The diachrony of light and auxiliary verbs in Indo-Aryan

Benjamin Slade

appeared in *Diachronica* 30.4 (2013): 531–578. this is a pre-press version which differs slightly from the version appearing in the journal (the journal version actually has a number of errors introduced by the type-setting process)

University of Utah

This study examines the historical development of light verbs in Indo-Aryan. I investigate the origins of the modern Indo-Aryan compound verb construction, and compare this construction with other light verb constructions in Indo-Aryan. Examination of the antecedents of the Indo-Aryan compound verb construction alongside other Indo-Aryan light verb constructions, combined with analysis of lexical and morphosyntactic differences between the compound verb systems of two Indo-Aryan languages (Hindi and Nepali), demonstrate that light verbs are not a stable or unchanging part of grammar, but rather undergo a variety of changes, including reanalysis as tense/aspect auxiliaries.

**Keywords:** syntax, morphology, compound verbs, light verbs, auxiliary verbs, reanalysis, grammaticalisation, chain shifts, Indo-Aryan, Hindi, Nepali, Sanskrit

1. **Introduction: Indo-Aryan compound verbs**

Compound verb (CV) structures are a conspicuous aspect of the syntax of Hindi and other modern Indo-Aryan (IA) languages. Masica (1991: 326) notes that “[i]t will be found that any descriptive grammar of a [modern IA] language has a section (usually inadequate) devoted to the phenomenon most commonly known as the compound verb.” The prototypical structure of a CV in the modern IA languages consists of an ‘absolutive’ form of the verb (the main verb, often referred to in the tradition as a ‘pole’) collocated with another verb (a light verb, traditionally ‘vector’), the latter acting as a modifier of the main verb, contributing Aktionsart (usually a perfective-completive sense), attitudinal information and other semantic nuances. See the examples from Hindi in (1) and (2): the (a) examples use simplex verbal forms; the (b) and (c) examples show corresponding CV constructions.

Abbreviations are: 1, first person; 2, second person; 3, third person; ABS, absolutive; ACC, accusative; AGT, agentive; CAUS, causative; CONV, converb; CV, compound verb; EMP, emphatic; FEM, feminine; GEN, genitive; GER, gerund; HON, honorific; IA, Indo-Aryan; IMP, imperative; IMPF, imperfect; IMPF_PTCP, imperfect participle; INF, infinitive; LOC, locative; MASC, masculine; MMDHN, middle-grade honorific; NEU, neuter; NOM, nominative; OBLQ, oblique; OPT, optative; PAST, past; PERF_PTCP, perfect participle; PL, plural; PRES, present; PROG, progressive; REL, relative pronoun; SG, singular; TA_PTCP, ta-participle; VOC, vocative.

1CV constructions are found throughout the modern IA languages, as well as in Indian members of other language families, that is, Dravidian, Tibeto-Burman and Munda, as well as in Central Asian languages (Masica 1976: ch. 3). This paper focuses, however, on the development of CVs and other light verb constructions within Indo-Aryan. This does not imply that I dismiss the possibility that the IA CV system reflects at least partial convergence with other language families. For general discussion of compound verbs in Indo-Aryan (and, more broadly, in South Asian languages), see Masica 1976, Kachru and Pandharipande 1980, Singh et al. 1986, Abbi and Gopalakrishnan 1991 and many of the papers in Verma 1993; on Hindi CVs, see, amongst others, Hook 1974, Nespoli 1997; on Nepali CVs, see Sharma 1980, Pokharel 1991, Peterson 2002.

2The absolutive is a fixed, indeclinable verbal form.

3Additionally, in (1) and (3) that the use of an intransitive vector blocks the appearance of the agentive marker on the subject; see §4.1.4.
(1) a. *us-ne khānā khāy-ā*
   he/she-AGT food.GEM eat.PERF_PTCP-GEM
   “He/she ate the food.”

   b. *vah khānā khā gay-ā*
   he/she.NOM food.GEM eat.POS go.PERF_PTCP-GEM
   “He ate up the food.” (perfective)  

(2) a. *us-ne kitāb parh-i*
   he/she-AGT book.GEM read.PERF_PTCP-GEM
   “He/she read the book.”

   b. *us-ne kitāb parh dī*
   he/she-AGT book.GEM read.POS give.PERF_PTCP-GEM
   “He/she read the book aloud (i.e. for someone else’s benefit).” (other-directed action)

   c. *us-ne kitāb parh lī*
   he/she-AGT book.GEM read.POS take.PERF_PTCP-GEM
   “He/she read the book silently (i.e. for his/her own benefit).” (self-directed action)

(3) a. *maiṁ-ne kyi-ā?*
   I-AGT what do.PERF_PTCP-MISC.GEM
   “What did I do?”

   b. *maiṁ kyiā kar baithā?*
   I.NOM what do.POS sit.PERF_PTCP-MISC.GEM
   “Oh what did I do?” (regret)

Indo-Aryan light verbs employed in compound verb constructions are usually form-identical with a full/main lexical verb. For example, the light verb used with a general-purpose completive sense in (1b), *gayā (< jānā “to go”), as a main verb means “went”; likewise, *dī* as a main verb means “gave” and *lī* “took”. The examples in (2) and (3) show that, in addition to perfective semantics, light verbs may also contribute other information: in (2b) *dī* “gave” contributes a sense of ‘other-directedness’, in (2c) *lī* “took” contributes a sense of ‘self-directedness’, and in (3b), *baithā* “sat” contributes a sense of ’regret’.

CV constructions historically derive from collocations involving ‘converbs’, where a converb is a non-finite verbal form which, in contrast to CVs, denotes an independent event or state from that expressed by the finite verb of the clause, but usually shares the subject-agent of the finite verb. Converbs, which persist into modern IA, typically denote actions or states completed before the beginning of the action or state expressed by the finite verb, as in the Hindi example in (4); though sometimes they indicate simultaneous action, as in the Hindi example in (5).  

(4) *vah khānā khā-ke ghar gayā*
   he/she-NOM food.GEM eat-CONV home go.PERF_PTCP-GEM
   “(After) having eaten food, he went home.”

(5) *vah haṁs-ke bol-i ˈhāṁ*
   he/she laugh-CONV say.PERF_PTCP-GEM-FEM ‘yes’
   “(Laughing) she said ‘yes’.”

In some modern IA languages the absolutive form of the verb used in CV constructions is identical in form to the converb. For example, though in Hindi converbs are usually followed by *kar or ke*, they may also appear with a zero-ending, thus being identical in form to an absolutive, resulting in potential ambiguity, as shown by examples (6a), (6b).

4 With *khā “eat”, lenā “take” is perhaps the more commonly used light verb; e.g. *us-ne khānā khā liyā “He ate up the food”.  
5 In general, in both early and modern Indo-Aryan, converbs are often used where English would typically use a VP and VP construction, e.g. (4) might be rendered as “He ate food and went home”.  
6 On the translation of *haṁs-ke, see §3.3.”
(6) usne mujhe ek thamz-ap kharid diyā (cp. Hook 1974: 54,h12)

a. us-ne mujhe ek thamz-ap kharid diy-ā
he/she-AGT me one Thums Up buy.conv give.PERF_PTCP-MSC.SG
“He/she bought a Thums Up (soda) and gave it to me.”

b. us-ne mujhe ek thamz-ap kharid diy-ā
he/she-AGT me one Thums Up buy.INS give.PERF_PTCP-MSC.SG
“He/she bought me a Thums Up (soda) (for my benefit).”

The morphological similarity or identity of IA converbs and absolutes ultimately derives from the same morphological source: the Old Indo-Aryan converb. Since I consider early IA examples which are at least potentially ambiguous between converb and CV readings, for Sanskrit and Pāli examples I utilise the term ‘gerund’, a term from the western Sanskritist tradition, simply to indicate the morphological form while remaining neutral on how it is to be interpreted. In addition to CVs and converbs, modern IA languages employ other verb-verb collocations (discussed in §3), where the first member appears as a past or present participle, rather than an absolute; these collocations typically exhibit continuative semantics, rather than the perfective semantics usual of CV constructions.

IA, given the length of its linguistic record, presents an ideal opportunity to evaluate the diachronic properties of light verbs, including the consideration of the possible antiquity of IA CV constructions, and whether light verbs exhibit any diachronically unusual degree of stability (see Butt and Lahiri 2002, Butt 2010, Bowern 2008).

The following section, §2, discusses light verbs and auxiliaries and potential differences between them. I establish rough criteria for distinguishing between light verbs and auxiliaries, though establishing a clear distinction between the two theoretical categories is difficult, particularly in any crosslinguistically robust fashion.

In §3, I investigate the history of verb-verb collocations in IA and show that the development of CV structures of the type found in Hindi, Nepali and other modern IA languages represents a true innovation. This conclusion accords with previous research; for example, in his study cited at the beginning of this section, Masica (1991: 326) remarks that “[compound verbs are] one of the true innovations of [modern IA], unknown to Sanskrit.”

An examination of lexical and morphosyntactic properties of CV constructions in two modern IA languages, to wit, Hindi and Nepali, in §4 further bolsters the position that light verbs are not exempt from diachronic change (contra Butt and Lahiri 2002, Butt 2010), revealing numerous differences in the CV systems of Hindi and Nepali.

In §5 I provide examples of light verbs being re-analysed as auxiliaries, including especially the reanalysis of light verbs as tense/aspect auxiliaries, filling the empirical gap pointed to by Bowern (2008: 174). Additionally, this section discusses how some of the developmental processes responsible for reanalysis of light verbs as auxiliaries do not involve evolution along grammaticalisation clines, but rather take place as part of larger structurally motivated changes, exhibiting changes which involve morphosyntactic chain shifts. Finally, in §6, I provide a summary and conclusions, along with discussion of avenues for future research.

2. Light verbs and auxiliaries

Languages employ a number of verbal or verb-like functional elements, most prominent being ‘auxiliaries’, including auxiliaries like English have and be, as well as ‘dummy’ auxiliaries like English periphrastic do.

Another category of functional elements is what has been termed ‘light verbs’ (Jespersen 1954; Catell 1984; Grimshaw and Mester 1988) constituting verbs which typically have (homophonous) full lexical verb counterparts, but which enter the derivation structurally deficient in some respect. This structural deficiency may include reduced or null semantic values, or the lack of θ-roles (i.e., failing to assign a semantic role to one or more arguments).

Prototypical light verbs occur in noun-verb complex predicates, as in English do dishes, do windows, etc. Since do in do dishes is essentially semantically inert, the question arises of how to distinguish this light verb use of do from the auxiliary verb do. The difference, simply put, is that the two uses of do have different syntactic properties, as shown by the fact that the periphrastic auxiliary do is still required in (7) despite the presence of the light verb do.

(7) a. I don’t do dishes.
b. *I don’t dishes.

Specifically from the form the Sanskrit converb appears in when formed from a prefixed verb, i.e. with the suffix -(t)ya/(t)yā (and not from the unprefixed form of the converb, which takes the suffix -tvā).
Modern Indo-Aryan languages form noun-verb and adjective-verb complexes in a similar way, using light verbs whose full verb meanings correspond to *be* or *do/make*, as shown in the Hindi examples in (8).

(8) a. mujhe yād hai.
   me.DAT memory.MSC.SG.NOM be.PRES.3SG
   “I remember.”

b. maiṁ-ne uskā intāzār kiyā
   I-ERG he/she.GEN waiting.MSC.SG.NOM make.PAST_PTCP.MSC.SG
   “I waited for him.”

Verb-verb complexes such as Indo-Aryan compound verbs have also been analysed as involving light verbs, e.g. Butt & Lahiri’s (2002) analysis of Hindi and Bengali verb-verb constructions. In contrast to the light verbs involved in noun-verb complexes, however, the light verbs of verb-verb complexes do make a semantic contribution to the predicate, as shown above in §1.8

Many historical treatments do not distinguish between light verbs and auxiliaries, as noted by Butt and Lahiri (2002: 4), who argue that such a distinction in fact should be made. Indeed, just as the light verb *do* and the periphrastic auxiliary *do* have different properties, so too are Indo-Aryan light verbs employed in compound verb constructions structurally distinct from more prototypical auxiliaries like tense-marking *honā* "be"; they do not occur in the same syntactic contexts nor do they contribute the same sort of semantics.

