

Paradigmaticization: A Case Study from South Asia

Author(s): Peter Edwin Hook

Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society (1988), pp. 293-303

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

Paradigmaticization: A Case Study from South Asia*

Peter Edwin Hook
University of Michigan

"Paradigmaticization" refers to that late stage in grammaticalization during which the distribution of an evolving grammatical element becomes general enough for it to enter into the system of options that characterize all (or nearly all) the members of some major word-class. On the basis of data from Indo-Aryan languages I show that this stage may exhibit what at first seem to be contradictory properties: While the text-frequency of the paradigmaticizing element shows a dramatic increase overall, the syntactic environments in which it may occur actually decrease in number and variety.

This finding emerges from a family-wide study of a class of auxiliaries in Indo-Aryan, all of which are homophonous with main verbs in their respective languages: GO, GIVE, TAKE, THROW, LEAVE BEHIND, PUT, etc. A "compound verb" (CV) comprises the finite form (FF) of one of these following the non-finite (NF) form of a primary verb:

1. compound (= primary + aux) vs. non-compound:

Kashmiri:	Tyin khuul-yith dyi		Tyin khoon <1>
Hindi-Urdu:	Tin khol d-o	~	Tin khol-o <2>
Gujarati:	Tin khol-i d-o	~	Tin khol-o
Marathi:	Tin ughaD-un dy-aa	~	Tin ughaD-aa
	can open-NF GIVE-FF	~	can open-FF
	'Open the can!'	~	'Open the can!'

As a syntactic structure the compound verb is an innovation in Indo-Aryan <3>, one that has evolved further in some languages than in others. This is reflected in wide variation in its text frequency: In some languages (Hindi-Urdu, Panjabi) it is encountered in running text up to six times more frequently than in other languages (Marathi, Kashmiri). The full gamut of frequencies between these two extremes is represented:

Relative text-frequency of compound verbs in NIA: <4>

Shina (Gilgit)	0	
Kashmiri	1	--
Marathi	2	----
Rajasthani	3	-----
Gujarati	4	-----
Panjabi	5	-----
Hindi-Urdu	6	-----

Along with text frequencies, the differences in meaning which stem from the use of a compound as opposed to a non-compound verb vary from one language to another. For some Indo-Aryan languages it can be shown that this opposition has become part of a system of regular semantic contrasts, a part of the verbal paradigm, while in others it has not. Furthermore, while more studies need to be done, research on older forms of Bengali and Hindi indicates that the construction is becoming more general over time (Zbavitel 1970, Kapp 1972). Assuming that the direction of change is the same for all Indo-Aryan languages, a comparative study of the compound verb provides us an unusual opportunity to examine the process of paradigmaticization at different points in apparent time and to follow it though its successive phases. In the first half of this paper (one part of a more comprehensive study <5>), I consider two stages that are quite widely separated, those represented by Hindi-Urdu and Marathi. In the second part I look at an intermediate stage of development, that presently found in Gujarati.

In a language like Marathi where the overall text frequency of compound verbs is low auxiliaries retain a greater fraction of their original lexical meaning. For instance, the use of the auxiliary GIVE is limited to either situations in which outward movement is imparted to the patient of an action <6>:

2m tyaana katsraa Taak-un dilaa 'He threw out
 he-ERG trash throw-NF GAVE the trash.'

or in which the result of the action is outwardly vectored, that is, in which it makes sense to think of the action as being done for another (as opposed to being done for oneself):

3m maajhaa saaThi he kaam kar-un de-Sil kaa
 my-OBL sake this job do-NF GIVE-FUT2 QM
 'Will you do this job for me?'

In a language like Hindi-Urdu where the compound verb is of much more frequent occurrence the connection of the auxiliary qua lexeme to its primary is harder to explain or predict. Combinations like those in 2m and 3m are, of course, not excluded <7>:

2h us-ne kacraa phEk diyaa 'He threw out
 he-ERG trash throw GAVE the trash.'

3h kyaa mere liye ye kaam kar d-oge
 QM my sake this job do GIVE-FUT2
 'Will you do this job for me?'

