

# Diachronic pertinacity of light verbs

Miriam Butt <sup>a,\*</sup>, Aditi Lahiri <sup>b</sup>

<sup>a</sup> University of Konstanz, Germany

<sup>b</sup> University of Oxford, United Kingdom

## Abstract

We contrast the historical data with respect to light verbs and auxiliaries in Indo-Aryan and show that light verbs are comparatively stable and unlikely to be subject to reanalysis or restructuring. We propose that there is a very tight connection between a light verb and its corresponding main verb, and that this connection differs markedly from the relationship an auxiliary bears to the main verb it is derived from. In particular, we depart from the received view that the existence of a light verb is due to a historical process of semantic bleaching. We instead propose that synchronically there is a single underlying lexical entry which tightly binds light verbs to their corresponding main verb. This proposal accounts not only for the simultaneous synchronic uses of light and main verbs, but also for the historical data.

*Keywords:* Urdu; Bengali; Auxiliaries; Historical pertinacity; Light verbs; Complex predicates; Reanalysis; Grammaticalization

## 1. Introduction

Complex predicate formation is an integral part of South Asian languages (see Chatterji, 1926; Masica, 1976 on South Asia as a linguistic area). We concentrate on V–V complex predicates of the type in (1).<sup>1,2</sup>

- (1) a. nadya=ne xat **lik<sup>h</sup> di-ya**  
Nadya.F=Erg letter.M.Nom write give-Perf.M.Sg  
'Nadya wrote a letter (completely).' (Urdu)
- b. nadya **a ga-yi**  
Nadya.F.Nom come go-Perf.F.Sg  
'Nadya has arrived.' (Urdu)

\* Corresponding author at: FB Sprachwissenschaft, Universität Konstanz, Fach 184, 78457 Konstanz, Germany. Tel.: +49 7531 885109; fax: +49 7531 884865.

E-mail address: [miriam.butt@uni-konstanz.de](mailto:miriam.butt@uni-konstanz.de) (M. Butt).

<sup>1</sup> The list of abbreviations used in this paper is as follows: A = Atmanepadam, Acc = Accusative, Aug = Augment, Caus = Causative, Cl = Classifier, Compl = Complement, Dat = Dative, Dem = Demonstrative, Emph = Emphatic, Erg = Ergative, F = Feminine, Fut = Future, Gen = Genitive, Imp = Imperative, Impf = Imperfect, Ind = Indicative, Inst = Instrumental, M = Masculine, Nom = Nominative, Obl = Oblique, Obj = Objective, Opt = Optative, P = Parasmaipadam, PP = Past Participle, Perf = Perfect, Pl = Plural, Pres = Present, Pron = Pronoun, Ptcp = Participle, Redup = Reduplication, Rel = Relative, Sg = Singular, Top = Topic, Voc = Vocative, W = Weak. A '-' indicates a morpheme boundary, a '=' a clitic boundary.

<sup>2</sup> Note that this is a revised and updated version of a paper whose previous incarnations were called 'Historical stability vs. historical change' and 'Verbal Passepartouts'.

In Urdu the form of the first verb is identical to the stem form and never carries any inflection in the V–V construction. This is the main or full verb. The second verb carries tense/aspect marking and is inflected according to the standard paradigms governing simple verb inflection. This verb is *light* in the sense that although it is form-identical to a main verb (and hence glossed with the main verb meaning in the close gloss), the predicational contribution is not that of a main verb (cf. [Jespersen, 1965](#), who first coined the term *light verb*). Rather, it serves to modify the main verb semantics by expressing such notions as completion, inception, benefaction, forcefulness, suddenness or volitionality ([Hook, 1974](#)). The parameters of completion and inception are responsible for the pervasive notion in the literature that these light verbs are aspectual (e.g., [Hook, 1991](#); [Singh, 1994](#)).

