The Compound Verb in Munda: An Areal and Typological Overview*

Peter Edwin Hook
University of Michigan

ABSTRACT

A study of six Munda languages shows that the syntactic category compound verb (which alternates with simple verb) may be identified in each one of them. However, while compound verbs in South Munda form systems which closely resemble those found in adjacent Indo-Aryan and Dravidian languages, North and Central Munda feature compound verbs of a very different sort. The South Munda type seems to have arisen as the result of cross-linguistic diffusion from its neighbors while that in North and Central Munda owes its origin to independent developments.

The compound verb is one of the syntactico-semantic phenomena common to most South Asian languages regardless of their genetic affiliations (Masica 1976: 141–58). It has been studied in some detail in Indo-Aryan (Pořízkova, Hook, Cardona, Zbavitel, etc.) and in Dravidian (Schiffman, Bhat, Annamalai, etc.), but has so far eluded the undivided attention of Austro-Asiaticists. Even for those languages whose compound verb systems have been analyzed with greater thoroughness the precise definition of the category has been a subject of uncertainty and controversy. Since I have taken an active part in trying to resolve such controversies vis-à-vis the compound verb in Indo-Aryan, I have some reason to hope that I may be able to make a contribution to a horizontal study of the phenomenon within the languages of the Munda family, even if I am not a Mundaist.

The present paper may be divided into four parts. First, using data from Hindi, I give a stipulative definition of the compound verb that I believe to be maximally effective in isolating corresponding constructions in other languages. Second, I apply this definition to data from three South Munda languages (Gtaq, Gutob, and Remo). Third, I describe some specific parallels between the South Munda compound verb and that of adjacent Indo-Aryan and Dravidian speech forms. Fourth, I take a brief look at published information on the compound verb in Central Munda (Kharia) and
North Munda (Santali and Mundari), for it is here—if anywhere—that there is evidence for the compound verb’s being indigenous to Munda rather than borrowed from Indo-Aryan and/or Dravidian.

(I) For the purpose of typological comparison I define the compound verb as a polyverbal sequence composed of a main or ‘lexical’ verb and one (or sometimes more than one) auxiliary verb such that: (1) they are homophonous with main verbs; and (2) they alternate with their absence. Thus, in:

(a) bijalī ā gāī
   electricity come WENT5
   ‘The electricity’s come on?’

(b) cālū kār do ɵb
   on do GIVE now
   ‘Put it on now.’

the forms kār and ā (‘do’ and ‘come’) are the main verbs bearing the main semantic load of the verb phrase in each sentence. The forms do and gāī are the auxiliaries (or, more precisely, the explicators of Masica 1976 or the vectors of Pray 1970 and Hook 1974). These have their homophonous counterparts among the main verbs of Hindi:

(c) bijalī gāī
   electricity went (gāī < jā ‘go’)
   ‘Electricity’s gone?’

(d) bil ke paise do, ɵ?
   bill’s money give no (do < de ‘give’)
   ‘Pay the bill, why don’t you?’

By specifying ‘homophonous with main verbs’ we exclude certain kinds of echoic formations (also very typical and useful in the establishment of the Indian linguistic area: see Emeneau 1956 and 1978):

(e) saj-daj rāhī hogī
   spruce-ECHO -ing be-FUT6
   ‘She must be getting all dressed up.’

By insisting that the auxiliary ‘alternate with its absence’ we exclude:

(1) sequences of main and modal verb or main verb and tensual auxiliary. Compare (a) and (f) with (g) and (h):
(a) bijāli ā gāi?
   electricity come WENT
   'Did the electricity come on?'

(f) bijāli āī?
   electricity come
   'Did the electricity come on?'

(g) bijāli ā sāk-egī
   electricity come can-FUT
   'The electricity may come on.'

(h) bijāli a-egi
   electricity come-FUT
   'The electricity will come on.'

or with (i) and (j):

(i) bijāli ā rāhī thi
   electricity com-ing was
   'The electricity was coming on.'

(j) bijāli āī thi
   electricity came was
   'The electricity had come on.'

