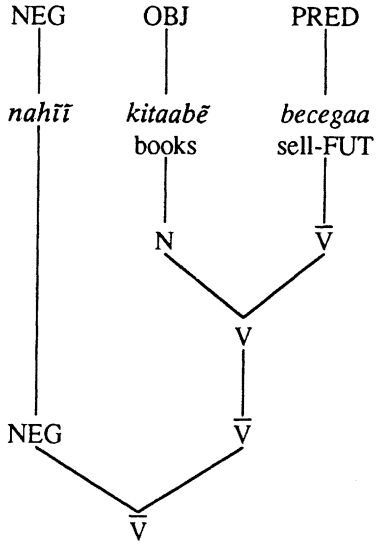
Thus, NEG placement in Hindi is sensitive to two distinct representations. The dual representation of NI along the dimensions of ARG/GF STR and GC STR ((55)), coupled with the independently required conditions on NEG placement in (79a) and (79b), correctly predicts an interesting syntactic gap: sentential negation is impossible in NI in Hindi.

As explained earlier, the NEG cannot occur between the N and V in the NIC, because it would create the ill-formed GC STR: this explains why sentential negation is impossible in (69) and (70b).

(80)



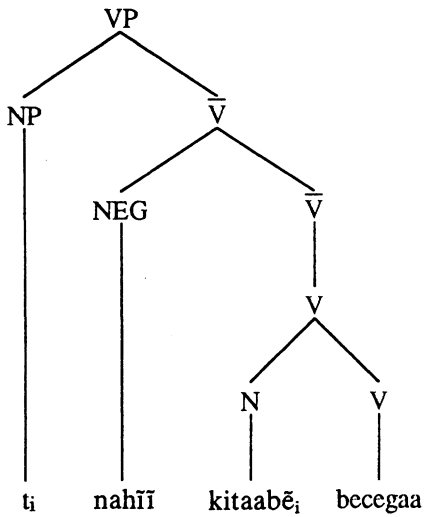
The NEG cannot occur to the left of the NI either, because it would create an ill-formed GF STR as in (77b), which violates (79b). This explains why sentential negation is impossible in (71):

(81) * 

5.2.4. *The Need for Factorization: GC STR vs. GF STR*

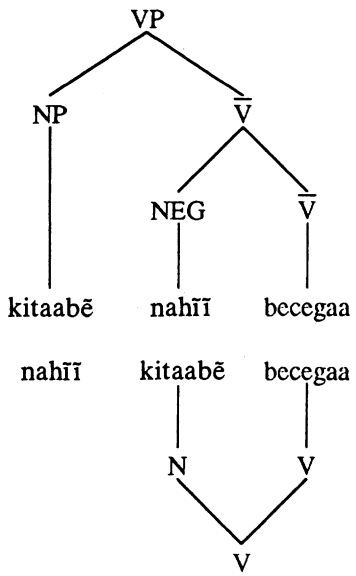
Our explanation for the impossibility of sentential negation in Hindi NIC is not available to syntactic theories which do not separate categorial structure and grammatical function structure as two dimensions of syntactic representation. Note that both (79a) and (79b) are statements on *syntactic* representations. Baker and Sadock distinguish between a syntactic representation in which the incorporated nominal is an NP of VP, and a morphological representation in which it is an N of V, but they do not distinguish between the representation of grammatical categories and the representation of grammatical functions. If objecthood is represented as NP of VP, then it is difficult to see what prevents the placement of NEG to the left of the NI in sentential negation, as in (82a) (head movement) and (82b) (coanalysis):

(82)a.



Head movement

b.



Coanalysis

As a result, the explanation for the impossibility of sentential negation in terms of (79a, b) is not available to them.

6. The Concept of Wordhood

Most debates on the Lexicalist Hypothesis (Chomsky 1970, Jackendoff 1972, Wasow 1977, Bresnan 1978, Lieber 1980, K. P. Mohanan 1982, Kiparsky 1982, Sproat 1986, Baker 1985, Ackerman 1987, Ackerman and

Webelhuth 1992, Bresnan and Mchombo 1993, and so on) center around the relation between linguistic structure within words and across words. These debates raise the question: "What is a word?"