But we cannot so easily give an explanation of GIVE's occurrence in 4h where the lexical meaning of GIVE is orthogonal to the meaning of the verb phrase as a whole:

4h mAI ne uske hoTh0 ko...TaTol-naa...Suruu kar diyaa
 I ERG her lips DAT feel-INF start do GAVE
 'I began to feel her lips.' Vaid 1970:13

nor in 5h where the semantic element of outward orientation is inherent in the meaning of the primary verb and the auxiliary appears redundant:

5h aaxirkaar us ne mujh ko paise de diye
 finally he ERG me to money give GAVE
 'At last he gave me the money'.

Even more problematic is any explanation in lexical semantic terms of compounds like those in 6h and 7h in which the literal meaning of the auxiliary verb seems to contradict or cancel the meaning of the primary:

6h is mE us ne teraa naam le diyaa
 this in he ERG your name take GAVE
 'He implicated you in this.'

7h idhar aa jaa 'Come here!'
 here come GO

In Marathi, the corresponding combinations are either not acceptable or (in the case of 7) can be interpreted only as expressing two separate actions:

4m mi titse oTh tsaatsap-na suru kela (*kar-un dila)
 I-ERG her lips touch-INF start did do-NF GAVE
 'I began to feel her lips.'

5m SevaTi tyaa na ma-laa paise dile (*de-un dile)
 finally he-ERG me-to money gave give-NF GAVE
 'Finally he gave me the money.'

6m hyaacaat tyaana tudzha naaw ghetla (*ghe-un dila)
 this-LOC he-ERG your name took took-NF GAVE
 'He implicated you in this.'

7m ikaDa ye 'Come here!' (*ye-un dzaa)
 here come come-NF GO
 (ye-un dzaa can only mean 'Come here and then go!')

Paradoxically, in CV-rich languages like Hindi-Urdu there are many more syntactic environments that exclude the compound verb altogether. Thus Marathi, with its far less frequent and less grammaticalized compound verb auxiliaries, permits them with negatives and semi-negatives (viz 'only'):

- 8m tyaa-na katsraa Taak-un dilaa naahi
 he-ERG trash throw-NF GAVE not
 'He did not throw out the trash.'
- 9m haa katsraa jaan-na-ts Taak-un dilaa
 this trash John-ERG-only throw-NF GAVE
 'John was the only one who threw out this trash.'

in construction with modal and phasal verbs:

- 10m aadz haa katsraa Taak-un de-u Sak-Sil kaa (modal)
 today this trash throw-NF GIVE can-FUT QM
 'Can (you) throw out this trash today?'
- 11m to katsraa Taak-un dy-aaylaa laaglaa (phasal)
 he trash throw-NF GIVE-INF began
 'He began to throw out the trash.'

and in the forms of past and conjunctive participles:

- 12m nuktyaa-ts ho-un gel-el-yaa goSti...
 recently-EMP happen-NF WENT-PP-fpl things
 'things that occurred very recently...' Kolhatkar 4
- 13m to katsraa Taak-un dy-un aat gelaa
 he trash throw-NF GIVE-CP inside went
 'He threw out the trash and went inside.'

while Hindi-Urdu, where the compound verb is from three to four times more frequent than in Marathi, does not:

- 8h us ne kacraa nahII phEkaa (*nahII phEk diyaa)
 he ERG trash not threw not throw GAVE
 'He didn't throw out the trash.'
- 9h ye kacraa jaan ne hii phEkaa
 this trash John ERG only threw
 'John was the only one who threw out this trash.'
 (*ye kacraa jaan ne hii phEk diyaa
 this trash John ERG only throw GAVE)
- 10h kyaa aaj ye kacraa phEk sak-oge
 QM today this trash throw can-FUT
 'Can (you) throw out this trash today?'
 (*kyaa aaj ye kacraa phEk de sak-oge
 QM today this trash throw GIVE can-FUT)
- 11h vo kacraa phEk-ne lagaa (*phEk de-ne lagaa)
 he trash throw-INF began throw GIVE-INF began
 'He began to throw out the trash.'
- 12h haal-hii-mE huii baate usko yaad nahII rahtII
 recently happened-fpl things him memory NEG remain
 'He doesn't remember things that happened recently.'
 (*haal-hii-mE ho gaii baate...
 recently happen GONE-fpl things...)
- 13h vo kacraa phEk-ke andar gayaa (*phEk de-ke)
 he trash throw-CP inside went throw GIVE-CP
 'He threw out the trash and went inside.'