Removing the modal sāk and tensual rāh has an easily observable or, better, an easily translatable effect on meaning. Removal of the vector ā, while affecting the meaning, never does so in a way that would affect the truth value of the utterance as a whole or change an English gloss of it.

(2) Another set excluded by the alternation criterion is that of serial verbs:

(k) pāpīṭā le jā-o-ge
   papaya take go-2pl-FUT
   'Will you take away the papaya?'

The components of a serial verb do not alternate with their absence. Removal of one of them, for instance, of jā 'go' in (k), yields an expression whose meaning is very different:
If one confines one's attention to a single language such as Hindi, it is possible to find a more interesting definition of the compound verb, one that shows significant parallels in semantic function and syntactic behavior among a large class of polyverbal sequences: a set of invariant unifying properties such as incollocability with negatives (Hook 1974: 98–103); expression of anteriority (Hook 1978b: 149–52), expression of perfective aspect (Pořížka 1972; 543–67; Hook 1987a), inability to express conation (Hook 1974: 163–78). However, as I have shown elsewhere (1982 and 1989), even within Indo-Aryan it is not possible to generalize the relatively rich functional definition of the compound verb that has been developed for Hindi. Therefore, we will examine the compound verb in Munda on the basis of the rather lean (and thus capacious) definition given above.

In the South Munda language Gta?, spoken in Koraput District of Orissa, polyverbal sequences appear with some frequency in the texts available to me. Sequences meeting the definition of the compound verb given above include those formed with auxiliary bi? (homophonous with bi? 'give'), we (homophonous with we 'go') and possibly bo (homophonous with bo 'put; keep'). Thus, bi? in (m):

\[(m)\] gte-la hun-dae akaen samwa bason bi?-ke7
then child-3pl this story say GIVE-PST
‘Then their child told this story.’ (MZ 2:23)

is homophonous with main verb bi? in (n):

\[(n)\] gro-gco ke samplae nae-n̄dre-hin bi?-e
officer to present our-people-pl give-FUT
‘Our people will give presents to the officers.’

Furthermore, these auxiliaries alternate with their absence:

\[(o)\] b?ba-rae bason-ke
father-3pl say-PST
‘Their father said . . . ’

\[(p)\] knwe?-rae gwe? we-ge
wife-3pl die GO-RPST
‘His wife had died.’
(q) nae remwa to gwe?-ge
our men indeed die-RPST
‘Our men indeed had died.’

In Gta?, as in Hindi-Urdu (see example k) and some other South Asian languages, sequences of main verbs can be formally indistinguishable from sequences of main verb plus auxiliary: se? pia? ‘break by tearing; rip’ or jog tlak ‘pick up and throw’ versus gwe? we ‘die-GO; die’. According to Mahapatra (1976: 818–22) these functionally different classes of verb sequences can be identified by their behavior with respect to Gta?’s rule of echo formation. If both verbs in sequence are main verbs, either one (independently of the other) can assume an echoic form:

(r) jog tlak > jog tlik, jag tlak, jig tlik, jig tlak
‘pick up (something) and throw (it somewhere)’

However, if the second verb in a sequence is a compound verb auxiliary then an echoic cannot be made from it:

(s) gwe? we > ga? we but *gwe? a, *ga? a
die GO > echoic ‘die’ (of inferior beings)

Gutob or Gadaba, another South Munda language spoken in Koraput District in Orissa, has a system which in broad outline resembles Gta?’s. Auxiliaries homophonous with main verbs include ber (cognate with Gta? be?) GIVE:

(t) uson gol-gol-te nom bobrig-o? ber-o? (DOT 19:6)
today smoothly you make-enter-PST GIVE-PST
‘Today you put it in smoothly.’ (ber as CV auxiliary)

(u) dabu be?-to-nom ki dio? loci-to-nom (DOT 14:4)
money give-HAB-2sg or free diddle-HAB-2sg
‘Do you pay or get it for free?’ (ber as main verb)

and ui GO:

(v) gol-gol-te gai-gi ui-to (DOT 20:17)
smoothly enter-PST GO-HAB
‘Smoothly (it) goes in…’ (ui as CV auxiliary)
aspatal-bo? ui-gi-nin ḍu-gu (DOT 15:14)
hospital-to go-PST-1sg be-PST
‘I had gone to the hospital.’ (ui as main verb)

Auxiliary ui can be shown to alternate with its absence:

(x) soli gai-gi ki ura? (DOT 14:20)
thing enter-PST or not
‘Did it go in? or not?’ (absence of ui; cf (v))

In addition Gutob uses sun (sū) THROW as a CV auxiliary.