Recent research has clearly shown that the term 'word' conflates a number of notions that must be kept distinct. If we use the term PHONOLOGICAL WORD (P-WORD) to refer to units of a phonological domain, then there are two types of phonological words, METRICAL WORD and PROSODIC WORD. Let us use the term SYNTACTIC WORD (S-WORD) to refer to wordlike units for the purposes of syntax. The expressions *John's* and *isn't* in (83a, b) are single P-WORD's, but not single S-WORD's:

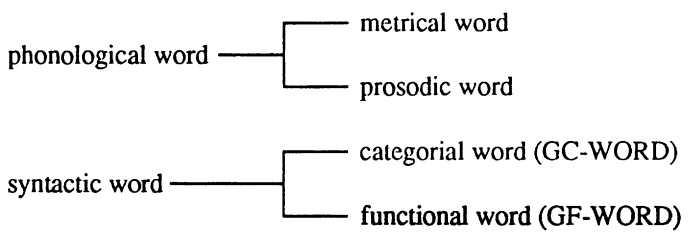
(83) a. *John's* coming here tomorrow.
b. *Isn't* he being promoted?

The representation of (83a) as (84) brings out the non-identity of P-WORD and S-WORD:

(84)

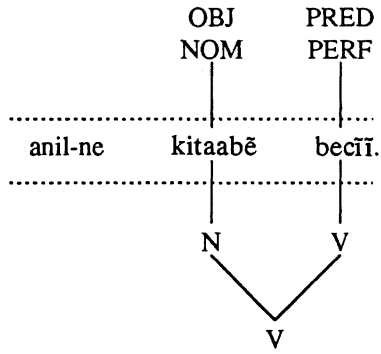
The non-convergence of the two notions of wordhood in representations like (84) are commonly acknowledged in the field. In sections 3 and 4, we saw the need to separate syntactic organization at GC-structure from syntactic organization at GF-structure. Given this factorization of dimensions, and the potential non-identity of wordhood across dimensions, it is reasonable to separate wordhood at GC-structure (GC-WORD) and GF-structure (GF-WORD). The different notions of wordhood, then, are as follows:

(85)



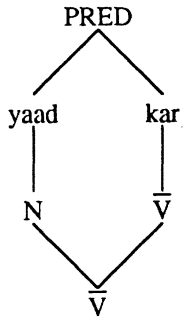
The distinction between GC-WORD and GF-WORD allows us to understand better the phenomenon of NI. In (48a), the relevant parts of which are repeated below, the N + V is a single GC-WORD (V^0). However, it is not a single GF-WORD: the N is an ARG, and the V is a PRED.

(48)a.



The structure of a type of complex predicates in Hindi in (73b) illustrates the reverse case:

(73)b.



In this complex predicate, the nominal and the verbal are two different GC-WORDS. However, together they form a single PRED. Hence they constitute a single GF-WORD.[44]

If the terms LEXICAL and SYNTACTIC refer to relations within and across words respectively, then we may say that Hindi NIC is lexical with respect to GC STR, and syntactic with respect to GF STR. Conversely, Hindi complex predicates are syntactic with respect to GC STR, and lexical with respect to GF STR. Given this picture, whether a construction is lexical or syntactic is not determinable unless the dimension of lexicality is specified.

[44] See Ackerman (1987), Ackerman and Webelhuth (1992), Matsumoto (1992) for the need to separate different notions of syntactic wordhood.

In order to avoid confusion resulting from ambiguity, I suggest that the term lexical be reserved for relations within a GC WORD. If this suggestion is accepted, then Hindi NIC is lexical, while Hindi complex predicates are not.

As discussed in section 4, the term lexical also involves an ambiguity between a modular concept (lexical module) and a representational concept (lexical category). It was our claim that Hindi NI is lexical in both these senses. These two interpretations of lexical come together if we accept the strong version of the Lexicalist Hypothesis, which I would like to state as follows:

(86) Lexical categories cannot be created in the phrasal (=postlexical) module, and phrasal categories cannot be created in the lexical module.

7. Concluding Remarks

My goal in this paper has been to provide a detailed grammar of the syntactic behaviour of NI in Hindi, and to outline a partial syntactic theory that yields the grammar. The puzzle of Hindi NI is its dual behavior with respect to lexicality. Fundamental to my account is the factorization of two different types of syntactic information into two dimensions of structuring, GC STR (category information) and GF STR (grammatical function information). The N + V compound is composed of two elements which form a single morphological unit at GC STR, but retain their syntactic independence at GF STR. The analyses of NI in both Sadock and Baker recognise the dual behaviour of NI. Their representations, however, do not attempt to separate two kinds of syntactic representations. As a result, the absence of modifier stranding, gapping, and sentential negation in Hindi NIC have no analysis in these theories.