However, in practice it is difficult to maintain a binary distinction between light and auxiliary verbs: on the one hand, these functional elements share a number of properties, making it difficult in some cases to distinguish light verbs from auxiliary verbs in a non-arbitrary manner; on the other, even where verbal functional elements show clear formal differences, these differences require a manifold rather than a binary division—e.g. the light verbs of Hindi CV constructions bear formal properties different from other light verb constructions found in IA; see §3 and §6.1. In other words, some light verbs are more auxiliary-like than others—a situation which makes sense if one accepts that light verbs can in fact become auxiliaries.

Such difficulties are particularly obvious in the case of trying to formulate crosslinguistically robust criteria for distinguishing between the two notional categories. As one reviewer points out, in languages like Modern English auxiliaries might reasonably be distinguished with respect to the special morphosyntactic properties that they bear (e.g. with respect to negation, inversion and ellipsis), while in languages like Modern Greek elements meaning “have” or “be” form periphrastic verbal combinations and thus appear comparable to similar elements in English, but unlike in English, these elements in Greek do not exhibit special morphosyntactic properties.

Given such difficulties, here I focus more on the evaluation of distinguishing criteria relevant to IA. Butt (2010: 65) suggests that one notable difference between light verbs and auxiliaries is that the former generally involve some sort of semantic contribution beyond that of purely functional tense/aspect information. Thus, while light verbs often do signal information regarding telicity, causation or temporal boundedness, they usually also involve semantic contributions of other sorts (including forcefulness, surprise, regret, benefaction, volitionality etc.). Additionally, Butt and Lahiri (2002) and Butt (2010) suggest that light verbs can be distinguished from auxiliaries in that the former are always form-identical to some (semantically contentful) full/main verb, and cannot undergo changes (phonological or otherwise) which do not simultaneously affect their main verb counterparts. A difficulty arises here since there do exist sporadic examples of light verbs without full verb counterparts (§4.2.3).

While recognising difficulties in distinguishing consistently (particularly in any crosslinguistically robust fashion) between light verbs and auxiliaries, the best distinguishing criterion, at least for IA languages, is whether the element in question contributes non-functional semantic information or is a more purely functional element required by the grammar in order to express some sort of paradigmatic information (tense, aspect, number etc.). §6.1 provides further discussion of the light verb-auxiliary distinction.

I now turn to an examination of light verbs in various stages of IA.

8Though Butt (2010: 65) points to noun-light verb constructions like *give an answer* as being distinct from the simplex verb *answer*, suggesting the former signals deliberateness of the action, citing Traugott (1999).
3. The origin of Indo-Aryan compound verb constructions

As discussed in §1, the morphological form of the main verb (pole) in modern IA CV constructions derives from the so-called Old Indo-Aryan (OIA) past gerund of prefixed verbs in -(t)ya/(t)yā (Chatterjee 1926; Hendriksen 1944; Tikkanen 1987), an element which has the "virtual value of an indeclinable participle" (Whitney 1879/1889. 1st/2nd edn: §989), which in general functions as a converb, meaning something like "having X-ed" (though it does not always have a past tense value; as discussed in §3.1). It is clear that the origin of modern IA CV structures ultimately lies in a reanalysis of structures involving a gerund combined with another verb (which, for ease of reference I will refer to as V₂, as it usually occurs following the gerund), so that the Sanskrit example in (9) can in a certain sense be seen as the precursor of Hindi (1b), repeated below as (10).

(9) annam sam-khād-ya gatah asti
food.acc together- eat-ger go.ta_ptcp.nom.sg be.pres.3sg
“He ate up the food and left” (Lit., “Having eaten up the food, (he) left.”)

(10) vah khānā khā gay-ā
he/she food.msc.sg eat.abs go.perf_ptcp-sg.masc
“He ate up the food.”

The development from the gerund + V₂ construction of type exemplified by (9) to the CV construction of the type exemplified by (10) took place via the reanalysis of gerund + V₂ as a single predicate, in which the semantic contribution of the gerund was taken to be primary, and the V₂ as a modifying element (i.e. a vector or light verb). This only happened with a subset of verbs occupying the V₂ position, specifically with verbs with broad semantic values, e.g. verbs meaning "go", "give", "take", etc. Further, converb constructions did not disappear from modern IA (see the converb examples from Hindi in §1); rather a sort of morphosyntactic split took place, with some gerund + V₂ constructions being reinterpretable/reinterpreted as CVs.

3.1. Verb-verb collocations in Vedic & Classical Sanskrit

Due to the etymological connection between absolutes and converbs discussed earlier, and the apparent necessity of assigning examples like (11) idiomatic, complex predicate readings (rather than sequential, converb readings), it might appear that CV constructions have long been part of IA.

Butt and Lahiri (2002) suggest that in fact the modern IA system of CVs does not represent an innovation but rather continues a system of light verb constructions found throughout all stages of IA, and argue that examples like (11), taken from Rgvedic Sanskrit (the earliest attested form of IA), may represent CV structures comparable to those found in modern IA. I will argue in this section, however, that Sanskrit actually has no constructions truly equivalent to modern IA CV constructions.

(11) imē te indra té vayām puruṣṭuta yē tvā+ārabhya
cāramasi prabhūvaso move.1pl.pres rich-one.voc
here yours Indra.voc rel_prone we.nom praised-by-many.voc rel_prone you.acc+grasp.ger
a. "We here, O Indra, are yours, O one praised by many, [we] who keep holding on to you, O rich one." (complex predicate reading)
b. "We here, O Indra, are yours, O one praised by many, [we] who having taken hold of you, move (around), O rich one." (l literal converb reading)

9In some languages, such as Nepali and Braj, the -ya survives as -i. In Hindi, the ending has been completely lost, so that Hindi absolutes have the appearance of bare stems. In most cases, modern IA absolutes are form-identical with converbs. Some languages show secondary differentiation of the converb forms, e.g. Nepali converbs end in -i, -i kar and -era (the first actually only orthographically distinct in Nepali, since in spoken Nepali there is no long/short distinction for high vowels, and thus absolutive -i is indistinguishable from converb -i). In Hindi, converbs are usually marked by the addition of -kar or -ke (though this is not obligatory). The Hindi forms in -kar, -ke appear to derive from an earlier pleonastic addition of a converb form of kar ‘do’—cf. the various possible converb forms in closely related Braj: mār-i, mār-i-ka, mār-i-ka-r-i “having struck” (Kellogg 1893. 2nd edn.)—thus Hindi mār kar “having struck” < “mār kari, lit. “having struck, having done” > “having struck”.

(Hindi data from this period is lacking, but Hindi is likely to have resembled closely related Braj—suggesting the appearance of pleonastic -kar/-ke before the loss of -i)

10Example (9) and all subsequent examples from Sanskrit are shown in unsandhiéd form.
This is in fact the interpretation of Butt and Lahiri (2002) and Butt (2010), who suggest that the modern IA system of CVs does not represent an innovation but rather continues a system of light verb constructions found throughout all stages of IA.

Now it is true, as Delbrück (1888: 406–7) observes, that car- “move” does seem to be able to bear an idiomatic sense when used with a gerund in examples like (11). However, gerund + car- is part of a larger set of constructions used in Sanskrit to indicate continuing action, a set which includes collocations involving not only gerunds, but also present participles. (12) provides constructed examples (with cal- “move” as the main verb, and different possible $V_2$s) of a number of these possible collocations, indicating the periods for which they are attested. Actual attested examples of type (12a) are given in (11) and (13), of type (12b) in (14) and of type (12c) in (15) and (16).

(12)  a. calitva carati / eti / tiṣṭhati  
   go.GER move.PRES.3SG / move.PRES.3SG / stand.PRES.3SG  
“keeps on going” (Rgveda onwards)

b. calan āste / tiṣṭhati / carati  
   go.PRES.PTCP sit.PRES.3SG / stand.PRES.3SG / move.PRES.3SG  
“keeps on going” (Vedic onwards)

c. calan bhavati /asti  
   go.PRES.PTCP be.PRES.3SG / be.PRES.3SG  
“keeps on going” (Classical Sanskrit)

(13)  eka eva _ asya _ doso hi gunān ākramya tiṣṭhati  
   one PART his fault.SG.NOM PART virtue.PL.ACC overpower.GER stand.PRES.3SG  
(a). “[He has] one fault which surpasses [~ keeps overpowering] [his] virtues”. (complex predicate reading)
(b). “[He has] one fault which stands having overpowered [his] virtues.” (literal converb reading)  
(Mahābhārata 3.2,22ab; cf. Tikkanen 1987: 176)

(14)  yathā sūcyā vāsah saṃdadhad iyād evam evāi  
   as needle.SG.INSTR clothing.SG.NOM mend.PRES.PTCP.SG.NEU.NOM go.PRES.OPT.3SG thus PART  
   'tābhir yañasya chidraṁ saṃdadhad eti  
   these.PL.INSTR sacrifice.SG.GEN defect.SG.NOM mend.PRES.PTCP.SG.NEU.NOM go.PRES.3SG  
(a). “Just as one would [habitually] mend a garment with a needle, so with these one [habitually] mends [any] defect of the sacrifice.” (complex predicate reading)
(b). “#Just as one would move mending a garment with a needle, so with these one moves mending [any defect of] the sacrifice.” (literal present participle reading)  
(Aitareya Brāhmaṇa 3,18,6; cf. Whitney 1879/1889. 1st/2nd edn: §1075a)

(15)  tasya _ ahaṁ tapaso viryam jānann āsam  
   his.SG.GEN I.SG.NOM asceticism.PL.Gen power.SG.ACC know.PRES.PTCP.SG.MSC.NOM be.PAST.1SG  
(a). “I kept being aware of the power of his asceticism”. (complex predicate reading)
(b). “#I was one knowing the power of his asceticism.” (literal present participle reading)  
(Mahābhārata 1.11.5; cf. Speijer 1886: §377.iv)

(16)  mā mṛtaṁ rudati bhava  
   not dead one.SG.ACC weep.PRES.PTCP.SG.FEM.NOM be.IMP.2SG  
(a). “Do not keep weeping for the dead one”.  
(b). “#Do not be one weeping for the dead one.”  
(Rāmāyaṇa 2.74,2; cf. Speijer 1886: §377.iv)

---

11 Example (11) is shown in unsandhiéd form.
12 Here and elsewhere _ represents the undoing of sandhi in order to make the glossing clearer.
There are several differences between modern IA CVs and the light verb constructions of (12)–(16). Firstly, unlike modern IA CVs, the constructions in (12) are grammatically peripheral, in the sense that they are not part of the central tense/aspect system of Vedic or Classical Sanskrit. Secondly, unlike modern IA CVs, which overwhelmingly impart a perfective sense, the constructions in (12) all contribute a durative/continuative sense.

If, as Butt and Lahiri (2002) and Butt (2010) maintain, light verbs—a category which they take to include the light (vector) verbs of modern IA CV constructions—are crosslinguistically stable, and further, unlike auxiliaries, never have lexical entries separate from those of their full verb counterparts, then the differences between the semantic contributions of modern IA light CV-type verbs and the Sanskrit light verbs of (12) are doubly surprising. Butt & Lahiri’s claim that light verbs do not have separate lexical entries entails that, for instance, the semantics of full verb go and light verb go can be somehow derived from a single (underspecified) lexical entry. Thus the perfective sense of the Hindi light verb jānā is supposed to be derivable from the same lexical entry from which the full verb go also derives. Even granting that both senses could be in fact somehow derived from the same underspecified lexical entry, this claim is problematic given that in Vedic Sanskrit i- “go” contributes a durative/continuative rather than the perfective sense of Hindi jā- “go”, as example (9) above.

Table 1 highlights the semantic differences between seemingly comparable Sanskrit gerund + light verb and Hindi absolutive (< Skt. gerund) + light verb collocations.