(y) dio? loēi sun-to-nin (DOT 14:6)
gratis diddle THROW-HAB-1sg
‘I get it free.’

(z) bāḍ gui-da? sū-o?-nom ki ito ḍu-to (DOT 17:18)
thing wash-water THROW-PST-2sg or so stay-HAB
‘Did you wash it or (just let it) stay like that?’

In his studies of the (South Munda) Remo verb Fernandez (1967: 35–41; 1983: 28–9) lists a cognate form sun ‘throw’ as an ‘intensifier’ in ‘complex roots’ such as bulo sun ‘boil over’ beside its use as the main verbal element in ik-tan sun ‘throw away cow-dung’. Other complex roots are formed from wiy ‘go’ (bana wiy ‘forget’) and iy ‘return’ (goy iy ‘die’). Use of bed ‘give’ in Remo appears to be restricted to serial verb sequences (1983:29).

We may unhesitatingly conclude from these data that South Munda does indeed have a compound verb that satisfies the definition presented in (I). Let us now examine some of the properties of the South Munda compound verb in its areal context.

(III) The compound verb in South Mundan does not seem to crop up as often as it does in a Hindi text of the same length. In the Gta? texts available to me, I was able to find no more than thirty instances of it out of more than 500 ‘opportunities’ (i.e. about 5%). In a text from Juray, another South Munda language, there appear not to be any instances of it at all (Zide 1983). This contrasts with a frequency in Hindi (dialogue) in the neighborhood of 15%. (Such figures must remain approximate since there is a wide degree of indeterminacy in establishing just what constitutes an ‘opportunity’. See Hook 1988 and 1989 for detailed figures and discussion.) A quantitative difference of this size implies a qualitative difference of some importance in the functional role of the construction in Hindi and Gta?. In fact, there appear to be some ‘gaps’ in the Gta? system: No auxiliary plays the role played by le (homophonous
with ‘take’) in Hindi. The kind of reflexive or ‘‘ingestive’’ (see Masica 1976: 48) verb that typically prefers le Hindi is simply not found in a compound form in Gta?: con ‘eat’; salia? ‘ask for’; etc. The one exception appears with bi?:

(aa) taen bha?-ke gsu? c?cwi bi?-la  mae gwe? we-ge
that head-to dog smell GIVE-when he die GO-RPST
‘When a dog sniffed the head he died.’ (MZ 6:25)

(Here, in Hindi, we would expect sügh liyā ‘sniff TOOK’.)

The appearance of we GO, too, seems to be restricted. In Mahapatra-Zide it occurs only with gwe ‘die’ (2:4; 3:16; 4:15, 26, 30), tar ‘come out, emerge’ (1:13, 14; 2:10, 17) and ga ‘enter’ (8:2, 14). In the entire collection it occurs not even once with the verb ‘become’, while Hindi-Urdu ho jii ‘become GO’ alone accounts for about one in six compound forms in dialogue. In this the Gta? system resembles that of Marathi where compound forms of transitive verbs occur relatively more freely than do those of intransitives (see appendices in Hook, 1991).

However, on other counts the divergences of the Gta? system from that of Hindi-Urdu can be explained most easily he comparing it with that of Oriya, the Indo-Aryan language of the surrounding population. One of the features which distinguish the compound verb system of Oriya from those of other Indic languages is the infrequent use as auxiliary of nebā (equivalent to Hindi’s lenā TAKE). Ingestive verbs in Oriya take debā GIVE rather than TAKE: pi debā ‘to drink’ (cf. Hindi pī lenā) and khāi nebā ‘to eat’ (cf. Hindi khā lenā). The expected pi nebā and khāi nebā struck a native speaker (from Puri) as being North Orissan or Bengalicized. The same is largely true of sensory verbs: dekhi debā ‘to see’, suni jib5 ‘to hear’ (where jibti is homophonous with ‘to go’), etc. For other verbs where the Hindi-knower would expect TAKE there simply is no compound form: for nebā ‘to take’, no *nei nebā).