A natural consequence of the factorization is that a wordlike unit at one level is not necessarily a wordlike unit at another level. This nonconvergence of wordhood at different levels has been the source of a number of controversies in linguistic theory, including debates surrounding the lexicality of NI. The separation of GC STR and GF STR, together with the recognition of the different notions of wordhood, leads to the issue of the relation between wordhood and lexicality. I have claimed in this paper that the notion of wordhood relevant for lexicality is the GC WORD. This proposal retains the basic insight underlying the lexicalist hypothesis by formulating it as a relation between a lexical category and

the lexical module. In many of the debates on lexicality, arguments for the lexicality of a construction are based on information that in our framework belongs in GC STR; arguments for the non-lexicality of the construction are based on information that belongs in GF STR. Given the perspective outlined above, many of the points of debate become non-issues.

References

Abney, S. P.: 1987, *The English Noun Phrase in its Sentential Aspect*, unpublished Ph.D. dissertation, MIT.

Ackerman, F.: 1987, *Miscreant Morphemes: Phrasal Predicates in Ugric*, unpublished Ph.D. dissertation, University of California, Berkeley.

Ackerman, F., and G. Webelhuth: 1992, Topicalization and German Complex Predicates, unpublished manuscript, Department of Linguistics, University of California in San Diego.

Allen, B., D. Gardiner and D. Frantz: 1984, 'Noun Incorporation in Southern Tiwa', *IJAL* **50**, 292–311.

Alsina, A.: 1993, *Predicate Composition: A Theory of Syntactic Function Alternations*, unpublished Ph.D. dissertation, Stanford University.

Baker, M.: 1985, *Incorporation: A Theory of Grammatical Function Changing*, Doctoral dissertation, MIT.

Baker, M.: 1988, *Incorporation: A Theory of Grammatical Function Changing*, The University of Chicago Press, Chicago.

Bauer, L.: 1983, *English Word-formation*, Cambridge University Press, Cambridge.

Bresnan, J.: 1978, 'A Realistic Transformational Grammar', in M. Halle, J. Bresnan and G. A. Miller (eds.), *Linguistic Theory and Psychological Reality*, MIT Press, Cambridge, Massachusetts.

Bresnan, J. (ed.): 1982a, *The Mental Representation of Grammatical Relations*, MIT Press, Cambridge, Massachusetts.

Bresnan, J.: 1982b, 'The Passive', in J. Bresnan (ed.) (1982a), pp. 3–86.

Bresnan, J., and S. Mchombo: 1993, 'The Lexical Integrity Principle: Evidence from Bantu', ms. Stanford University and University of California at Berkeley, California.

Chomsky, N.: 1970, 'Remarks on Nominalization', in R. Jacobs and P. Rosenbaum (eds.), *Readings in English Transformational Grammar*, The Hague, Mouton.

Chomsky, N.: 1981, *Lectures on Government and Binding*, Foris, Dordrecht.

Cole, J.: 1990, 'The Minimal Word in Bengali', in A. Halpern (ed.), *The Proceedings of the Ninth West Coast Conference on Formal Linguistics*, CSLI, Stanford, California, pp. 157–170.

Di Sciullo, A. M., and E. Williams: 1987, *On the Definition of Word*, MIT Press, Cambridge, Massachusetts.

Gair, J. W., and K. Wali: 1989, 'Hindi Agreement as Anaphor', *Linguistics* **27**, 45–70.

Hale, K.: 1983, 'Warlpiri and the Grammar of Nonconfigurational Languages', *Natural Language and Linguistic Theory* **1**, 5–48.

Hale, K. and S. J. Keyser: 1991, *On the Syntax of Argument Structure*, Center for Cognitive Science, MIT, Cambridge, Massachusetts.

Hopper, P. J. and S. A. Thompson: 1984, 'The Discourse Basis for Lexical Categories in Universal Grammar', *Language* **60**, 703–52.

Inkelas, S.: 1989, *Prosodic Constituency in the Lexicon*, Doctoral dissertation, Stanford University, Stanford, California.