<table>
<thead>
<tr>
<th>Sanskrit light verb semantic contribution</th>
<th>Sanskrit stem</th>
<th>Full verb meaning</th>
<th>Hindi light verb semantic contribution</th>
<th>Hindi stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>durative/continuative</td>
<td>i-</td>
<td>“go, move”</td>
<td>perfective</td>
<td>jā-</td>
</tr>
<tr>
<td>durative/continuative</td>
<td>ās-</td>
<td>“sit”</td>
<td>perfective</td>
<td>baith-</td>
</tr>
<tr>
<td>durative/continuative</td>
<td>sthā-</td>
<td>“stand”</td>
<td>perfective</td>
<td>khaṛā ho-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(‘deliberate’ or ‘immediate’)</td>
</tr>
</tbody>
</table>

Table 1. Skt. ger + light verb vs. Hindi abs + light verb collocations

Thus the light verb constructions of Sanskrit differ significantly from the light verb constructions of modern IA, both in terms of their integration into the grammar and in terms of their semantic contribution. What the light verb constructions of Sanskrit do resemble are the present participle+light verb constructions of modern IA, as exemplified by Hindi in (17) and (18), and Nepali in (19) and (20). Note that these constructions are morphologically and semantically distinct from CV constructions.

(17) vah din bhar so-tā gay-ā
He kept on sleeping all day.”

(18) vah din bhar so-tā rah-ā
“He kept on sleeping all day.”

(19) ma bhan-dai jān-chu, tam lekh-tai jā
“I will keep dictating, and you keep writing.” (Pokharel 1991: 194)

See below, §3.2, for discussion of central versus peripheral.

An example of khaṛā honā used as a light verb is given below.

(i) pulis-ko ātā dekh gunde bhāg khare hue
“Having seen the police coming, the villains took to their heels.” (Nespital 1997: 274)

This light verb is itself composite, consisting of an adjectival form and a (non-CV-type) light verb.
With certain verbs whose perfective participles effectively express present states, we find these sorts of constructions formed with perfective participles, e.g. (McGregor (1995. 3rd edn.: 150)):

(21) larēkī ek ghantā vahāṁ baiṭī rahī
girl one hour there sit.PERF_PTCP.FEM.SG remain.PERF_PTCP.FEM.SG
“The girl stayed sitting there for an hour.”

We also find collocations of perfect participles (with invariant default masculine singular agreement morphology) with *karnā* “to do”, which serve to indicate the habitual (rather than the continuative-durative in (17), (18), (19), (20)) nature of the action, e.g. (from McGregor (1995. 3rd edn.: 151)):

(22) pahle mere yahāṁ āyā kartī thī, ab nahīṁ
before my here come.PERF_PTCP.MSC.SG do.IMPF_PTCP.FEM.SG be.PAST.FEM.SG, now not
ātī
come.IMPF_PTCP.FEM.SG
“She used to frequently come to my place, but now she doesn’t.”

Not only are modern IA constructions in (17)–(20), like the Sanskrit constructions in (13)–(16) above, grammatically peripheral (see below, §3.2), in that they are not part of the core morphologically-realised tense/aspect system of Hindi or Nepali, but these modern IA light verbs also make the same sort of semantic contribution as do those of Sanskrit, namely durative/continuative. The peripherality of such constructions is such that a speaker could still sound idiomatic and nativelike without using them. The same would not be true of a speaker who failed to use CV constructions, on which see further Hook (1974), who notes the fact that CVs are nearly obligatory in realis contexts (at least in Hindi; this is less true for certain other IA languages, such as Nepali). Further, as noted above in (12), the Sanskrit light verb constructions include not only gerund + light verb, but also—as in Hindi and Nepali—present participle + light verb, both of which contribute durative/continuative senses. Therefore, examples claimed by Butt & Lahiri to be early CV constructions in Sanskrit are actually much more comparable to the grammatically-peripheral present participle + light verb constructions of (17)–(20).

Both in old and modern IA, a variety of different V-V collocations involve the light verb of the construction selecting for morphologically different types of V₁s—where V₁ is the verb contributing the core semantic sense of the collocation, itself appearing in a non-finite form. In old IA, both the collocations involving gerund V₁s and those involving present participle V₁s are grammatically-peripheral and involve continuative semantics. Thus both are actually more comparable to the modern IA collocations involving present participle V₁s than to the modern IA CV construction.

Butt and Lahiri (2002: 23) cite an example of a Skt. gerund + light verb construction, (23), where they take the light verb as contributing a perfective sense comparable to that of light verbs of modern IA CV constructions.

(23) tataḥ māṣikā+ uḍḍiya gatā
then fly fly-up.GER go.TA_PTCP.FEM.SG
a. “Then the fly flew off.” (complex predicate reading)
b. “Then the fly, having flown up, left.” (literal converb reading)
(Pañcatantra 1.22, Tikkanen 1987: 176)16

They claim that here “the verb ‘go’ cannot be seen as the main predication of the sentence...[r]ather it modifies the event semantics of the participle ‘fly’”, comparing it to modern IA examples like Hindi (24).

(24) makkhī uṛ gayi
fly.FEM.SG.NOM fly.ABS go.PERF_PTCP.FEM.SG
“The fly flew off”

16Example (23) is shown in unsandhied form.
Although (23) looks superficially like (24), it is perfectly possible to interpret the gerund literally, as indicated by the translation in (23b). In fact, the overall absence of examples of this sort from Sanskrit suggests that positing a perfective light verb go in Sanskrit would require special pleading.17

More generally, the close examination of the putative examples of early IA CV constructions has revealed nothing truly comparable to the CV constructions of modern IA. Some of these examples were shown to be perfectly acceptable with their literal converb readings, while the non-spurious examples of light verb constructions in Sanskrit were shown to most closely resemble the grammatically-peripheral light verb constructions in modern IA involving imperfect or perfective participles. In sum, there is no evidence which would suggest that we attribute CVs to Old IA.

Table 2 provides a summary of the various types of constructions discussed in this section and throughout the remainder of this paper (note, however, that for the most part these constructions are all formally distinct, in that they involve different morphological forms—although there is some amount of overlap/ambiguity between converb and compound verb constructions in terms of their surface forms as noted above).

<table>
<thead>
<tr>
<th>Semantics</th>
<th>Centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONVERB</td>
<td>sequence of two events</td>
</tr>
<tr>
<td>PRESENT PARTICIPLE + LIGHT VERB</td>
<td>light verb contributes a continuative/durative sense to the interpretation of the (present participle) verb</td>
</tr>
<tr>
<td>PART PARTICIPLE + LIGHT VERB</td>
<td>light verb contributes a habitual sense to the interpretation of the (present participle) verb</td>
</tr>
<tr>
<td>ABS + LIGHT VERB (&quot;compound verb&quot;)</td>
<td>light verb contributes (generally) perfective Aktionsart &amp; other evaluative/subjective nuances to the interpretation of the (ABS) verb</td>
</tr>
</tbody>
</table>

Table 2. Overview of different types of modern IA verb-verb constructions discussed herein

3.2. **Excursus: periphery vs. core**

In §3.1 above, I distinguished between peripheral light verb constructions, like the continuative constructions of Sanskrit (12) and Hindi (17), (18) (and Nepali (19), (20)), and core grammar constructions, like Hindi CV constructions. In this section I consider examples of core vs. peripheral constructions in English in order to make this distinction more concrete.

In general, core grammar constructions are less specific in their functions than more peripheral constructions, and are involved in the expression of more basic functions. Thus in (25) below, will X is the basic construction for expressing future tense in English, whereas the constructions to the left of the dotted line on the periphery-core continuum can be considered to be more peripheral, non-basic constructions associated with future time.