Another peculiarity distinguishing both Gta? and Remo (as well as Oriya and Marathi) from Hindi is the ability of their compound forms to occur as conjunctive participles:

(bb) tar we-ce (MZ 1:13) (cf. Hindi *nikāl jā-kār)
emerge GO-CP
‘having gone out . . . ’

(cc) len-o7 sun-o7 sit . . . Remo (Fernandez 1983:45)
thresh-PST THROW-PST CP (?)
‘having threshed . . . ’
(dd) hli?-ke cu bi?-ce wig-ke (MZ 7:12)
  bamboo-to smear GIVE-CP return-PST
  ‘Having smeared (it) on the bamboo they came on home.’
(cf. Hindi *lāgā de-kār)
  apply GIVE-CP

(ee) bhoji khā-i de-i cāli-golā Oriya
dinner eat-CP GIVE-CP went-away
  ‘Having eaten dinner he left.’
(cf. Hindi *khā de/le kār)
  eat GIVE/TAKE CP

The corresponding Hindi forms can only be nikāl-kār, lāgā-kār, khā-kār, etc. Elsewhere (Hook 1988), I have shown that the compound verb’s ability to occur in different syntactic environments has an inverse correlation with the degrees to which vectors are semantically bleached in different languages and with their overall frequency.

The compound verb systems of Gutob and of Remo are to be distinguished from that of Hindi by the prevalence of an auxiliary homophonous with the verb for ‘throw’: sun or sū. It is as frequent as auxiliary ber GIVE; whereas dāl, its equivalent in Hindi, is much less frequent and semantically more marked than Hindi de GIVE. In this Gutob and Remo seem to have undergone the influence of neighboring Dravidian languages. In fact, two of them Parji (Burrow and Bhattacharya 1963:44) and Ollari (Bhattacharya 1956:47) have the same threefold choice of auxiliaries as Gutob: GIVE, GO, and THROW, with THROW in both Parji and Ollari the most important of the three.

Of course, as persuasive as data of this sort are, we cannot usually be as confident of the direction of influence as we can be of the fact of influence. Indeed, were the compound verb in all of Munda to have the characteristics that we find for it in Gta?, and Gutob, it would not be possible to say whether it came into Munda from Indo-Aryan and Dravidian or vice versa. However, in North Munda and Central Munda (Kharia) there are compound verb systems unlike the ones encountered in Indo-Aryan, Dravidian or South Munda. Because of the isolation and divergence of these systems it is probable that they represent a compound verb proper to Munda, if not to Austro-Asiatic in general.

(IV) In his study of South Asia as a linguistic area Masica shows that a construction which appears to meet my definition of the compound verb exists not only in Indo-Aryan and Dravidian but also (inter alia) in Tajiki, Altaic and (if they are to be considered different from Altaic) Mongolian and Korean. Although he states that the compound verb exists in Munda (Masica 1976:144), he provides no examples or
evidence except for a table on p. 147 where two unusual items are listed as compound verb-forming auxiliaries in Santali: *jom* (as a main verb, 'eat') and *got?* (as a main verb, 'pluck'). These items are *sui generis*: they appear nowhere else on Masica's chart, and, oddly enough, none of the common (or, for that matter, uncommon) items found in the other languages are to be found in Santali.

Examples of these auxiliaries are available from the excellent dictionaries of Santali and Mundari compiled by Bodding and Hoffman. First, examples of Santali *got?*:

(ff) malhan god-me  
beans pluck-IMPER  
‘Pick some beans.’

(gg) hcc got?-cn-a-c  
come PLUCK-PST-IND-3sg  
‘He came quickly/suddenly.’

(hh) ol god-me  
write PLUCK-IMPER  
‘Write quickly!’

(ii) ncl got?-kcd-c-a-n  
see PLUCK-PST-3sg-IND-1sg  
‘I had a glimpse of him.’

A similar example from Mundari:

(jj) hukum namjante hiju god-me  
order on-getting come PLUCK-IMPER  
‘Come as soon as you get the order.’