We concentrate on the South Asian languages Urdu/Hindi<sup>3</sup> and Bengali in particular (see [Zbavitel, 1970](#); [Hook, 1974](#); [Ramchand, 1990](#); [Singh, 1994](#); [Butt, 1995](#) among others).<sup>4</sup> Some Bengali examples are shown in (2). The form of the first verb in the sequence carries what has often been dubbed “perfective” morphology. We gloss this morpheme as a participle, which reflects its Old Indo-Aryan origin.<sup>5</sup>

- (2) a. ram        bag<sup>h</sup>-ṭa-ke    mer-e p<sup>h</sup>el-l-o  
       Ram.Nom tiger-Cl-Acc hit-Gd throw-Past-3  
       ‘Ram killed the tiger.’ (Bengali)
- b. ram        ef-e        por-l-o  
       Ram.Nom come-Gd fall-Past-3  
       ‘Ram arrived.’ (Bengali)

The first verb in the sequence is again the main verb. The second verb carries tense marking and is inflected according to the standard paradigms governing simple verb inflection. Bengali light verbs in V–V constructions function like the Urdu ones in terms of syntax and semantics.

In this paper, we argue that there is a very tight connection between a light verb and its corresponding main verb, and that this connection differs markedly from the relationship an auxiliary bears to the main verb it is derived from. We depart from the received view that the existence of a light verb is due to a historical process of *semantic bleaching* (e.g., [Hopper and Traugott, 1993](#); [Hook, 1991](#)) and instead propose that there is a single underlying lexical entry, which can account not only for the simultaneous synchronic uses of light and main verbs, but also for the historical data.

[Section 2](#) first revisits the evidence that light verbs form a syntactic class that is distinct from auxiliaries and main verbs. We establish that light verbs are characterized by the following properties: they co-predicate with a main verb in a syntactically monoclausal construction and they serve to semantically modulate or structure the event predication of the main verb. In contrast to auxiliaries, they do not situate the event predication temporally or aspectually, i.e., they cannot be seen as instances of I or T, but must be analyzed as a subclass of V. We thus work with a narrower, but more precise understanding of light verbs than is often assumed in the literature.

[Section 3](#) presents historical evidence showing that light verbs tend to be stable or *pertinacious* in diachronic terms. This is again a property that sets light verbs apart from auxiliaries, which in contrast are very mutable across time. In [section 4](#), we discuss our proposals with respect to the single underlying lexical entry that accounts for the form-identical light and main verb uses synchronically and across time. [Section 5](#) provides a comparative look at data from other language families and discusses existing views which contrast with ours. [Section 6](#) concludes the paper.

## 2. Syntactic status

There are several important points to be made about the syntactic status of light verbs. For one, light verbs enter into a copredicational domain with the main verb. That is, the main verb and the light verb together function as the predicate of the clause. [Butt \(1995\)](#) presents detailed argumentation involving anaphora resolution, control, and verb agreement which shows that the V–V constructions examined here are monoclausal in the sense that the clause contains only a single subject, a single object, etc. This contrasts with constructions as in (3), which are similar, but consist of a main verb (‘go’) and an embedded clause (‘having written a letter’).

<sup>3</sup> The South Asian languages Urdu and Hindi are closely related. Both are among the 18 official languages of India and are spoken primarily in the north of India. Urdu is the national language of Pakistan. The data presented in this paper are drawn primarily from the dialect of Urdu spoken in Lahore, Pakistan, as well as from examples cited in the literature on both Urdu and Hindi.

<sup>4</sup> Note that while complex predication exists in Dravidian languages, the morphosyntax of those constructions differs quite markedly from that of the Indo-Aryan languages. It is not possible to address these issues competently within the scope of this paper and we therefore concentrate exclusively on examples from Indo-Aryan.

<sup>5</sup> The ancestral construction functioned like a participial and was dubbed a gerund in the Western tradition of Sanskrit scholarship.

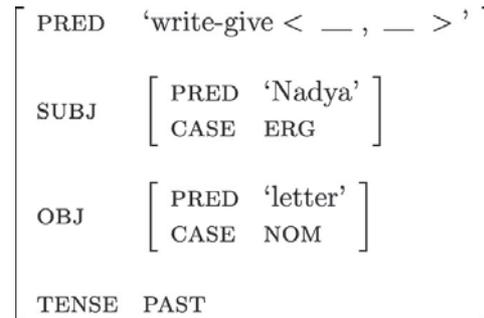


Fig. 1. Monoclausal F-structure of complex predicate.