(25) English future-oriented constructions: (X = bare infinitive verb)

```
<table>
<thead>
<tr>
<th></th>
<th>BE</th>
<th>BE</th>
<th>BE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periphery</td>
<td>under oath to X</td>
<td>sworn to X</td>
<td>bound to X</td>
</tr>
<tr>
<td>Core</td>
<td>will X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

17The only other example involving an apparent light verb meaning "go" I have seen is also from Tikkanen (1987: 176):

(i) tato yāvat sā tāṁ sārikāṁ galamotanapūrvaṁ vināśayati tāvad
then when she.sg.nom that.sg.acc Sarika-bird.sg.acc before wringing the neck cause the destruction of pres.3sg then
uddiya yaya
fly-up.ger go.perf_ptcp.3sg
a. "Then, before she had time to strangle that Sarika-bird to death, it flew away." (complex predicate reading)
b. "Then, before she had time to strangle that Sarika-bird to death, it flew up and left." (literal converb reading)

for which the same concerns may be raised as for (23).
The core-periphery distinction is at least somewhat gradient in nature, as shown by the fact that the construction BE bound to X can have a much less specific sense than either BE sworn to X or BE under oath to X, namely “have a logical necessity to X”. In origin, BE bound to X is nearly equivalent to the other two constructions, namely bearing the sense “having entered into a contract binding to service; under legal or moral obligation” (OED); showing that BE bound to X has undergone grammaticalisation—displaying a bleaching of its earlier sense, with the construction now bearing a more general future-oriented sense.

Turning to an English example which is semantically closer to the sense of the peripheral constructions considered above, namely Sanskrit (12) and Hindi (17), (18), consider the English continuative constructions in (26).

(26) English continuative constructions: \( (X = \text{bare infinitive verb}) \)

<table>
<thead>
<tr>
<th>Periphery</th>
<th>Core</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE occupied in X-ing</td>
<td>BE X-ing</td>
</tr>
<tr>
<td>BE engaged in X-ing</td>
<td>keep on X-ing</td>
</tr>
<tr>
<td>BE in the process of X-ing</td>
<td>keep X-ing</td>
</tr>
</tbody>
</table>

Here the construction BE X-ing is the most basic expression of continuative action in English, a core grammar construction. The constructions to the left of the dotted line are more peripheral constructions with less basic continuative senses: BE X-ing is the basic grammatical form for continuatives in English. Again, as in the case of BE bound to X, some of these peripheral constructions, namely keep X-ing and keep on X-ing, have undergone grammaticalisation and become more general—but still not as general as the core construction BE X-ing, as shown by the contrast between (27a) and (27b) below.¹⁸

(27) a. John was working all day.
   b. John kept on working all day.

The core-periphery distinction between BE X-ing and keep (on) X-ing is neatly paralleled by the distinction in Hindi between X-tā rahnā “keep on X-ing” and X rahā HONĀ “is X-ing”, see (28).

(28) Hindi continuative constructions: \( (X = \text{bare verb stem}) \)

<table>
<thead>
<tr>
<th>Periphery</th>
<th>Core</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-tā rahnā “keep on X-ing”</td>
<td>X rahā HONĀ “is X-ing”</td>
</tr>
<tr>
<td>CVs</td>
<td>X-tā HONĀ “X-es”</td>
</tr>
</tbody>
</table>

CV constructions—in contrast to X-tā rahnā—can be considered part of the core grammar of Hindi in view of the fact that they are nearly obligatory even in very basic expressions, in the sense that even the simple statement he died is much more idiomatically expressed with a CV (29a) than with a simplex verb (29b); (29a) in most contexts will not sound native-like, though it is not ungrammatical.

(29) a. vah mar gayā.
   he/she die.ABS go.PERF_PTCP.MSC.SG
   “S/he died.”

b. vah marā.
   he/she die.PERF_PTCP.MSC.SG
   “S/he died.”

3.3. Verb-verb collocations in Middle Indo-Aryan

Some possible examples of early CV constructions are found in Pāli texts, largely from Sri Lanka (Tikkanen 1987; Hook 1993), with the most convincing examples involving gerund + give, as in (30), (31).

(30) ath’assa sathā udakasātkāmi khipitvā adāsi
   then him master water-cloak throw.GER give.PAST.3SG

¹⁸Constructions similar to, though not completely identical with, keep on X-ing include get on X-ing, carry on X-ing, press on X-ing. See Cappelle (1999) for discussion, including the semantic distinctions and historical differences between keep X-ing and keep on X-ing.
a. “Then the master threw a bathrobe to him.” (complex predicate reading)
b. “Then the master gave him a bathrobe, throwing (it).” (literal converb reading)

(Dhammapad-Aṭṭhakathā II,61,10; post 5th-c. C.E., cited Hendriksen 1944: 134)

(31) so tassā saddāṁ sutvā . . . assamapadāṁ ānetvā aggim katvā adāsi
he her cry.ACC hear.GER . . . hermitage.ACC bring.GER fire.ACC make.GER give.PAST.3SG
a. “He, having heard her cry, having brought her to his hermitage, made a fire (for her).” (complex predicate reading)
b. “He, having heard her cry, having brought her to his hermitage, having made a fire, offered it (to her).” (literal converb reading)

(Jātaka I,296,10; between 300 B.C.E. and 400 C.E.; cited Hendriksen 1944: 134)

(30) and (31) look much more promising than the other supposed early examples of CV constructions; however even here (as indicated by the two possible translations) it is not necessary that we interpret such examples as involving complex predicates.

In example (31) a literal converb reading is possible, as indicated by the translation, and thus again there is no evidence which forces us to interpret give here as a light verb—in fact, given that (31) involves a string of gerunds (“having heard her cry, having brought her to his hermitage, . . .”), a complex predicate interpretation seems even more unlikely.¹⁹

Example (30), at first blush, looks much less likely to be interpretable as involving a literal converb reading. However, gerunds sometimes bear ‘non-past’ interpretations; see the Sanskrit examples (32) and (33).²⁰

(32) sitā mad-vacanāt vācyā samāśvāsya prasādaya ca
Sita.FEM.SG.NOM my-command.ABL speak.GER console.GER and
a. “By my command Sita is to be spoken to, having consoled and calmed (her).” (past-tense converb reading)
b. “By my command Sita is to be spoken to consolingly and calmingly.” (non-past converb reading)

(Mahābhārata 3.264,56, cited in Tikkanen 1987: 123)

(33) vikramārko nītim uḷaṅghya rāyjaṁ na karoti
Vikramārka.NOM rules-of-ethics.ACC transgress.GER rule.ACC not do.PRES.3SG
a. “Vikramarka, having transgressed the rules of ethics, does not rule (his kingdom).” (past-tense converb reading)
b. “Vikramarka does not rule (his kingdom), transgressing the rules of ethics.” (non-past converb reading)

(Vikramacarita 18.0, cited in Tikkanen 1987: 124)

As indicated by the translations, in these examples the past-tense converb reading is unavailable. In Pāli as well we find examples where gerunds must be interpreted with a non-past reading, as in example (34) below.

(34) . . . atha so . . . bhatiṁ katvā jīvati
. . . thus he.NOM . . . wages.ACC do.GER live.PRES.3SG
a. “. . . and so he, having worked as a day-labourer, lives.” (past-tense converb reading)
b. “. . . and so he lives by working as a day-labourer.” (non-past converb reading)

(Jātaka 41, cited Tikkanen 1987: 125)

Thus a literal non-past converb reading of (30) is perfectly plausible, i.e. “then the master gave him a bathrobe, (by) throwing (it)”. The Pāli examples in (30) and (31) thus need not be interpreted as involving complex predication. They do look like excellent starting points for the reanalysis of gerund constructions as complex predicates, but there is no reason to assume that such a reanalysis has already occurred by the stage of Pāli since in all cases the literal converb readings are perfectly possible.²¹

¹⁹See §1 for the definition of “gerund” which is relevant here.
²⁰On ‘non-past’ readings of Sanskrit gerunds, see Hock (1992).
²¹Even if one were to assume that examples (30) and (31) represent nascent CV constructions, since these texts are from Sri Lanka, it is not clear what bearing they have on the development of compound verb constructions in mainland IA, especially given that they may reflect the influence
3.4. Verb-verb collocations in late Middle Indo-Aryan and early Modern Indo-Aryan

It has been suggested that CV constructions are found in Apabhraṁśa (Hook 1977; Bubenik 1998). However, most if not all of Bubenik’s examples seem to be more felicitously interpreted in other ways, as suggested to me by Eva De Clercq (p.c.). Singh (1980: 164–167) lists, without context, a number of examples of potential CV constructions in Apabhraṁśa. Given the uncertainties of interpretation of the Apabhraṁśa data, then I leave them out of consideration. This is an area requiring further investigation.

The first unequivocal examples of CV constructions do not appear until the early modern IA period (16th–18th c.).

See the examples from early Braj Bhāṣā (a close relative of the ancestor of modern Hindi) in (35), (36), and the Nepali examples in (37), (38).

(35) nātaru jau vacchā marī jāi, tau gāi cchīṃḍāī jāi
   otherwise if calf.nom die.abs go.pres.3sg, then cow.nom take.away.perf_ptcp.fem.sg go.pres.3sg
   a. "Otherwise, if the calf dies, then the cow is taken away." (complex predicate reading)
   b. "Otherwise, if the calf having died goes, then the cow is taken away." (literal converb reading)
      (ca. 1600 c.e., Indrajit of Orchā’s commentary on the Nītiśataka of Bhartr̥hari, f.18b4; McGregor 1968: 57)

(36) bhayabhīta hvai-kari samudra kau mathivau na cchāḍ-i dayau
   without fear be-conv ocean of churning not leave-abs give.perf_ptcp.msc.sg
   a. "Having become fearless, (he) did not leave off the churning of the ocean." (complex predicate reading)
   b. "Having become fearless, (he), having not left the churning of the ocean, gave." (literal converb reading)
      (ca. 1600 c.e., Indrajit of Orchā’s commentary on the Nītiśataka of Bhartr̥hari, f.17a; McGregor 1968: 54)

(37) tahāṁ kāji-kana vujhāi samjhāi ṇahā pathā-i- dinu havas
   there Kazi.instr understanding understand.caus.conv here send-abs-give.imp
   a. "Persuade the Kazi there, and send him here." (complex predicate reading)
   b. "Persuade the Kazi there, and having sent him here, give (him).” (literal converb reading)
      (1755/6 c.e. ~ 1812 V.S., Prithvinarayan Shah, letter to Paṇḍit Rājivalocana; Pokharel 1963: 196)

(38) tyo cāṁḍo māgera pathā-i- deu
   that(thing) quickly request.conv send-abs-give.imp
   a. "Quickly request that thing and send it (to me).” (complex predicate reading)
   b. "Quickly request that thing and having sent it, give (it) (to me).” (literal converb reading)
      (1767/7 c.e. ~ 1824 V.S., Prithvinarayan Shah, letter to Haripaṇḍit, Pokharel 1963: 211)

As indicated by the translations, in these examples a literal converb reading is unavailable, demonstrating that by the early modern period of IA, CVs have made their appearance.

In conclusion, sound evidence for CV constructions in IA thus does not appear until the early modern period (pace Butt and Lahiri 2002). Verb-verb collocations apparently involving light verbs do appear early on in IA, but these are grammatically-peripheral constructions, distinct in their semantics from CV constructions, resembling instead the (likewise peripheral) present participle + light verb constructions of modern IA.

4. Morphosyntactic and lexical change in IA CVs

If light verbs were “diachronically stable” (Butt and Lahiri 2002, Butt 2010) we should expect the morphosyntactic and lexical properties of the modern IA CV systems to be very similar. This however does not obtain. Hook and Pardeshi (2005) point out a number of light verb constructions in IA which show evidence of historical change, and herein I provide a number of additional instances of historical changes involving light verbs in IA. I examine particularly two closely related IA languages, Hindi and Nepali, and show that, both in terms of morphosyntax and the lexicon,
the CV systems of these two languages differ significantly. These indicate that change has taken place in the CV systems of one or both languages (assuming CVs to have originated in IA no earlier than in Apabhramśa).

4.1. Morphosyntactic change in IA CVs

I examine four differences between the morphosyntax of Hindi and Nepali CV constructions. These properties I refer to as ‘interruptibility’, the ability of elements to intervene between the main verb and the light verb of the CV construction; ‘recursion’, whether a main verb may be modified by more than one light verb; ‘construction-specific restrictions’, i.e. which syntactic constructions light (vector-type) verbs can occur in; and ‘dominance’, whether agentive marking of subject of transitive verbs in perfective tenses is controlled by the transitivity of the main verb or the transitivity of the light verb.

4.1.1. Interruptibility

Hindi CVs are interruptible; that is, other words may occur in between the main verb and the light verb components of the CV, as shown in (39).

(39) mai-ne khānā khā to liy-ā, lekin phir ulī bī ā-ī
I-AGT food eat.ABS EMP take.PTCP-SG.MASC but then vomit also come.PTCP-SG.FEM