A cognate form *god* ‘pluck’ is used as an auxiliary in the Central Munda language Kharia:

(kk) in ina? alsi ob-sid  
I my axe CAUS-lose PLUCK-have-1sg  
‘I have lost my axe.’

There is a second auxiliary in Santali with a function similar to that of *got?* and homophonous with a main verb in the same semantic field. This is *hot?* which as a main verb has the meaning of ‘strip’:
(II) chïtîrîc? hod-me
     switch      strip-IMPER
     'Strip the stick (of leaves).'

As a compound verb auxiliary hod expresses speed or vehemence of the action expressed by the main verb:

(mm) lai hod-me
     tell STRIP-IMPER
     'Tell (it) quickly.'

The other anomalous item in Masica’s chart, namely, jom EAT, is found in Mundari (nn) and Ho (oo):

(nn) en horoko lel jom-me
     those people see EAT-IMPER
     'Take a look at those people.'

(oo) umbul-re dub jom-pe
     shade-in sit EAT-IMPER
     'Sit (at ease) in the shade.'

This auxiliary is reported to mean ‘well’, to one’s advantage’ (Deeney 1975:67–71) and ‘for one’s benefit or comfort’ (Hoffman 7:2098). In this jom seems to parallel (at least in part) the connotations of Hindi’s le TAKE (while got? PLUCK and hot? STRIP parallel Hindi’s dâl THROW).

Every one of the twelve auxiliaries (GO, COME, RISE, etc.) in Masica’s table of “chief explicator auxiliaries” (p. 46) is found thoroughly mixed and scattered among languages of both the Indo-Aryan and Dravidian families and every one of the languages examined has at least 6 members of the set of 12 (except for Sinhalese which has only 3). The evidence for linguistic convergence here is overwhelming but the mixing has been thorough enough to make the quest for origins or direction of influence very difficult if not impossible. North and Central Munda which: (1) have auxiliaries found nowhere else and which; (2) seem not to have any of the auxiliaries found everywhere else differ sharply from South Mundan languages like Gta?, Remo and Gutob which look to be completely South Asian in their compound verb systems. This means that if there was a compound verb system in Proto-Munda it must have resembled those of North Munda and Kharia. Secondly, the fact that N. Munda and Kharia do not share any auxiliaries with the rest of South Asia (yet even so have developed a compound verb system) makes the possibility of separate (but parallel)
Compound Verb in Munda

origins and development of compound verb systems in Indo-Aryan and Dravidian seem stronger (Hook 1977, Herring ms.)

NOTES

* An earlier version of this paper was presented at the Second International Conference on Austroasiatic Linguistics (SICAL) under the title: "The Compound Verb in South Munda: A Typological and Areal Sketch." Since then, I have been able to expand its scope and the information on which it is based to include an examination of the compound verb in North and Central Munda as well. I am indebted to Norman Zide who suggested I undertake this study, who invited me to SICAL, and who gave unstintingly of his time and attention in helping me interpret some of the unpublished Munda texts that are at his disposal (DOT and MZ).

1. Which is not to say they have ignored it completely. Most descriptions of whole languages (Deeney, Burrows, Aze, Biligiri, etc.) have a paragraph or two devoted to it. But no one to my knowledge has attempted to make a comparative study before now for all of Munda.

2. An idea of the degree to which descriptions of Hindi disagree on the definition can be had from the chart in Hook 1974:19–20.


4. I do not exclude situations where the main verb can be found only in some other, cognate language.

5. Glosses in capitals (GO, GIVE, THROW, etc.) refer not to the meanings of the auxiliaries themselves but to the meanings of the main verbs with which they are homophonous.

6. Abbreviations are to be interpreted as follows:

   - **CAUS** causative affix
   - **INTR** intransitive
   - **CP** conjunctive participle
   - **MZ** Mahapatra and Zide texts
   - **DOT** DeArmonde oral texts
   - **pl** plural
   - **ECHO** echoic formation
   - **PST** past tense
   - **ERG** ergative
   - **RPST** recent past
   - **FUT** future tense
   - **sg** singular
   - **HAB** habitual aspect
   - **TR** transitive
   - **IMPER** imperative mood
   - **1** first person
   - **3** third person
   - **IND** indicative mood

7. In the Munda transcriptions <ae> represents the front low vowel. In general I have tried to normalize cited data to one uniform transcription wherever possible.