These reduced privileges of external co-occurrence are not in contradiction with, but are rather a consequence of the greater freedom in internal co-occurrence relations shown by the Hindi-Urdu compound verb as compared to Marathi's. They follow from the presence of general, across-the-board, paradigmatic values (perfectivity, relative tense, etc.) which the opposition of compound to non-compound forms has acquired in CV-rich languages like Hindi-Urdu.

Traveling north over the thousand or so miles that separate Bombay from Delhi, one encounters languages and dialects with greater and greater compound verb frequencies. Marathi in the south and Hindi-Urdu in the north are at both geographical and, with respect to the CV construction, evolutionary extremes. Applying diagnostic tests to the Hindi-Urdu CV, we can demonstrate that (among other functions) it stands in opposition to the corresponding simple verb (SV) in a system of contrasts that include 1) relative tense and 2) perfective aspect: In a context that explicitly contrasts the time of occurrence of a pair of actions, the CV must be used to express the prior action and the SV, the later one:

- 14h jab tak aap yahAA aae us ne mujhe ciT de dii thii
 when by you here came he ERG me note give GIVEN had
 'By the time you got here he had given me the note.'
 15h jab tak aap ne mujhe ciT dii vo yahAA aa gayaa thaa
 when by you ERG me note gave he here come WENT had
 'By the time you gave me the note he had come here.'

The distribution of compound and simple forms in 14h and 15h is controlled by the relative location in time of the two actions expressed. In 14h and 15h the CV/SV opposition functions to co-express <8> relative tense.

Independent of (and at times in conflict with) this the opposition of compound and simple forms also functions to express perfective aspect in the sense which that term has in Slavistics. Since discussion of aspect in languages which do not have it (like English) tends to generate unsatisfyingly discursive explanations with recourse to visualization and impressionistic analogies (when those explanations fail), I will simply note that in aspectological studies of Russian it has been noted that clauses dependent on expressions of fear and anxiety show a preference for perfectives that borders on being categorical (Forsyth 1970:258-61, 297):

- 16r mat' bojalas' kak-by ee syn ne zabolet
 mother feared lest her son NEG took-sick-PV
 'The mother was afraid that her son might get sick.'

There is a similar (although not quite so unequivocal) preference for perfective forms in clauses dependent on expressions that mean 'until':

17r zvonite poka ne otvetyat
sound while NEG answer-PV
'Keep ringing until they answer.' (Forsyth 1970:133)

These asymmetries in distribution of aspectual forms in Russian have precise counterparts in Hindi-Urdu:

18h mujhe Dar thaa ki kahII tum use ciT na de do
to-me fear was that lest you him note NEG give GIVE
'I was afraid that you might give him the note.'

19h tum yahAA Thaharo jab-tak vo tumhE ciT na de de
you here stay until he you note NEG give GIVE
'Wait here until he gives you the note.'

Applying these same tests to Marathi we find that its CV/SV opposition does not regularly express either relative tense <9> or perfectivity:

14m tu ye-Nyaacaa agodar tyaa-na ma-laa patr dila hota
you come-INF before he-ERG me-to letter gave was
'By the time you got here he'd given me the letter.'

15m tu patr de-Nyaacaa agodar haa itha aalaa hotaa
you letter give-INF before he here came was
'By the time you gave me the letter he'd come here.'

18m ma-laa kaaLjii hoti ki tu tyaa-laa patr de-Sil
me-to anxiety was that you him-to letter give-FUT
'I was afraid that you might give him the letter.'