"I did eat food, but then I also vomited."

Nepali CVs are not interruptible, even by particles like ta (equivalent to Hindi to):

(40) *mai-le khānā khā-ī ta sak-em.
I-AGT food eat-ABS EMP finish-PAST.SG.MASC
"I did eat food, . . ."

This difference is possibly reflected in writing, as Hindi CVs are written as separate words, but Nepali CVs usually as a single word. The inseparability of Nepali CVs suggests that they are either formed in the lexicon, or else composed at a much lower level of syntax than Hindi CVs. In other words, Nepali light verbs appear to be more affixal in nature than are Hindi light verbs.

4.1.2. Recursion

Hindi allows for only one light verb per main verb, as shown in (41).

23Nepali and Hindi are fairly closely related languages: as Matthews (1998: i) remarks, "[Nepali and Hindi] are in fact so close to each other that early Western grammarians regarded Nepali merely as a dialect of Hindi". The relationship between them is roughly paralleled by that between Italian and Spanish.

24These properties represent a subset of the properties derived in a bottom-up fashion based on consideration of the major differences in Hindi and Nepali CV constructions as found in these data. These data were drawn from previous accounts of compound verbs in Hindi and Nepali (see fn. 1), from grammars (Hindi: Kellogg 1893. 2nd edn., McGregor 1995. 3rd edn.; Nepali: Verma and Sharma 1979, Matthews 1998), and from consultation with native speakers. Early Nepali examples are taken from texts contained in Pokharel 1963.

25All Hindi speakers consulted allowed for particles like to to intervene between main verb and light verb, as in (39); most speakers also allow pronouns to intervene, as in (ia); some also allow for full NPs, as in (ib), (ic).

(i) a. ā to vah gay-ā hai, lekin sabzi nāhin lāy-ā
come.ABS EMP he go.PTCP-SG.MASC be.3P.SG.PRES but vegetables not bring.PTCP-SG.MASC
"He has indeed come, but he didn’t bring the vegetables."

b. %kha to khānā liy-ā
eat.ABS EMP food take.PTCP-SG.MASC
"(I/he/she) did eat food."

c. %liy-ā to khānā khā
take.PTCP-SG.MASC EMP food eat.ABS
"(I/he/she) did (in fact) have food."

26This is unlikely to be an artifact of the script, as both languages employ the same writing system (devanāgarī). However, a reviewer points out that that Nepali CVs are not invariably written as if single words, though this is the majority pattern.
a. *us-ne kican sāf kar di
   he/she-AGT kitchen clean do.ABS give.PAST.PTCP.SG.FEM
   “He/she cleaned the kitchen for me.”

b. *us-ne kican sāf kar dāl-i
   he/she-AGT kitchen clean do.ABS put.PAST.PTCP-SG.FEM
   “He/she cleaned the kitchen straightaway.”

c. *us-ne kican sāf kar de dāl-i
   he/she-AGT kitchen clean do.ABS give.ABS put.PAST.PTCP-SG.FEM
   “He/she cleaned the kitchen for me straightaway.”

In Nepali, on the other hand, CVs may involve up to two light verbs, as in (42), where the main verb, *gari*, is modified by both *dii* (itself in absolutive form) and *hālin*.

(42) un-le kican sapāhā gar-i- di-i- hāl-in
   he/she.MidHON-ERG kitchen clean do.ABS- give.ABS- put-PAST.3SG.FEM
   “She cleaned the kitchen for me straightaway.” (Peterson (2002: 107))

In certain respects this makes Nepali CV appear more affixal in nature, given that main verb + light verb behaves like a simplex verb for purposes of additional operations (including the addition of other light verbs).

While it is not entirely clear which of these patterns represents the more innovative and which the more conservative, it is obvious that the CV system of at least one of the two languages has undergone change.\(^\text{27}\)

### 4.1.3. Construction-specific restrictions

Nepali CVs can occur in conjunctive participles. Thus, for example, a converb in Nepali may be composed from a CV, as in (43).

(43) bhāt khā-i- sak-era u sut-na gay-o
    food eat-ABS- finish-CONV he/she sleep-INF.OBLQ go.PAST-3SG
    “Having finished dinner, he went to sleep.” (Peterson (2002: 108))

In contrast, Hindi converbs can only be formed from simplex verbs. Therefore while (44b) is grammatical, (44a) is not.\(^\text{28}\)

(44) a. ’khānā khā le kar vah gay-ā
    food eat.ABS take CONV he/she go.PAST_PTCP-SG.MASC
    “Having eaten up the food, he left.”

b. khānā khā kar vah gay-ā
    food eat.ABS CONV he/she go.PAST_PTCP-SG.MASC
    “Having eaten the food, he left.”

Hindi appears to have innovated: what is now the overt converb marker in Hindi, *kar* or *ke*, derives ultimately from the pleonastic addition of a converb form of *kar* “do” (see fn. 9), pointing to the possibility of forming converbs of CV collocations in earlier Hindi, and suggesting that with respect to this property Nepali is conservative.

### 4.1.4. Dominance and transitivity issues

In both Hindi and Nepali, the question arises of how to reconcile clashes in transitivity between main verb and light verb in CV combinations. Before going into detail, it is important to understand the basic patterns of case-marking

\(^{27}\)The common ancestor of Hindi and Nepali would be “Proto-Modern-Indo-Aryan”, not directly attested (though presumably similar in certain aspect to the literary Apabhramśa) and therefore not available for examination in determining which language is innovative here. Given that the predecessor construction to IA CVs, early IA gerund + verb, could involve multiple gerunds (“V₁’s”), as in example (31) above, it is possible that the limitation to a single “V₁” is the innovation.

\(^{28}\)Examples like (44a) are possible in Dakhkhini Hindi (spoken in Hyderabad), which has undergone convergence with Dravidian Telugu; see Arora (2004).
morphology in these two languages. Both Hindi and Nepali (more or less)\textsuperscript{29} employ a special marking for agents of transitive verbs in perfective tenses, -ne in Hindi, -le in Nepali. In Hindi, the marking of agents correlates with verbal agreement: agentive case-marked nouns do not control predicate agreement.\textsuperscript{30} In Nepali, (non-experiencer) subjects control predicate agreement regardless of whether they are agentive case-marked or not.

More relevant for the purposes of CVs is the fact that in Hindi it is the light verb member of the CV which determines whether the entire CV is treated as transitive or intransitive (and thus whether the subject receives agentive casemarking)—except in the case of intransitive main verb combined with transitive light verb (where we find variation in the assignment of transitivity to the compound for the purposes of case-assignment);\textsuperscript{31} in Nepali, it is always the main verb which determines the transitivity of the entire verbal compound. This is shown in the examples in (45) and (46).

(45) **Hindi CVs**

a. **Intransitive main verb + intransitive light verb = Intransitive:**

\[
\text{vah ā gay-ā} \\
\text{he come.ABS go.PAST.PTCP-MSC.SG}
\]

“He came.”

b. **Transitive main verb + intransitive light verb = Intransitive:**

\[
\text{vah khān ā khā gay-ā} \\
\text{he food eat.ABS go.PAST.PTCP-MSC.SG}
\]

“He ate up the food.”

c. **Transitive main verb + transitive light verb = Transitive:**

\[
\text{us-ne khān ā liy-ā} \\
\text{he-AGT food eat.ABS take.PAST.PTCP-MSC.SG}
\]

“He ate up the food.”

\textsuperscript{29}Nepali also sometimes displays agentive case-marking in non-perfective contexts; the exact conditions on the use of -le in such cases is not entirely clear: it seems more frequent with third-person subjects. Li (2007) suggests that it is obligatory with inanimates, while Poudel (2006) suggests that -le may distinguish individual-level predicates (a property that holds for a slice of the referent’s spatio-temporal existence) from stage-level predicates (denoting a lasting/inherent property of the referent), whose subjects are Ø-marked. Some dialects of Nepali have generalised agent marking to all transitive verbs regardless of tense/aspect, predicate-type, animacy—this probably reflects convergence with Tibeto-Burman.

\textsuperscript{30}The unmarked object usually controls verb agreement; if this is blocked by the oblique postposition ko, the verb takes default masculine, singular, third-person agreement.

\textsuperscript{31}Combinations of intransitive main verb and transitive light verb are very rare in Hindi, apparently largely avoided by Hindi speakers. I have found only two examples from Nespital (1997: 1108-1109) where the case-assignment is clear (i.e. where the finite verb appears as a perfect participle).

(i) \[
\text{śatru senā acānak hī hamārī senā par ā li} \\
\text{enemy army suddenly EMPH our army on come.ABS take.PERF.PTCP.FEM.SG}
\]

“The enemy army suddenly pounced on our army.”

(ii) \[
\text{jab satīś apne ghar se nīkl-ā to uskā kuttā bhi uske piche ho liy-ā} \\
\text{when Satish his own house from exit-PERF.PTCP.MSC.SG then his dog also his behind be.ABS take-PERF.PTCP.MSC.SG}
\]

“When Satish left his house, then his dog also followed him.”

In both cases, the subject is not marked with an agentive marker, despite the light verb member being transitive. Remarkably, in both (i) and (ii), not only is the light verb transitive, but the CV itself appears to be semantically transitive (in the sense that both pounce and follow require complements). However, more generally combinations of intransitive main verb and transitive light verb seems to result in variation in Hindi (probably as a result of the rarity of such combinations). That is, some speakers prefer (iii), others prefer (iv) (yet others prefer to avoid such combinations altogether).

(iii) \[
\text{vah baith liy-ā} \\
\text{he sit.ABS take.PAST.PTCP-MSC.SG}
\]

“He sat down.”

(iv) \[
\text{us-ne baith liy-ā} \\
\text{he-AGT sit.ABS take.PAST.PTCP-MSC.SG}
\]

“He sat down.”

Speakers show similar variation with respect to other combinations, such as dikh diyā “appeared”, and ghum liyā “roamed”.

15
d. Intransitive main verb + transitive light verb = (variation) (see fn. 31)

\[ \text{vah/us-ne} \ \text{baith} \ \text{liy-ā} \]
he.NOM/he-.AGT sit.ABS take.PERF_PTCP-MSC.SG
“He sat down.”

(46) NEPALI CVs

a. Intransitive main verb + intransitive light verb = Intransitive:

\[ \text{u} \ \text{mandir-mā} \ \text{ga-i-} \ \text{ā-eko} \ \text{cha} \]
he.temple-LOC go-ABS- come-PERF_PTCP-MSC.SG be.PRES.3SG
“He kept going to the temple.”

b. Transitive main verb + intransitive light verb = Transitive:

\[ \text{us-le} \ \text{yo} \ \text{kām} \ \text{gar-i-} \ \text{ā-eko} \ \text{cha} \]
he-.AGT this work do-ABS- come-PERF_PTCP-MSC.SG be.PRES.3SG
“He has continued to do this work.”

c. Transitive main verb + transitive light verb = Transitive:

\[ \text{us-le} \ \text{bhāt} \ \text{kā-i-} \ \text{di-yo} \]
he-.AGT food eat-ABS- give-PAST.3MSC.SG
“He ate up the food.”

d. Intransitive main verb + transitive light verb = Intransitive:

\[ \text{u} \ \text{rām} \ \text{ko-lāgi} \ \text{bajār} \ \text{ga-i-} \ \text{diy-o} \]
he.Ram for-sake-of market go-ABS- give-PAST.3MSC.SG
“He went to the market for Ram’s sake/in place of Ram.”

Thus, in Hindi the light verb usually determines the transitivity of the entire compound, for purposes of assignment (or non-assignment) of agentive marking to the subject. In Nepali, it is the main verb which determines the transitivity. Again, though it is unclear which of these represents the more conservative system, it is clear that change has taken place in the CV system of one or both languages. The uncertainty stems from the absence of extant Proto-Modern-Indo-Aryan, and the fact that for the predecessor gerund construction in earlier IA such issues did not arise: we do not find an ergative/absolute-type system in Sanskrit or Pāli.

4.2. Lexical change in IA CVs

While there is much commonality in the sets of light verbs employed by the different modern IA languages in CV constructions, they are not identical. Thus, in some cases the same light verb sense is expressed using verbs with the same full verb meaning, e.g. Hindi \text{ḍāl-} and Nepali \text{hāl}, Hindi \text{de-} and Nepali \text{di-} (cognates, in the latter case); some light verb senses are expressed using verbs with distinct full verb senses, e.g. Hindi \text{baith-} and Nepali \text{paṭhāu-}; in some cases a light verb sense is expressable only in one of the languages, e.g. Hindi \text{le-}. See Table 3.

<table>
<thead>
<tr>
<th>HINDI STEM</th>
<th>HINDI FULL VERB MEANING</th>
<th>LIGHT VERB SENSE</th>
<th>NEPALI FULL VERB MEANING</th>
<th>NEPALI STEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{ḍāl}</td>
<td>throw</td>
<td>immediacy</td>
<td>throw</td>
<td>\text{hāl}</td>
</tr>
<tr>
<td>\text{de}</td>
<td>give</td>
<td>other-directed</td>
<td>give</td>
<td>\text{di}</td>
</tr>
<tr>
<td>\text{le}</td>
<td>take</td>
<td>self-directed</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>\text{cuk}</td>
<td>finish</td>
<td>completive</td>
<td>complete / be able to</td>
<td>\text{sak}</td>
</tr>
<tr>
<td>\text{baith}</td>
<td>sit</td>
<td>regret</td>
<td>send</td>
<td>\text{paṭhāu}</td>
</tr>
</tbody>
</table>

Table 3. Selected Hindi and Nepali light verbs compared

16
4.2.1. Against shared full verb/light verb entries 1: light verbs of regret

In their analysis of IA light verbs, Butt and Lahiri (2002: 43–46) suggest that each full verb/light verb pairing in IA CV constructions can be derived from a single underspecified entry, e.g. that both the full and light verb uses of Hindi *denā* “give” derive from a single lexical entry, as in (47).