- (3) nadya [xat lik<sup>h</sup> kar] ga-yi  
 Nadya.F.Nom letter.M.Nom write having go-Perf.M.Sg  
 'Nadya left, having written a letter.' (Urdu)

In order to be maximally clear about the contrast between monoclausal vs. biclausal constructions, we here show the relevant differences in terms of an LFG (Lexical-Functional Grammar) analysis that is based on Butt (1995). The f(unctional)-structure in Fig. 1 corresponds to the monoclausal complex predicate in (1), the f-structure in Fig. 2 to the biclausal sentence in (3). The precise details of the LFG analysis are not important here (for details see Butt, 1995). Important is that complex predicates correspond to a monoclausal functional structure, not a biclausal one.

Since complex predications are functionally monoclausal, there is a tendency in the literature to analyze light verbs as being analogous to auxiliaries. That is, as being functional elements. However, this tendency only serves to obscure an important syntactic dimension: it can be shown for language after language that light verbs form a distinct class that behaves quite differently from that of auxiliaries (Butt, 2003, 2010). Some of the relevant evidence is reviewed in the next sections.

### 2.1. Light verbs vs. auxiliaries

Analyses across the theoretical spectrum have tended to analyze light verbs as a type of auxiliary. In recent developments within the Minimal Program (MP), for example, v has been regarded as a natural position for light verbs (Adger, 2003, 134), particularly as it is associated with a notion of cause and can thus be used to account for causative formation in languages like Romance, where the causative verb 'make' has been analyzed as a light verb (Rosen, 1989; Alsina, 1996). Butt and Ramchand, 2005 propose v as the relevant position for Urdu/Hindi light verbs of the type investigated here as well as for permissives.

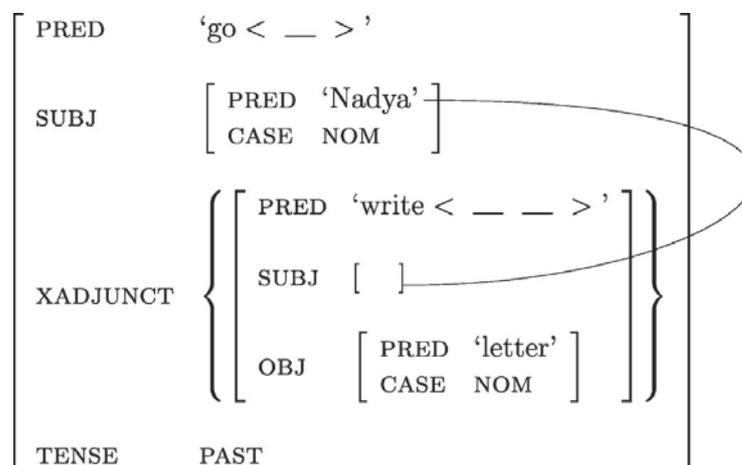


Fig. 2. F-structure of biclausal sentence.

However, *v* is not universally reserved for light verbs. The position has, for example, also been associated with auxiliaries (Chomsky, 1957) and (root) modals (e.g., Roberts and Roussou, in press, 47). In Head-Driven Phrase Structure Grammar (HPSG) Abeillé et al. (1998) analyze both tense auxiliaries and causative *faire* constructions in French as complex predicates, thus conflating a causative light verb (*faire* ‘make’) with other auxiliaries in the language.<sup>6</sup> Within the more functionally oriented literature, Hook (1991), for example, sees light verbs as a stage in the on-going grammaticalization from main verbs to auxiliaries.

We would therefore like to make clear that light verbs need to be recognized as a distinct syntactic class. The point can be illustrated quite straightforwardly with respect to Urdu, where auxiliaries and light verbs show distinct syntactic behaviors with regard to case marking, word order, reduplication and topicalization. In this paper, we restrict ourselves to presenting evidence from word order and reduplication (see Butt and Geuder, 2001 for further evidence).

Urdu/Hindi is an SOV languages with fairly free word order. The verb may expand into a relatively large verbal complex, as shown in (14).<sup>7</sup> Word order within the verbal complex is very rigid and auxiliaries and light verbs do not appear in the same positional slot. As shown in (4), light verbs and auxiliaries are not interchangeable.