8. Although he does not say so it seems from one of Mahapatra’s examples that the compound verb auxiliary can apply to the base plus echo as a whole: gwe? ga? we 'die, etc.'
9. According to Mahapatra and to Zide (1976), Gta? echoics differ in meaning from the more widespread South Asian formations which Mahapatra (1976) refers to as 'tags'. These are formed by changing an initial consonant (or syllable) and generally express and approximation to or variation on the meaning of the base (for example, Hindi-Urdu kapre wapre 'clothing and what-not', jāl ūl gayā 'burned up, etc.'). Thus, the replacement of a base vowel with a expresses some inferiority in the subject; with i, its small size, etc.

10. Notice that ber GIVE in this example is preceded by the past tense form of the main verb whereas sun THROW in (y) is not. In Remo we find similar variation (from Fernandez 1983):

(a) bana wiy 'forget' vs ajur-o? wiy 'leap over'
(b) gwisun sun 'wash feet' vs len-o? sun 'thresh'
(c) gay sun 'enter' vs bad o? sun 'slap'

Perhaps the variation has a phonological condition with -o? appearing after all consonantal bases except those in engma. The use in compound verbs of a finite form of the main verb followed by a finite form of the auxiliary is found in some incipient systems such as that of Baluchi and some dialects of Tajiki spoken in Uzbekistan (Rastorgueva 1964: 101–2).

11. As Norman Zide has pointed out (personal communication) this restriction in Gta? cannot reflect Oriya in which the compound form of 'become' (he-i jā) is fairly common.

12. In his discussion of god in Kharia, Biligiri (1965:47) objects to taking auxiliary god as related to main verb god 'pluck' for the reason that main verb god must always take the transitive set of verb endings whereas the endings used with auxiliary god are transitive or intransitive as the main verb preceding god is transitive or not. Thus

(a) gitag-god-ki
   sleep-PLUCK-PST (INTR) 'He went to sleep.'
(b) gil-got-?og
   beat-PLUCK-PST (TR) 'He beat (someone) up.'

But this is actually no different from the situation obtaining in Hindi where the compound verb auxiliaries le and de homophonous with the transitive verbs TAKE and GIVE exhibit intransitive morphology if the main verbs are intransitive:

(c) māi aj cāl diyā hū
   I today go GAVE am 'I've set out today.'
(d) māi ne gārī cālā dī hai
   I ERG car made-go GAVE is 'I've started the car.'
(e) wo mere sāth ho li
   she me with be TOOK 'She came along with me.'
(f) us ne moze dho liye
   she ERG socks wash TOOK 'She washed her socks.'
13. Address correspondence to: Dr P. E. Hook, Program in Linguistics, The University of Michigan, 1076 Frieze Building, Ann Arbor, MI 48109-1285, U.S.A.

REFERENCES

Annamalai, E.

Aze, R. R.

Bhat, D. N. S.

Bhattacharya, S.

Biligiri, H. S.

Bodding, P. O.

Burrow, T., and S. Bhattacharya
1953 *The Parji Language, a Dravidian Language of Bastar*, Herford: Stephen Austin and Sons, Ltd.

Burrows, L.

Cardona, George

Dasgupta, Probal

Deeney, J.
Ekka, Francis

Emeneau, Murray B.
1956 "India As a Linguistic Area," *Language* 32. 3–16.

Fernandez, Frank

Grierson, George A. (compiler)

Herring, Susan

Hoffman, Rev. J.

Hook, Peter E.

Jenner, P., L. Thompson and S. Starosta (eds)


Mahapatra, K.


Masica, Colin P.


Munda, R. D.


Pinnow, H. J.


Porížka, Vincenc


Prabodhachandran Nayar, V. R.


Pray, Bruce

1970 *Topics in Hindi Urdu Grammar*. Berkeley: Center for South and Southeast Asian Studies, University of California.


Zide, Norman H.