(47) “give” — Underspecified Entry

Butt and Lahiri (2002) adopt a semantics involving subevents,\(^{32}\) where the light verb use of “give” can be “taken to modify the superordinate event by contributing lexical semantic information which is loosely based on the predicational force of the main verb ‘give’” (45). More specifically they posit that this information consists of an adverbial event modification where “give” contributes to the main verb semantics the fact that the event (a) involves the application/mission of force and (b) is agentive/purposeful. Generally-speaking, it is far from clear what the nature of the underspecified entry of (47) would be, such that both the full verb and light verb uses of *give* could be systematically derived from it. However, even if we grant that this theory could be maintained for light verbs like *denā*, it still requires that the system of derivation of full and light verbs from a single underspecified lexical entry be cross-linguistically robust, in the sense that the connections between the full verb and light verb senses should be predictable. This might seem reasonable in the case of light verbs like *give*—which do at least tend crosslinguistically to involve similar light verb semantic contributions\(^{33}\)—but even within IA, languages vary greatly in the association between full and light verb meanings.

For example, both Hindi and Nepali have light verbs used when the speaker wishes to express regret; however, Hindi and Nepali use light verbs whose full verb counterparts bear completely different senses: the Hindi light verb indicating regret is *baithnā* “to sit” (48), the Nepali light verb indicating regret is *paṭhāunu* “to send” (49).

(48) *main kyā kar bais-hā?*
    I.NOM what do.ABS sit.PERF PTCP-MSC.SG
    “Oh what have I done?”

(49) *maile ke gar-i paṭhā-eṁ?*
    I.AGT what do.ABS send.PERF PTCP-1SG
    “Oh what have I done?”

Not only is it unclear what sort of single underspecified lexical entry could underlie senses of Hindi *baithnā* or Nepali *paṭhāunu*, but the very fact that these languages, for the same light verb semantic contribution, employ verbs with completely different full verb senses strongly suggests that, even where there is homophony between full and light verbs, light verbs have their own lexical entries, separate from those of their full verb counterparts (where counterparts even exist). More relevantly for current purposes, the lexical differences in the light verb inventories of Hindi and Nepali are another indication of historical change affecting light verbs independently of their full verb counterparts.

4.2.2. Against shared full verb/light verb entries 2: independent phonological change

The Nepali light verb *baksinu* provides another example of change affecting a light verb independently of its full verb counterpart. Nepali *baksinu* is a light verb employed in CV constructions as an honorific when referring to Nepali royalty and other persons due great respect, e.g.:


\(^{33}\)See e.g. Masica (1976: 141-158), who notes, beside the ubiquity of *give* in Indo-Aryan, similar functions for light verb uses of *give* in Burmese, Tajik and Uzbek. However, even in other languages employing light verb type constructions, for instance Japanese and Turkish, *give* appears with rather different functions (as an honorific in Japanese, as a light verb indicating “rapidity, ease, suddenness” in Turkish). Even in Indo-Aryan *give* does not always contribute a sense of “other-directedness”; in Nepali, alongside this function, it also can simply mark general perfectivity of the action.
Both baksinu and a phonologically-reduced form are also employed in upper-class Kathmandu families (e.g. by children to parents, wife to husband etc.), as discussed in Sharma (1980: 130–2). An example of the reduced form of baksinu is given in example (51).

(51) buwā-le bhujā khā-i- s-yo
dad-AGT rice eat-ABS- HON-3SG
“Dad ate rice.” (Sharma 1980: 132)

While baksinu is employed as a main verb (52), -s- is not (53).

(52) mahārāni-le ma-lāi takmā baksi-yo
queen-AGT 1-OBLQ medal bestow-PAST.3SG
“The queen bestowed a medal upon me.” (Ibid.)

(53) *mahārāni-le ma-lāi takmā s-yo
queen-AGT 1-OBLQ medal hon-PAST.3SG

That baksinu as a light verb, but not as a full verb, can appear in the reduced form -s- is problematic for Butt & Lahiri’s (2002) claim that light verbs cannot undergo any change not also experienced by their full verb counterparts, since baksinu as a light verb can occur in both a reduced and an unreduced phonological form, but the full verb form only occurs in unreduced form.  

4.2.3. Light verbs lacking full verb counterparts

Though there seems to be some tendency for light verbs to continue to have full verb counterparts, this is not always the case. Nepali, for example, employs ṭopalnu as a light verb indicating pretense, as in example (54) below, but ṭopalnu does not exist as a full verb.

(54) u gā-i- ṭopal-dai- cha
he/she sing-ABS- pretend-PROG_PTCP-be.PRES.3SG
“He is pretending to sing.” (Pokharel 1991: 195)

Turner (1931: 247) suggests that a full verb counterpart may have once existed, possibly meaning “to cover” (cf. Nepali ṭopi “helmet”, cf. Hindi ṭopī “hat”, Hindi ṭopnā “to cover”), cf. (55).

(55) *u ṭopal-dai- cha
he/she pretend-PROG_PTCP-be.PRES.3SG
“He is pretending.”

The lack of a full verb counterpart for ṭopalnu demonstrates that the claim that light verbs always share a single lexical entry with their full verb counterparts cannot be maintained, if we assume that ṭopalnu is a light verb. However, given that ṭopalnu makes a rather more substantial semantic contribution than do most light verbs, it might be tempting to treat it as either a (modal-like) auxiliary, or else as a full lexical verb which obligatorily takes an absolutive as its complement.

A similar case is found in Hindi, where the element saknā “to be able to” selects for an absolutive complement, just as do light verbs in Hindi. Like Nepali ṭopalnu, Hindi saknā cannot appear as a main verb, in the sense that it obligatorily takes an absolutive complement.

(56) a. vah gā sakt-ā hai
he/she.NOM sing.ABS be_able.IMPF_PTCP-MSC.SG be.PRES.3SG
“He can sing.”

Nepali baksinu is ultimately a loanword from Persian bakhi “to give”, and obviously was borrowed as a full verb, not as a light verb, since, as shown by example (52), it can still be used as a full verb.
b. ‘vah sakt-ā hai
   he/she.NOM be_able.IMPF_PTCP-MSC.SG be.PRES.3SG
   “He can.”

Not only does saknā select for an absolutive complement, but it also behaves like a light verb in that it follows the Hindi rule of a maximum of one light verb per main verb (§4.1.2), i.e. saknā cannot co-occur with a(nother) light verb (see Nespital 1997: 1116).35

(57)  a. us-ne khānā khā liy-ā
      he/she-AGT food eat.ABS take.PERF_PTCP-MSC.SG
      “He/She ate up the food.”
  b. vah khānā khā sakt-ā hai
      he/she.NOM food eat.ABS be_able.IMPF_PTCP-MSC.SG be.PRES.3SG
      “He can eat food.”
  c. ‘vah khānā khā le sakt-ā hai
      he/she.NOM food eat.ABS take.ABS be_able.IMPF_PTCP-MSC.SG be.PRES.3SG
      “He can eat up food.”
  d. ‘vah khānā khā sak le-ā hai
      he/she.NOM food eat.ABS be_able.ABS take.IMPF_PTCP-MSC.SG be.PRES.3SG
      “He can eat up food.”

Thus both Nepali topalnau and Hindi saknā are ungrammatical as main verbs, and behave morphosyntactically like light verbs. The analysis of these elements in terms of categorical distinctions between lexical verb vs. light verb vs. auxiliary is therefore unclear, pointing again to inherent difficulties in distinguishing between different functional categories.

4.3. Conclusions

In summary, we have seen that the CV systems of Hindi and Nepali differ not only in their light verb inventories, but also in various aspects of morphosyntax.

Regarding morphosyntactic differences, §4.1.1 demonstrated that while in Hindi certain elements external to the CV can intervene between sequences of main verb-light verb, in Nepali no element may do so. This is the first piece of data suggesting that Nepali light verbs are more affixal in nature.

In §4.2, it was shown that Hindi and Nepali light verbs in CV constructions show only partial overlap in terms of light verb meanings; and in some cases, the same light verb meaning is realised but with each language using verbs with very different full verb senses (sti in Hindi vs. send in Nepali for the light verb sense of "regret"). Also not all light verbs have full verb counterparts—notably, the Nepali topal- has no (synchronic) full verb counterpart.

§4.1.2 presented data which show that Hindi allows for one light verb per main verb, while Nepali allows for at least two light verbs to occur with a single main verb. Here again, Nepali light verbs appear more affixal, since a Nepali CV behaves like a simplex verb in that it can combine with another light verb. With respect to construction-specific conditions on light verbs, as discussed in §4.1.3, Nepali CVs again behave like simplex verbs, in that they can appear in the same set of environments in which simplex verbs can appear—one again pointing to the affixal nature of Nepali light verbs.

Finally, with respect to dominance and transitivity marking, §4.1.4, Nepali light verbs again are found to act more like affixes than do Hindi light verbs, since the transitivity of the entire CV compound is always based solely upon the transitivity of the main verb in Nepali. In Hindi, conversely, it is usually the light verb which determines whether the CV is treated as transitive or intransitive for purposes of case-marking on the subject.

Thus, with respect to all four of these properties, Nepali light verbs in CV constructions appear more affix-like than do Hindi light verbs—whether this means that Nepali light verbs are ‘more grammaticalised’ than Hindi light verbs will be considered in §6.1.

35Other modal verbs, selecting for infinitive complements, do not have this restriction upon them:

(i) us-e khānā khā lenā cahiye.
    he/she-DAT food eat.ABS take.INF ought-to
    “He/she ought to eat up the food.”
The variation in morphosyntactic and lexical properties of light verbs in Nepali and Hindi is not characteristic of a 'stable system', contra Butt and Lahiri (2002), Butt (2010), etc. Change has taken place in the CV systems of one, or more likely, both of these languages.

5. Evolution from light verb to auxiliary

Indo-Aryan provides clear examples of light verbs becoming auxiliaries, specifically tense/aspect auxiliaries. In modern Hindi, rahnā "to stay, to remain", which functioned as a light verb in 19th-century Hindi, has become a grammatically central auxiliary. The case of Nepali rahanu "to stay, to remain", discussed in §5.2 is a particularly interesting example, as it is apparently a light verb which is currently in the process of becoming an auxiliary. The Indo-Aryan evidence demonstrates that light verbs can in fact become auxiliaries, filling the gap pointed to by Bowern (2008). Further, these changes, involving a grammatically-peripheral light verb contributing a durative/continuative sense being reanalysed as a core continuative aspect auxiliary represent a repetition of history: the auxiliary honā of the modern Hindi simple present (which was in earlier Hindi a present progressive construction) derives from an early IA use of be as a grammatically peripheral durative/continuative light verb.

It is important to note that these changes do not appear to take place via gradual grammaticalisation along a cline (cf. Hopper and Traugott 1993), but rather as part of a larger morphosyntactic chain shift, triggered by the reanalysis of old simple present forms as modal.

5.1. The development of the Hindi auxiliary rahnā

The Hindi auxiliary rahnā itself originated as a light verb. In Modern Hindi, the simple present and the present continuous are clearly distinguished, as shown in (58).

(58) a. main skul jatā hūm
I school go.IMPF_PTCP.MSC.SG be.PRES.1SG
"I (habitually) go to school."

b. main skul jā rahā hūm
I school go.ABS remain.PERF_PTCP.MSC.SG be.PRES.1SG
"I am going to school (just now)."

But Kellogg (1893. 2nd edn.: §404) cites forms like main jatā hūm as meaning either “I go” or “I am going”.36 Further, he categorises rahnā as a light verb (Kellogg 1893. 2nd edn.: §428) rather than an auxiliary, suggesting that—as in the case of Modern Nepali rahanu—as indicating continuous action in 19th c. Hindi was more peripheral and not yet integrated as part of the core grammar as a clear aspectual auxiliary; see (59) and (60), taken from Kellogg (1893. 2nd edn.: §404), retaining his translations.

(59) a. donom larke khelte the
both boy.PL play.IMPF_PTCP.MSC.PL be.PAST.3PL
"The two children were playing."

b. donom larke khele rahē the
both boy.PL play.ABS remain.PERF_PTCP.MSC.PL be.PAST.3PL
"The two children were engaged in play."

(60) a. vah suntā hai
he hear.IMPF_PTCP.MSC.SG be.3SG
"He hears." / "He is hearing."

b. vah sun rahā hai
he hear.ABS remain.PERF_PTCP.MSC.PL be.3SG
"He is occupied in hearing."

---

36 A situation which persists marginally in modern Hindi, in much the same way as the modern English simple present can be employed with a progressive sense in certain contexts, e.g. "I am attaching a document to this email..." vs. "I attach a document to this email...".