(4) Main Verb (Light Verb) (Passive) (Aspectual Auxiliaries)  
(Tense Auxiliary)

(5) a. batʃʃa so dʒa rah-a hɛ  
child.M.Nom sleep go stay-Perf.M.Sg be.Pres.3.Sg  
‘The child is going to sleep.’ (Urdu)

b. \*batʃʃa so rah ga-ya hɛ  
child.M.Nom sleep stay go-Perf.M.Sg be.Pres.3.Sg  
‘The child is going to sleep.’ (Urdu)

A reviewer notes that the ordering rigidity exhibited in the Urdu verbal complex does not by itself necessarily argue against an analysis of the light verb as an auxiliary. The light verb could in principle be a special kind of auxiliary that is restricted to a certain slot, just as the aspectual and tense auxiliaries are (and as found in the English auxiliary system). This is true, which is why this is not the only evidence we adduce (see below). However, note the position of the passive auxiliary and also note that Urdu has a morphological causative which attaches to the verb stem (see Saksena, 1982 for details). The parts of the verbal complex which effect changes in the argument structure of the main verb, i.e., passives and causatives, are closer to the verb than the aspectual and tense auxiliaries. In Minimalist analyses this would follow nicely from the assumption that aspectual auxiliaries are situated in AspP and tense auxiliaries in TP/IP, which are both considered to be above the vP/VP, which in turn would contain the argument changing elements such as the light verb, the passive auxiliary and the causative morpheme (e.g., see Julien, 2002 for a crosslinguistic perspective and Ramchand, 2008 for Urdu/Hindi in particular). The position of the light verb as occurring close to the verb and before the passive auxiliary thus suggests that it is not situating the event described by the verb, as the tense and aspect auxiliaries do, but is part of the vP/VP in which the overall argument structure of the verb is determined.

Another difference between light verbs and auxiliaries is found in the possibilities for reduplication. In complex predicates, either the light verb or the main verb or both may be reduplicated (see Fitzpatrick-Cole, 1994, 1996 for a detailed study on Bengali, cf. Abbi, 1991). Examples (6) and (7) illustrate reduplication of the main and the light verb in Urdu, respectively. An auxiliary may not be similarly reduplicated, as (8) and (9) show.

(6) a. vo so-ti t<sup>h</sup>-i  
Pron.3.Sg.Nom sleep-Impf.F.Sg be.Past-Sg.F  
‘She used to sleep.’ (Urdu)

b. vo so-ti voti t<sup>h</sup>-i  
Pron.3.Sg.Nom sleep-Impf.F.Sg sleep.Redup be.Past-Sg.F  
‘She used to sleep (and such like things).’ (Urdu)

<sup>6</sup> Most of the work that we are aware of within HPSG treats light verbs as either auxiliaries or as a variation on raising. One exception is Choi and Wechsler’s (2001) analysis of Korean N–V formations with ‘do’. Here, the light verb follows the general schema of a raising verb, but special mechanisms are introduced to reflect the monoclausal properties of the construction.

<sup>7</sup> Indeed, (4) is a simplification of the actual situation because the slot occupied by the main verb could also be occupied by an N–V or Adj–V complex predicate – the light verb slot shown in (4) would stack on top of these. Noun incorporation structures could also appear in the main verb slot and we have not said anything about the interaction with other types of complex predicates like the Urdu permissive (see Butt, 1995; Butt and Ramchand, 2005). Detailing all these interactions, including a more fine-grained look at the set of available aspectual auxiliaries (Butt and Rizvi, 2010) would take us too far afield from the central point, which is that the distribution of auxiliaries does not overlap with those of the light verbs.