19m to tu-laa patr de-i paryant itha thaamb
he you-to letter give-INF until here wait
'Wait here until he gives you the letter.'

In a CV-poor language like Marathi, where each auxiliary preserves more of its lexical semantics, it plays no role in any system of general semantic oppositions that determines its presence or absence. That is, it has a tendency to form a tighter bond with its primary and to co-occur with it in all environments open to the primary itself. For the CV-construction to participate in a system of contrasts that includes nearly every lexical verb in their domain it must be possible to find a CV form for nearly every lexical verb. This is not possible in Marathi, especially for intransitive primaries. Hence, the lack of reactivity in 14m through 19m. Since the compound forms that do occur in Marathi do not have paradigmatic values, it is not surprising that they are less sensitive to their grammatical environments and may occur in a variety of contexts (viz, 8

through 13) which are closed to their structural counterparts in Hindi-Urdu.

Hindi-Urdu and Marathi are at fairly extreme points along the scale of paradigmaticization postulated here. Let us look at a language whose CV's are intermediate in frequency: high enough to enter into a system of paradigmatic contrasts but still not so high as Hindi-Urdu's. Do the CV's in such a language have a greater degree of flexibility in their privileges of (external) co-occurrence than Hindi-Urdu's CV's have?

Gujarati, spoken in an area to the north of Marathi and to the south of Hindi-Urdu, has a CV-construction that meets this description: A compound form exists for nearly every lexical verb and by the tests in 14, 15 and 18 the SV/CV opposition participates fairly systematically in the expression of relative tense and of perfective aspect <10>:

- 14g tame aavyaa te pehlAA eNe paisaa aapi didhaa hataa
 you came that before he-ERG money give GAVE was
 'He had given (me) the money by the time you came.'
- 15g tame mane paisaa aapyaa te pehlAA e aavi gayo hato
 you me-to money gave that before he come WENT was
 'He had come by the time you gave me the money.'
- 18g mane bhik hati ke tame kadaac ene paisaa aapi do
 me fear was that you maybe him-to money give GIVE
 'I was afraid that you might give him the money.'

However, while adherence to the distribution of CV/SV forms shown in 14, 15 and 18 is nearly one hundred per cent for speakers of Hindi and Urdu and close to zero for speakers of Marathi, speakers of Gujarati and its dialects, (on the basis of a sample of some eighteen speakers) observe the same patterns only about 75% of the time. That is, the CV/SV opposition is well on its way to acquiring paradigmatic values in Gujarati but is not yet used with perfect regularity by every speaker. As a correlate of this incomplete incorporation of the CV/SV opposition into the verb paradigm, we may note another peculiarity of the Gujarati data which may be explained as a shift from external to internal flexibility in privileges of occurrence. In Hindi-Urdu, CV's do not in general <11> occur in negative environments. It is possible to see in this a reflection of their role in expressing anteriority and perfectivity: The time of the non-occurrence of an action is not often contrasted with the (later) time of some other action. Nor is complete or holistic conceptualization likely to be as important in talking about actions that do not occur as it is in talking about those that do. In Hindi-Urdu a disfavoring effect may be observed not

only for prototypically negative particles like nahII 'not' and mat 'do not' (prohibitive particle) but for semi-negatives like sirf...hii 'only' and Saayad hii 'hardly' as well:

- 20h buuRhaa haathii mar gayaa 'The old elephant died.'
old elephant die WENT
21h koi haathii nahII maraa 'No elephant died.'
any elephant NEG died
22h sirf ek hii haathii maraa 'Only one elephant died.'
only 1 EMP elephant died

In 21h and 22h we see the disfavoring effect of a negative and a semi-negative on a primary which (perhaps because of its inherent semantics) has a strong preference in Hindi-Urdu for manifestation as a CV. In Marathi, the same predicate almost always shows up as a SV and 20, 21 and 22 do not show us anything. In Gujarati, however, where the paradigmaticization of the CV/SV opposition is almost but not quite complete, we find a less sensitive reaction to negative contexts. While nearly every speaker uses the SV with full negatives like nathi 'not' more than half of them prefer the CV with semi-negatives:

- 20g gharDo haathi mari gayo 'The old elephant died.'
old elephant die WENT
21g ekei haathi maryo nathii 'No elephant died.'
even-one elephant died NEG
22g fakt ek-aj haathi mari gayo 'Only 1 elephant died.'
only 1-EMP elephant die WENT

In the first part of this paper we saw that the CV construction in a CV-rich language like Hindi-Urdu has poorer privileges of occurrence than it does in a CV-poor language like Marathi. This difference is understandable if we think of the less grammaticalized CV auxiliaries in Marathi as being themselves more like lexical verbs, while the more grammaticalized ones in Hindi-Urdu are less like their homophonous primaries, more like modals. As such they do not occur in places where other modals (eg, sak 'can', caahiye 'should' and paR 'have to') do not occur, while the CV auxiliaries of Marathi (in general) may occur wherever their homophonous counterparts among primary lexical verbs occur. This, of course, is a coarse kind of explanation, true in general but not applicable to the graded differences in sensitivity to negation shown by a three-way comparison of the compound verb in Hindi-Urdu, Marathi, and Gujarati (particularly since the use of modals is not at all disfavored by negation). To fully understand

that sensitivity to negation we would need to look at the semantics of relative tense and perfective aspect in more detail than space here allows <12>.

Summary: On the basis of the evolution of the CV in Indo-Aryan, I propose that there is a correlation among the following three concomitants of paradigmaticization: (1) a dramatic increase in textual frequency, (2) a universalization of privileges of occurrence vis-à-vis some major word class, and (3) a reduction in privileges of occurrence with respect to some inventory of constructions.

Notes

*. I gratefully acknowledge the organizational and the financial support of the American Institute of Indian Studies, the U.S. Office of Education (Fulbright-Hays) and the Horace H. Rackham School of Graduate Studies (University of Michigan) for studies in the field which led to the findings presented here. I am also indebted to the insightful comments and suggestions of friends and colleagues, in particular, to those of John Myhill and Madhav Deshpande.

<1>. Standard Kashmiri word order has been altered here toward the South Asian norm in order to facilitate the comparison of its compound verb with those of other NIA languages. (Kashmiri is V-2 in most types of clauses.)

<2>. Transcription systems used for data from Marathi, Gujarati, Kashmiri, Hindi-Urdu and other NIA languages are those generally found in the Indo-Aryan linguistics literature except that contrastive length in vowels is shown by doubling the symbol (not by macron or colon), nasality in vowels is shown by capitalization (not by tilda or following capital N), retroflexion of apical stops and flaps is shown by capitalization (not by a sublinear dot), and palatal fricatives are represented by capitalization of the symbols for the corresponding dental fricatives (not by diacritics).

<3>. The earliest attestations of the compound verb in Indo-Aryan are to be found in Buddhist Pali, in Jataka Tales composed in Sri Lanka during the early centuries of the Christian era: Hook 1977:336. It is also found in the philosophical writing of Buddhaghōṣa (example A located by Madhav Deshpande in sumangalavilāsinī:180):

A yam yam virujjhati tam tam ujukam katvā denti
 what what conflicts that that straight make-NF GIVE
 '(True scholars) correct whatever is contradictory.'

It may have come into Indo-Aryan as a calque on similar (and more ancient) structures in Dravidian or Altaic: see Hook 1987:163-4.

<4>. Both names and numerical values are very approxi-

mate. A map of higher resolution incorporating a 250 by 250 kilometer grid is in Hook 1977.

<5>. Presented at the Symposium on Grammaticalization at the University of Oregon, May 1988.