The later auxiliary nature of Hindi rahnā represents a reanalysis which is part of larger reconstructing of the Hindi verbal system; see §5.3, especially Table 4.

### 5.2. Nepali rahānu: light verb > aspectual auxiliary

Like baksinu (§4.2.2), Nepali rahānu “remain, stay” behaves morphologically like other light verbs, selecting for an absolutive participle in -i. Rahānu is also identical to other light verbs with respect to restrictions on recursion and construction-specific restrictions (see §4.1.2 and §4.1.3, respectively).

However, unlike other Nepali CV constructions, CVs involving rahānu (or rākhu “keep”) never require agentive-marking on the subject, regardless of the transitivity of the main verb. In this, the construction including rahānu in its light verb function appears to be a periphrastic counterpart of the Nepali synthetic imperfect construction, verb-stem + dai + inflected form of be. In other words, rahānu as a light verb, (61)—with respect to agentive casemarking assignment—paṇa pattern with the synthetic imperfect in -dai, (62), not with other CVs (such as the light verb di- as shown in (63)).

(61) a. ma mandir-mā ga-i- rah-eko chu
   I temple-LOC go-abs- remain-perf_ptcp.msc.sg be.1p.pres
   “I am going to the temple.” / “I have been going to the temple.”

   b. ma yo kāṃ gar-i- rah-eko chu
   I this work do-abs- remain-perf_ptcp.msc.sg be.1p.pres
   “I am doing this work.” / “I have been doing this work.”

(62) a. ma mandir-mā jāṁ-dai chu
   I temple-LOC go-impf_ptcp be.1p.pres
   “I am going to the temple.”

   b. ma yo kāṃ gar-dai chu
   I this work do-impf_ptcp be.1p.pres
   “I am doing this work.”

(63) a. ma mandir-mā gi-i- di-eko chu
   I temple-LOC go-abs- give-perf_ptcp.msc.sg be.1p.pres
   “I have gone to the temple (for someone).”

   b. mai-le yo kāṃ gi-i- di-eko chu
   I-agt this work do-abs- give-perf_ptcp.msc.sg be.1p.pres
   “I have done this work (for someone).”

In contrast to normal CV constructions (like those in (63)), constructions in which rahānu functions as a light verb do not require agentive marking on the subject (even when the main verb is transitive); see (61). In this, the CV construction with rahānu closely resembles the synthetic imperfect construction, as in (62).

The Nepali light verb rahānu appears thus to be in the process of becoming an auxiliary like Hindi rahnā,— see the Hindi examples in various tenses in (64).37

(64) a. maiṁ cal rahā hūṁ
   I move.abs remain.perf_ptcp.msc.sg be.pres.1sg
   “I am going.”

   b. maiṁ cal rahā thā
   I move.abs remain.perf_ptcp.msc.sg be.past.1sg
   “I was going.”

37Though, as a reviewer points out, at this point Nepali rahānu may be best analysed as an atypical light verb.
Likewise, in Nepali the construction absolutive particle + rahanu can be used in various tenses, a subset of which are shown in (65). (See Matthews (1998: 234-237) for more examples.)

(65) a. ma cal-i- raheko chu
   I move-abs- remain.PERF_PTCP.MSC.SG be.1SG.PRES
   “I am going.”

b. ma cal-i- raheko thieṁ
   I move-abs- remain.PERF_PTCP.MSC.SG be.1SG.PAST
   “I was going.”

c. ma cal-i- raheko huṁlā
   I move-abs- remain.PERF_PTCP.MSC.SG be.1SG.FUT
   “I (probably) shall be going.”

However, there are differences between the Hindi rahnā and Nepali rahanu constructions. Hindi rahnā can be used as an auxiliary in the present tense with future connotations, as in (66).

(66) kal main ghar jā rahnā hūṁ
   tomorrow I home go.Abs remain.PERF_PTCP.MSC.SG be.PRES.1SG
   “I’m going home tomorrow.”

In Nepali this is not the case. While the synthetic imperfective construction can be used in the present tense with future reference, as in (67a), the periphrastic absolutive + rahanu—which is formally similar to the Hindi construction in (66)—cannot; see (67b).

(67) a. bhōli ma ghara jāṁ-dai chu
   tomorrow I home go-IMPF.PTCP be.1SG.PRES
   “I am going home tomorrow.”

b. *bhōli ma ghara ga-i- raheko chu
   tomorrow I home go-ABS remain-PERF.PTCP.MSC.SG be.1SG.PRES

Thus Nepali rahanu has not acquired all of the properties of an auxiliary like Hindi rahnā since it cannot be used in future contexts, as can imperfects in -dai (Sharma 1980: 108–109). Further, it has not so far supplanted the synthetic -dai imperfective construction. Finally, rahanu exhibits some properties prototypical of light verbs used in CV constructions, some prototypical of auxiliaries—and therefore appears to represent a stage intermediate between light verb and auxiliary. Hindi rahnā combined with bare verb stems (morphologically identical to a CV construction) clearly functions as an auxiliary in present day Hindi (see (64)), but just as clearly acted more as a light verb as late as the 19th century, as discussed above. Present day Nepali rahanu, in its use in structures like those shown in (61) above, represents a stage intermediate between light verb and auxiliary—at least in the sense that it exhibits certain properties typical of light verbs and certain properties typical of auxiliaries.

5.3. Development of the Hindi simple present

The Hindi auxiliary rahnā “to stay, to remain” was shown to have evolved from a light verb, and the Nepali light verbs raḥanu/raḥkhu appear to be in the process of becoming auxiliaries. In fact, even the auxiliary honā “to be” as it appears in the modern Hindi simple present, e.g. (68), evolved in a fashion parallel to that of the auxiliary rahnā. Both changes appear to have involved a morphosyntactic chain shift (discussed further below)—thus not representing a gradual cline-like grammaticalisation process, but rather a structurally-induced change.

(68) jā-tā hai
   go-IMPF_PTCP.MSC.SG be.3SG
   “goes”
Hindi honā “to be” and its finite forms including hai “is” derive from earlier Sanskrit bhū- “become/be/exist” (e.g. bhavati “becomes”). One might suppose that this auxiliary use of honā developed directly from the full verb honā “to be”. But even in this case we find that hai in the simple present construction, (68), derives from an earlier light verb construction.

In §5.2, it was noted that forms like (68) were earlier (into the 19th c.) present progressives, and that the modern Hindi present progressives, (64), were recruited from a peripheral construction involving a light verb use of rahnā. The construction in (68), with its earlier present progressive value, derives from a structure represented in Classical Sanskrit by the type shown in (70) (cf. the simple present in (69)), as discussed in §3.1.

(69) calati
go.pres.3sg
“goes; is going”

(70) calan bhavati / asti
go.pres.ptcp be/become.pres.3sg / be/become.pres.3sg
“keeps on going”

However, recall from above (see (12)), that in Sanskrit, bhū- was only one of a number of verbs which could appear in this construction. Other verbs which could appear in this peripheral construction include ās- “sitting” and sthā- “standing”— whose (full verb) meanings are those typical of light verbs (Hock 2008).

In later IA (Apabhraṁśa), the construction present participle + be (cf. (70)) became grammaticalised as a progressive present, with the old simple present tending to be used as a non-continuative (though it is still possible in continuative contexts as well), as shown in (71) (see Singh 1980: 138).

(71) a. calai
go.pres.3sg
“goes; (is going)”

b. calantu acchāi
go.impf.ptcp.msc.sg be.pres.3sg
“is going”

In Hindi, the value of the earlier simple present (Skt. (69), Apa. (71a)) shifts from present to modal, leaving the periphrastic progressive present of Apabhraṁśa (71b)—itself a grammaticalisation of an earlier peripheral continuative/durative construction; see examples like (70)—as the new unmarked present indicative. This is the situation in 19th-century Hindi. By the 20th century, the Apabhraṁśa shift repeats, with the grammaticalisation of the peripheral continuative absolutive + rah- construction as the new progressive present. These changes are summarised in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>Core Grammar</th>
<th>Peripheral</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Modal</td>
<td>Simple Present</td>
</tr>
<tr>
<td>Stage A (Sanskrit)</td>
<td>calati</td>
<td>calati</td>
</tr>
<tr>
<td>Stage I (early Hindi)</td>
<td>cale</td>
<td>caltā hai</td>
</tr>
<tr>
<td>Stage II (19th-c. Hindi)</td>
<td>cale</td>
<td>caltā hai</td>
</tr>
<tr>
<td>Stage III (modern Hindi)</td>
<td>cale</td>
<td>caltā hai</td>
</tr>
</tbody>
</table>

Table 4. Morphosyntactic Chain Shift in Hindi (adapted from Hock 2008)

This ‘morphosyntactic chain shift’ resulted, ultimately, from the old simple present (cale-type) taking on modal functions, leaving a ‘gap’ in the paradigm which was filled by the present progressive (caltā hai-type)—which then bore two values (simple present and present progressive). This ambiguity is resolved with the reanalysis of the old
peripheral continuative construction of the type cal rahā hai as the (core-grammar) present progressive, as described above. For more on the this morphological chain shift in Hindi, see Hock 2008 (with reference to Bloch 1920).

As diagrammed in Table 4, the change between Stages II and III is essentially a repetition of the change which occurred between Stages A and I. In both cases, a peripheral construction, involving a light verb, is recruited into the core grammar, at which point the light verb is reanalysed as an auxiliary.

5.4. Conclusions

I have noted several examples of light verbs becoming auxiliaries in IA: (a) the reanalysis of the light verb (Skt. as-/bhū-; Apa. acch- “be”) of the durative/continuative present participle + light verb construction as the auxiliary used in the progressive tenses; (b) the reanalysis of the early Hindi light verb rah- in CV constructions as a progressive-marking auxiliary; (c) the ongoing reanalysis of the Nepali light verbs raha-, rakh- as progressive auxiliaries.38

6. Findings and implications

6.1. Grammaticalisation and the light verb/auxiliary contrast

Butt and Lahiri (2002: 12-13) use three major criteria to distinguish between light verbs and auxiliaries in IA: (1) (non)-interchangeability, (2) interruptibility, (3) ability to undergo reduplication.

With regard to the first, they mean simply that light verbs like jā- and auxiliaries like rah- are not grammatical in the same set of syntactic contexts. This seems a poor measure, since even in the case of elements which are clearly auxiliaries, different auxiliaries are grammatical in different syntactic contexts, e.g. the English auxiliaries have and be have distinct distributions.

The property of interruptibility refers to whether other elements may intervene in a V1 + V2 collocation (§4.1.1). Butt and Lahiri (2002) observe that in Hindi a main verb can be separated from a light verb, but they claim that the main verb may not be separated from its accompanying auxiliary and therefore that sentences like (72) are ungrammatical. However, my Hindi consultants in fact accept examples like (72b) as grammatical—this may reflect dialectal differences between the grammars of my consultants and Butt’s Hindi/Urdu.39

Finally, they observe that light verbs, but not auxiliaries, can undergo reduplication (Butt and Lahiri 2002: 13).

If we apply these tests to the whole range of light verb constructions in Hindi, the distinction between light verb and auxiliary becomes even murkier. Recall that in Hindi we also find various peripheral light verb constructions, including imperfect participle + rah-/jā- and perfect participle + kar- (see Hindi (17), (18), (19), (20); and (22) in fn. (20), respectively). These elements are interruptible, as shown by (73):

38 A reviewer objects that these changes could simply involve the light verbs undergoing a lexical change, whereas Bowern’s (2008) claim applies specifically to constructions. However, it is clear that these changes did in fact involve reanalysis of the light verbs in the context of particular constructions, as discussed in detail above. The same reviewer also proposes that the Nepali change (c) might be due to the influence of Hindi and thus be exempt from Bowern’s generalisation. There is no evidence which would suggest that ongoing changes raha- is experiencing are motivated by contact with Hindi; furthermore, the development of rakh- towards becoming a progressive auxiliary in unparalleled in Hindi. Hindi does possess a homophonous verb rakṣ-, which, in its main verb usages, is largely synonymous with the main verb uses of Nepali rakh-. Hindi rakṣ- is in fact also employed as a light verb in CV constructions (mostly contributing a sense that the action was completed some time ago, often with a certain purpose in mind relating to a later event; see Nespital 1997: 1032–1033), but is not employed in any progressive-auxiliary-like function.

39 A reviewer confirms that (72b) is grammatical.
“The child kept sleeping.”

b. \( \text{āy-ā} \) to \( \text{larki kar-ti} \) thī
dom-PFV-ABS.FEM.GEN girl-PFV-ABS.FEM.GEN be-PFV-ABS.FEM.GEN

“The girl used to come frequently . . .”