- (7) a. vo so **dʒa-ti** t<sup>h</sup>-i  
 Pron.3.Sg.Nom sleep go-Impf.F.Sg be.Past-Sg.F  
 ‘She used to go to sleep.’ (Urdu)
- b. vo so **dʒa-ti vati** t<sup>h</sup>-i  
 Pron.3.Sg.Nom sleep go-Impf.F.Sg go.Redup be.Past-Sg.F  
 ‘She used to keep going to sleep (at inopportune moments).’ (Urdu)
- (8) a. vo so **rah-i** t<sup>h</sup>-i  
 Pron.3.Sg.Nom sleep stay-Perf.F.Sg be.Past-Sg.F  
 ‘She was sleeping.’ (Urdu)
- b. \*vo so **rah-i vahi** t<sup>h</sup>-i  
 Pron.3.Sg.Nom sleep stay-Perf.F.Sg stay.Redup be.Past-Sg.F  
 ‘She was sleeping.’ (Urdu)
- (9) a. vo so-ti t<sup>h</sup>-i  
 Pron.3.Sg.Nom sleep-Impf.F.Sg be.Past-Sg.F  
 ‘She used to sleep.’ (Urdu)
- b. \*vo so-ti t<sup>h</sup>-i si  
 Pron.3.Sg.Nom sleep-Impf.F.Sg be.Past-Sg.F be.Redup  
 ‘She used to sleep.’ (Urdu)

Finally, unlike auxiliaries, light verbs always span the entire verbal paradigm. That is, while auxiliaries tend to appear in only part of the verbal paradigm, light verbs may appear in any tense/aspect (cf. Butt and Rizvi, 2010). For example, the present and future tense are expressed by the auxiliary version of *ho* ‘be’ (cf. (5)). This form inflects for number and person. However, the past tense is expressed by a suppletive form *t<sup>h</sup>-*, which inflects for number and gender (see examples above). The verb *rah-* ‘stay’ forms the simple progressive by using the morphology that usually expresses perfect/past: *rah-a/i/e*. There are good historical explanations for these seemingly confusing patterns and the historical explanations are entirely in line with what is known about the grammaticalization of verbs into tense/aspect markers crosslinguistically (Bybee et al., 1994). However, defective paradigms of idiosyncratic patterns of verbal marking are never found with light verbs. Light verbs inflect just like their main verb counterparts. Light verbs and auxiliaries thus have a very different distribution.

## 2.2. Light verbs vs. main verbs

Light verbs also contrast with main verbs in terms of distribution and syntactic behaviour. In this section, we discuss some representative data to show that light verbs cannot be analyzed as main verbs.

Recall that V–V sequences are sometimes ambiguous between a complex predicate and an embedded adverbial structure. An overt realization of *kar* ‘having’ precludes any potential disambiguity, as it is felicitous with an embedded adverbial structure but not a complex predicate, as shown in (10).<sup>8</sup>

- (10) a. nadya=ne abu=ko **makan**  
 Nadya.F=Erg father.M.Sg=Dat house.M.Nom  
 [bana \*kar **di-ya**]  
 make having give-Perf.M.Sg  
 ‘Nadya built a house (completely) for (her) father.’ (Urdu)
- b. nadya=ne abu=ko [**makan** bana kar] **di-ya**  
 Nadya.F=Erg father.M.Sg=Dat house.M.Nom make having give-Perf.M.Sg  
 ‘Having built a house, Nadya gave (it) to (her) father.’ (Urdu)

It is possible to insert another direct object in the adverbial construction, but not in the complex predicate one. This is because the adverbial construction in (11b) consists of two independent verbs, one of which (*bana* ‘make’) heads an

<sup>8</sup> Example (10) is well-formed (in the right context) because Urdu is a pro-drop language which allows any and all arguments to be dropped as long as they can be recovered from the discourse. This type of pro-drop is characteristic of South Asian languages in general.

embedded clause whose subject is controlled by the matrix subject. The complex predicate, on the other hand, is an instance of copredication where both the verbs combine to provide a single predicational head and the light verb cannot act as an independent predicator. As such, there is no subject control, no embedded clause, and only one possible object. This is shown in (11).