<6>. Marathi data confirmed by Madhav Deshpande. Abbreviations used in this paper are as follows:

2.....second person
CP.....conjunctive participle marker
CV.....compound verb
DAT.....dative case or postposition
EMP.....emphatic particle
ERG.....ergative case or postposition
f.....feminine gender
FUT.....future tense
INF.....infinitival form
LOC.....locative case or postposition
NEG.....negative particle
NF.....non-finite form
OBL.....oblique case
pl.....plural number
PP.....past participle
PV.....perfective form
QM.....yes/no question marker
SV.....simple (=non-compound) verb

<7>. Hindi-Urdu data confirmed by Afroz Taj.

<8>. I use the term co-express because the past perfect also expresses anteriority of action here. Other situations that pattern like 14h-15h exist in which perfect tenses cannot be used (Ah is from Nespital 1981:311):

Ah is-se pahle ki vo tumhE piite biicbacaav kar dEge
this before that he you beat rescue do GIVE
'Before he can hit you they'll come to your rescue.'

<9>. This is not to say that the use of the compound verb in Marathi has nothing to do with these semantic categories. On the contrary, the expression of anteriority, for example, is facilitated in Am by use of a CV:
Am mi aadhi haa ker kaadh-un gheu. mag tu vaats.

I first this dust remove-NF TAKE then you read

'First let me sweep. Then (sit down to) read.'

But a CV is not required. In fact, for many primaries (especially intransitives) there is no compound form available even when the environment is such as to favor the use of one:

Bm aapaN aadhi sTeSanaalaa pohotsu. mag fon-karu.

we first station-DAT arrive. then telephone

'Let's first get to the station and then phone.'

<10>. Gujarati data were collected in Gujarat in 1982.

<11>. See Gaeffke 1967:13-38 for discussion of negation and the Hindi-Urdu compound verb.

<12>. For discussion of the interaction of perfective aspect and negation see Hook 1974:201-229 and Forsyth 1970:241-261.

References

- Buddhaghoṣa's sumangalavilāsinī. Vol I. M. Tiwary, Ed. 1974. Nalanda: Nava Nalanda Mahavihara.
- Comrie, B. 1976. Aspect. Cambridge University Press.
- Forsyth, J. 1970. A Grammar of Aspect: Usage and Meaning in the Russian Verb. Cambridge University Press
- Gaefcke, P. 1967. Untersuchungen zur Syntax des Hindi. The Hague: Mouton and Co.
- Givón, T. 1984. Syntax: a Functional-Typological Introduction. Amsterdam: J. Benjamins.
- Heine, B., U. Claudi and F. Hunnemeyer. Ms. From Cognition to Grammar: Evidence from African Languages.
- Hook, P. 1987. 'Linguistic Areas: Getting at the Grain of History'. In Festschrift for Henry Hoenigswald, On the Occasion of his Seventieth Birthday. George Cardona and Norman H. Zide, Eds. Tuebingen: Gunter Narr Verlag. Pp. 155-168.
- _____. 1978. 'Perfecting a Test for the Perfective: Aspectual Parallels in Russian, Lithuanian, Modern Greek, Hindi and Pashto'. In University of Michigan Papers in Linguistics, 2:89-104.
- _____. 1977. 'The Distribution of the Compound Verb in the Languages of North India and the Question of its Origin'. In International Journal of Dravidian Linguistics, 6:336-51.
- _____. 1974. The Compound Verb in Hindi. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.
- Kapp, Dieter B. 1972. Das Verbum paraba in seiner Funktion als Simplex und Explikativum in Jayasis Padumavati. Wiesbaden: Verlag Otto Harrassowitz.
- Kolhatkar, S. K. 1910. sudaamyatse pohe (It's the Idea that Counts). 7th edition. Pune: Modern Book Depot.
- Masica, Colin. 1976. Defining a Linguistic Area: South Asia. Chicago: University of Chicago Press.
- Nespital, H. 1981. Das Futursystem im Hindi und Urdu. Wiesbaden: Franz Steiner Verlag.
- Vaid, Krishna Baldev. 1970. duusre kinaare se... (From the other shore...). Delhi: Radhakrishna Prakashan.
- Zbavitel, Dušan. 1970. Non-finite Verbal Forms in Bengali. Prague: Czechoslovak Academy of Sciences.