But the light verb construction perfect participle + \( \text{kar-} \) does not behave as we would expect it to with respect to dominance (see §4.1.4). That is, we would expect that, like other transitive light verbs (i.e. like the light verbs of Hindi CV constructions), it should require agential marking on the subject in perfective tenses—but it does not; see example (74). Rather, like the Hindi progressive present (absolute + \( \text{raha} \)), this construction always behaves as if the verb were in a non-perfective tense, for purposes of case-assignment.

\[
\begin{align*}
\text{(74) a.} & \quad \text{tere pūrvaj} \quad \text{ghās chil-ā} \quad \text{ki-e} \\
& \quad \text{your ancestors.MSC.PL grass scrape-PFV-ABS.MSC.PL do-PFV-ABS-MSC.PL} \\
& \quad \text{"Your ancestors scraped grass! (i.e. did menial work)" (McGregor 1995. 3rd edn.: 151)} \\
\text{b.} & \quad \text{\textquotesingle tere pūrvajōm} \quad \text{ne ghās chil-ā} \quad \text{ki} \\
& \quad \text{your ancestors.MSC.OBLQ.PL -ABS grass.FEM.MSC scrape-PFV-ABS.MSC do-PFV-ABS-MSC.FEM} \\
\end{align*}
\]

The situation becomes even more complicated once we include Nepali CVs and other light verb constructions. Table 5 summarises the relevant properties of seven types of auxiliary and light verb constructions in Hindi and Nepali. The morphosyntactic structure of each type is given, followed by an example, the semantic contribution of the auxiliary/light verb and two further properties. These properties are interruptibility—whether or not other elements may intervene between the V1 and V2—and dominance—which member of the collocation determines the overall transitivity for the purposes of case-marking (see §4.1.1 and §4.1.4, respectively). The possible values for this final property are: \( V_1 \) (=transitivity is determined based on the transitivity of \( V_1 \)), \( V_2 \) (=transitivity is determined based on the transitivity of \( V_2 \)) and ‘‘’ (=construction is always treated as a non-perfective tense).

Based on the data summarised in Table 5, how are we to classify these types as light verbs or auxiliaries? Butt and Lahiri (2002) consider only the contrast between types I and VII, categories which are perhaps disjoint in features (if we accept Butt’s judgement that such sequences are non-interruptible). Considering a fuller range of constructions, we find considerable intertype overlap of properties.

We cannot distinguish these grammatical elements on the basis of the morphological category of the \( V_1 \), since types both at the top and bottom of the table use an absolutive as the \( V_1 \)—and, in fact, all of the Nepali constructions involve an absolutive as \( V_1 \).

Interruptibility certainly does not distinguish light verb from auxiliary, since none of the Nepali constructions are interruptible, and Butt’s claimed non-interruptibility of type I elements is questionable (or at least subject to dialectal variation). Dominance fares no better, since it is irrelevant for types I–V (as they are treated as non-perfective), and types VI and VII—which we would like to classify together, as they are both CV constructions—differ on which member controls transitivity.

In sum, I see no single formal property or set of properties which suffices to distinguish light verbs from auxiliaries in both Hindi and Nepali, despite the great overall grammatical and lexical similarities of these two languages. At best we might divide the types of Table 5 into three sets: (1) types I–II, progressives; (2) types III–V, continuatives; (3) types VI–VII, CV constructions, with perfective/subjective senses—but note that this is a semantic/pragmatic

\[\text{[40]In both Hindi and Nepali agentic case-marking on the agent of a transitive verb is only required in perfective tenses. ‘Subjective’ refers to Tikkkanen’s (1987) ‘subjective aspect’, Abbi & Gopalakrishnan’s (1991) ‘attitudinal meaning’, and includes the addition of shades of meaning referring to the role or perspective of the speaker, e.g. Hindi light verbs de- “other-directedness”, baith- “regret”. I label this type as perfective/subjective for the sake of precision, for—though CVs usually have a perfective sense as shown by Hook (1974, 1991)—Butt and Lahiri (2002: 14) observe that CVs can occur where perfectivity is excluded, as in (i) below.}

\[
\begin{align*}
\text{(i) mariam inel likh mar rahī thi jab villī kamre men āy-ā} \\
& \quad \text{Miriam.FEM email.FEM write-ABS hit PROG.FEM.MSC be-PFV-ABS.FEM.MSC when Willi.MSC room.MSC.OBLQ in code-PFV-ABS.MSC.MSC} \\
& \quad \text{“Miriam was dashing off an e-mail when Willi came into the room.”}
\end{align*}
\]

\[\text{[41]But see §4.1.4 for complications involving transitive \( V_1 \) and intransitive \( V_2 \).}

\[\text{[42]The other two properties in §4.1, namely restrictions on recursion and construction-specific restrictions, were we to extend them to other light verb/auxiliary constructions, would only serve to distinguish type VI (which can occur in conjunctive participles and can undergo one level of recursion) from all other types.}\]
<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
<th>Semantics</th>
<th>Interrupt.</th>
<th>Dominance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ia. (Hindi) ABS + PERF_PTCP of rah-</td>
<td>cal raha hai</td>
<td>progressive</td>
<td>yes</td>
<td>—</td>
</tr>
<tr>
<td>Ia. (Hindi) ABS + PERF_PTCP of rah-</td>
<td>cal raha hai</td>
<td>progressive</td>
<td>no</td>
<td>—</td>
</tr>
<tr>
<td>Ib. (Bu' s Hindi/Urdu)</td>
<td>cal raha hai</td>
<td>progressive</td>
<td>no</td>
<td>—</td>
</tr>
<tr>
<td>Ib. (Bu' s Hindi/Urdu)</td>
<td>cal raha hai</td>
<td>progressive</td>
<td>no</td>
<td>—</td>
</tr>
<tr>
<td>II. (Nepali) ABS + PERF_PTCP of raha-</td>
<td>cali-raheko cha</td>
<td>progressive</td>
<td>no</td>
<td>—</td>
</tr>
<tr>
<td>III. (Hindi) IMPF_PTCP + raha-</td>
<td>cal tā rathā hai</td>
<td>continuative (durative)</td>
<td>yes</td>
<td>—</td>
</tr>
<tr>
<td>IV. (Nepali) ABS + finite form of raha-</td>
<td>cali-rahancha</td>
<td>continuative (durative)</td>
<td>no</td>
<td>—</td>
</tr>
<tr>
<td>V. (Hindi) PERF_PTCP + kar-</td>
<td>cala kartā hai</td>
<td>continuative (habitual)</td>
<td>yes</td>
<td>—</td>
</tr>
<tr>
<td>VI. (Nepali) ABS + light verb</td>
<td>cala diyo</td>
<td>perfective/subjective</td>
<td>no</td>
<td>V₁</td>
</tr>
<tr>
<td>VII. (Hindi) ABS + light verb</td>
<td>cala gaya</td>
<td>perfective/subjective</td>
<td>yes</td>
<td>V₂</td>
</tr>
</tbody>
</table>

Table 5. Classification of selected Hindi and Nepali light verbs & auxiliaries

classification, not a formal one. And even this classification involves a degree of overlap: types III-V include different types of continuatives (durative vs. habitual), and type II, unlike type I, cannot be used with future reference (see (66), (67)).

Though it may be tempting to view the data in Table 5 as evidence for a grammaticalisation continuum—with type I most auxiliary-like and fully grammaticalised and type VII most light verb-like and least grammaticalised—the multiply overlapping morphosyntactic properties and semantic/pragmatic functions suggests that the types in Table 5 simply represent different grammaticalisations. Nepali CVs (type VI) developed in such a way that they are non-interruptible, unlike Hindi CVs (type VII); while non-interruptibility certainly seems to be a property associated with more grammaticalised elements, I would not want to claim that Nepali CVs are more grammaticalised than, say, the Hindi durative construction cali tā rathā hai (type II), which is interruptible. With respect to semantic bleaching and obligatoriness, Nepali CVs have the same status as Hindi CVs, though they display different morphosyntactic properties.

Grammaticalisation appears then not to be a one-dimensional process—rather linguistic elements may undergo grammaticalisation along various axes, and an element grammaticalised with respect to certain features is not necessarily grammaticalised with respect to others. Thus, a one-dimensional representation of grammaticalisation, like the cline shown in (75) may be misleading.43

(75) (a) parataxis > (b) hypotaxis > (c) serialisation > (d) light verb > (e) auxiliary verb > (f) univerbated affix (Andersen 2005: 331)

For instance, based on certain criteria (i.e. the four properties discussed in §4.1, including non-interruptibility), Nepali light verbs appear to be univerbated affixes; but semantically and pragmatically they behave like Hindi light verbs (which based on the same criteria, are definitely not univerbated affixes), and therefore do not have the same status as, for instance, an obligatory progressive aspect affix (e.g. type I). Another ‘grammaticalisation axis’ we might consider is the continuum between core-grammar and peripheral constructions (see §3.2), a distinction referred to throughout this study. As mentioned earlier, types II-V represent peripheral constructions, and types I, VI, VII represent core-grammar constructions—a classification which cuts across many of the other distinctions discussed above. Peripheral constructions can be considered less grammaticalised in the sense that they are less fully integrated into the grammar. From this perspective, types I, VI, VII are more fully grammaticalised than types II-V, despite the

fact that these two groupings are heterogeneous with respect to other properties considered.

Viewing grammaticalisation in this multidimensional way offers an explanation of why it is difficult to divide the types in Table 5 into two discrete categories of light verb and auxiliary. It might be useful in a synchronic description—of a particular language—to distinguish between light verbs and auxiliaries, so long as these are understood to be prototypical categories and not universal primitives. However, it may be more useful to consider these types of grammaticalised verbal elements in terms of clusters of properties—clusters which turn out, at least in the case of Nepali and Hindi, to involve a large degree of (multiply) overlapping properties, as diagrammed in Table 5.

In terms of such clusters of properties, the Nepali grammaticalised verbal elements are less distinguished from one another in terms of formal properties—all are constructed from an absolutive $V_1$ and none are interruptible, whereas the Hindi grammaticalised verbal elements exhibit less overlap of properties, showing variation both in the morphology of the $V_1$ and with respect to interruptibility, and thus can be more easily divided into notional discrete categories.

6.2. Conclusion

I have investigated the origin and development of the IA CV construction. Though Butt and Lahiri (2002) point to the existence of grammaticalised verb-verb constructions in Old IA, a closer examination of these constructions reveals that—though they do exist, as part of the more “peripheral” grammar—they are formally and functionally distinct from the CV constructions of modern IA. The earliest unequivocal examples of CVs are found in Early Modern IA (17th-18th c. C.E.). An examination of two closely related modern IA languages, Hindi and Nepali, reveals that numerous morphosyntactic and lexical changes have taken place in these languages, suggesting that light verbs, like all other elements, are not historically stable but rather are subject to reanalysis and change (pace Butt and Lahiri (2002) and Butt (2010)), including reanalysis as auxiliaries. Finally, Indo-Aryan fills in the data gap Bowern (2008) points to, as such changes include reanalysis of light verbs as auxiliaries.

The reanalysis of the descendant of the Sanskrit light verb bhu- “be” as an aspectual auxiliary in early Hindi, as well as the parallel reanalysis of the Hindi light verb raha- as an aspectual auxiliary are both noteworthy from the standpoint that they do not take place via grammaticalisation along a cline, but instead as part of larger, structurally-motivated morphosyntactic chain shifts. Such changes highlight the importance of looking beyond the grammaticalisationist paradigm when evaluating historical changes.

Much more work remains to be done on IA CVs. The most complete collections of data and examinations of the properties of IA CVs are those of Hindi (e.g. Nespital (1997), Hook (1974)). A more complete picture requires much more data and analysis from other IA languages.\[^{44}\]

References


\[^{44}\]A careful consideration of Apabhraṁśa data is also highly desirable, given the possible examples of Apabhraṁśa CVs cited by Singh (1980), Bubenik (1998), and others.


**Résumé**

Cette étude examine le développement historique des verbes légers en indo-aryen. On explore les origines, en indo-aryen moderne, de constructions avec verbe composé, pour ensuite comparer ces constructions avec d’autres constructions verbales à verbe léger en indo-aryen. L’examen des formes anciennes du verbe composé et la comparaison avec d’autres structures à verbe léger en indo-aryen, ainsi que l’analyse des différences lexicales et morphosyntaxiques entre les systèmes des verbes composés de deux langues indo-aryennes (le hindi et le népalais), montrent que les verbes légers ne sont pas une partie stable ou invariable de la grammaire, mais subissent plutôt une série de changements, y compris la réanalyse comme auxiliaires marquant le temps et/ou l’aspect.

**Zusammenfassung**


**Author’s address**

Benjamin Slade
Department of Linguistics
University of Utah
Languages and Communication Building
255 South Central Campus Drive, Room 2300
Salt Lake City, UT 84112-0492

b.slade@utah.edu