- (11) a. \*nadya=ne abu=ko **makan** bana p̄ɛisa di-ya  
Nadya.F=Erg father.M.Sg=Dat house.M.Nom make money.M.Nom give-Perf.M.Sg  
'Nadya built (her) father a house (completely) money.'  
(Urdu)
- b. nadya=ne abu=ko [**makan** bana kar] p̄ɛisa di-ya  
Nadya.F=Erg father.M.Sg=Dat house.M.Nom make having money.M.Nom give-Perf.M.Sg  
'Having built a house, Nadya gave (her) father money.'  
(Urdu)

Another difference between monoclausal complex predicates and biclausal constructions such as in (11b) is prosodic. Complex predicates are known to have special prosodic properties in Urdu and Bengali (Bayer and Lahiri, 1990; Hayes and Lahiri, 1991; Lahiri and Fitzpatrick-Cole, 1999; Fitzpatrick-Cole, 1996) in that the two members of the joint predication are contained within one phonological phrase. Relevant contrasts from Urdu and Bengali are shown in (12) and (13), respectively.<sup>9</sup>

- (12) a. nadya=ne xat (lik<sup>h</sup> di-ya)<sub>φ</sub>  
Nadya.F=Erg letter.M.Nom write give-Perf.M.Sg  
'Nadya wrote a letter (completely).'  
(Urdu)
- b. nadya=ne xat (lik<sup>h</sup> (kar))<sub>φ</sub> (di-ya)<sub>φ</sub>  
Nadya.F=Erg letter.M.Nom write having give-Perf.M.Sg  
'Having written a letter, Nadya gave (it).'  
(Urdu)
- (13) a. j̄æmoli arjola-ʒa-ke (mer-e p<sup>h</sup>el-e-tʃ<sup>h</sup>-e)<sub>φ</sub>  
Shamoli.F.Nom coakroach-CI-Acc hit-Gd throw-Perf-be-3  
'Shamoli killed the coakroach (completely).'  
(Bengali)
- b. j̄æmoli arjola-ʒa-ke (mer-e)<sub>φ</sub> (p<sup>h</sup>el-e-tʃ<sup>h</sup>-e)<sub>φ</sub>  
Shamoli.F.Nom coakroach-CI-Acc hit-Gd throw-Perf-be-3  
'Having killed the coakroach, Shamoli threw it.'  
(Bengali)
- c. j̄æmoli arjola-ʒa-ke (mer-e)<sub>φ</sub> (p<sup>h</sup>el-e di-e-tʃ<sup>h</sup>-e)<sub>φ</sub>  
Shamoli.F.Nom coakroach-CI-Acc hit-Gd throw-Gd give-Perf-be-3  
'Having killed the coakroach, Shamoli threw it (away).'  
(Bengali)

Light verbs thus have properties which differ markedly from those of main verbs.<sup>10</sup> In the next section, we turn to evidence which shows that light verbs must also be differentiated from serial verbs.

<sup>9</sup> Native Bengali speakers would prefer to disambiguate (13b) more clearly by adding a light verb, as in (13c). Urdu/Hindi native speakers prefer the explicit use of *kar* 'having' as in (12b) in the biclausal context.

<sup>10</sup> At this point, one reviewer wonders about tests for monoclausality. As mentioned in section 2, Butt (1995) adduces several arguments for the monoclausality of these V–V complex predicates based on anaphora resolution, control and agreement. The arguments have already been repeated in several papers and the reader is referred to Butt (1997, 2010), Butt and Ramchand (2005). The reviewer also wonders whether modification with time adverbials or negation can be used as diagnostics. This raises a point that is well worth making. We have found that time adverbials and negation are tricky to use as diagnostics for *syntactic* monoclausality as what they pick out are (sub)events of a given predication. With regard to the V–V complex predicates it is indeed not possible to use time adverbials in the same clause (e.g., \*Nadya today build-gave a house for her father tomorrow.). However, it is marginally possible to use two time adverbials in Urdu permissives, which are also syntactically monoclausal by the syntactic tests (agreement, control, anaphora) (e.g., Nadya yesterday allowed her father to build a house next year.). The adverbial test is thus not conclusive. Negation, interestingly, can be used to pick out just the light verb in the V–V complex predicates, though it is to date not clear just what is being negated semantically. Negation of Urdu/Hindi and Bangla V–V complex predicates in general is actually very restricted and only holds under a set of conditions which are at present not understood. Negation thus shows that the V–V complex predicates are special in some sense, but the test is again not conclusive as the monoclausal Urdu permissives do allow negation of each subevent (e.g., Nadya didn't allow her father to build a house. vs. Nadya allowed her father to not build a house.). For a good discussion of some problems with these diagnostics with respect to the monoclausal vs. biclausal status of English *have*-constructions (as in Nadya had her father build a house.), see Tantos (2008).