Jackendoff, R.: 1972, *Semantic Interpretation in Generative Grammar*, MIT Press.
Kachru, Y., B. Kachru and T. Bhatia: 1976, 'The Notion 'Subject': A Note on Hindi-Urdu, Kashmiri and Panjabi', in M. K. Verma (ed.), *The Notion of Subject in South Asian Languages*, South Asian Studies, University of Wisconsin, Madison Wisconsin, pp. 79–108.
Kiparsky, P.: 1982, 'Lexical Phonology and Morphology', in I.-S. Yang (ed.), *Linguistics in the Morning Calm*, Hanshin, Seoul.
Knecht, L.: 1986, 'Lexical Causatives in Turkish', in D. Slobin and K. Zimmer (eds.), *Studies in Turkish Linguistics*, Benjamins, Amsterdam, pp. 93–122.
Liberman, M. and A. Prince: 1977, 'On Stress and Linguistic Rhythm', *Linguistic Inquiry* **8**, 249–336.
Lieber, R.: 1980, *On the Organization of the Lexicon*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
Mahajan, A.: 1989, 'On the A/A-bar Distinction: Scrambling and Weak Crossover in Hindi', ms., MIT, Cambridge, Massachusetts.
Matsumoto, Y.: 1992, *On the Wordhood of Complex Predicates in Japanese*, Doctoral dissertation, Stanford University, Stanford, California.
McGregor, R. S.: 1972, *Outline of Hindi Grammar*, Clarendon Press, Oxford.
Merlan, F.: 1976, 'Noun Incorporation and Discourse Reference in Modern Nahuatl', *IJAL* **42**, 177–191.
Mithun, M.: 1984, 'The Evolution of Noun Incorporation', *Language* **60**, 847–895.
Mohanan, K. P.: 1982, *Lexical Phonology*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
Mohanan, K. P.: 1984, 'Lexical and Configurational Structures', *The Linguistic Review* **3**, 113–139.
Mohanan, K. P.: in press, 'Organization of the Grammar', in J. Goldsmith (ed.), *A Handbook on Phonological Theory*, Chicago University Press.
Mohanan, T.: 1990, *Arguments in Hindi*, Doctoral dissertation, Stanford University, Stanford, California.
Mohanan, T.: 1993, 'Case Alternation on Objects in Hindi', *South Asian Language Review* **3**,. 1–30.
Reimsdijk, H. van, and E. Williams: 1981, 'NP-structure', *The Linguistic Review* **1**, 171–217.
Rosen, S. T.: 1989, 'Two Types of Noun Incorporation: A Lexical Analysis', *Language* **65**, 294–317.
Sadock, J. M.: 1980, 'Noun Incorporation in Greenlandic: A Case of Syntactic Word Formation', *Language* **57**, 300–319.
Sadock, J. M.: 1985, 'Autolexical Syntax: A Proposal for the Treatment of Noun Incorporation and Similar Phenomena', *NLLT* **3**, 379–440.
Sadock, J. M.: 1986, 'Some Notes on Noun Incorporation', *Language* **62**, 19–31.
Sadock, J. M.: 1991, *Autolexical Syntax: A Theory of Parallel Grammatical Representations*, The University of Chicago Press, Chicago.
Simpson, J.: 1983, *Aspects of Warlpiri Morphology and Syntax*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
Simpson, J.: 1991, *Warlpiri Morpho-syntax: A Lexicalist Approach*, Kluwer Academic Publishers, Dordrecht.
Sproat, R.: 1986, *On Deriving the Lexicon*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
Srivastava, M.: 1969, *The Elements of Hindi Grammar*, Motilal Banarsidass, India.
Verma, M. K.: 1971, *The Stucture of the Noun Phrase in English and Hindi*, Motilal Banarsidass, India.
Wasow, T.: 1977, 'Transformations and the Lexicon', in P. Culicover, T. Wasow, and A. Akmajian (eds), *Formal Syntax*, Academic Press, New York, pp. 327–360.

Zubizaretta, M-L.: 1987, *Levels of Representation in the Lexicon and the Syntax*, Foris, Dordrecht.

Received 15 August 1991
Revised 31 March 1993

Linguistics Programme/Dept. of English Language & Literature
National University of Singapore
Kent Ridge, Singapore 0511
elltaram@nusvm.bitnet