### 2.3. Light verbs vs. serial verbs

The complex predicates investigated here consist of V–V sequences. A comparison could therefore be drawn to another well known multiverb phenomenon: that of serial verbs. This section draws a distinction between prototypical complex predicates and prototypical serial verbs. A quick comparison of one typical sort of serial verb from Akan in (14) with a typical Urdu V–V complex predicate in (15) illustrates some concrete syntactic and semantic differences. Serial verbs typically stack several events in a single clause. This is not possible for complex predicates. The light verb *le* ‘take’ in (15) merely contributes additional information to the extracting event, but does not contribute an event of its own. In (14), on the other hand, there are three distinct events: a getting event, a building event and a selling event.

(14) Gyasiba nyá-à sika sí-i dan ton-èè  
 Gyasiba get-Compl money build-Compl house sell-Compl  
 ‘Gyasiba got money, built a house and sold it.’ (Akan)  
 (Osam, 2003)

(15) anjum=ne pat<sup>h</sup>ar=ko bahar nikal li-ya  
 Anjum.F.Sg=Erg stone.M=Acc out extract take-Perf.M.Sg  
 ‘Anjum extracted the stone.’ (Urdu)

Syntactically, the kind of object sharing made famous by Baker (1989) for serial verbs is not a characteristic of complex predicates. Another well-known property associated with serial verbs is the sharing of a single subject and the requirement that each verb be separately and uniformly marked for tense/aspect. This is true for Akan, for example. It is also true for Paamese and (16) provides an example which makes the point very nicely: each member of the serial verb in Paamese displays agreement features which have to be consonant with one another.

(16) iire rehe-sooni vakillii rehe-haa  
 1Pl.Incl 1Pl.Incl-Distant.Throw canoe 1Pl.Incl-Distant.Go  
 ‘We will go, putting (throwing) our canoe to sea.’ (Paamese)  
 (Crowley, 1987:47)

This property is not exhibited by typical complex predicates: only one of the members of the verbal complex carries tense/aspect or agreement features. This in turn indicates that what is at stake is a difference at the level of event structure: complex predicates describe one event, whereas serial verbs denote a complex event.

We follow Butt and Ramchand (2005) and Butt and Geuder (2001) in analyzing the function of light verbs as providing a modulation or modification of the main event semantics. In contrast, as shown by Durie (1997), serial verbs typically appear to denote a *complex* conceptual event. Serial verbs can thus only be used to describe events which are “normal” in a given cultural context. For example, *buy take fish* is an expected sequence of events, while *sell take fish* is not. In Sranan the former therefore exists as a serial verb, while the latter does not. Another example comes from Alambak (Bruce, 1988), where an action which involves climbing a tree in order to look for insects is a reasonable complex event, but an action which involves climbing a tree in order to look at the moon is not.

The literature on serial verbs and verbal complexes is vast and there are several works which attempt a formal characterization of the properties of serial verbs (e.g., Sebba, 1987; Durie, 1997; Osam, 2003). However, much more needs to be done for a precise characterization of both serial verbs and complex predicates that can lead to a comprehensive theory about differing types of verbal complexes. The work presented here is intended as a step towards a more precise characterization of the properties of complex predicates and light verbs. For a discussion of serial verb constructions which proposes a typology that is very much in line with the views presented here, see Ameka (2006).<sup>11</sup>

<sup>11</sup> Ameka’s characterization includes diachronic arguments in that serial verbs are known to give rise to prepositions, complementizers and aspectual markers (e.g., Lord, 1993). The claim in this paper is that light verbs are historically inert. A reviewer has suggested that the serial verb data invalidates the diachronic evidence adduced here for light verbs. However, note that while the evidence presented in this paper is drawn from historical sources, the evidence for the grammaticalization of serial verbs is based primarily on comparative synchronic data (since no or only fragmentary diachronic evidence is available for most of the serializing languages that have been investigated) and an *expectation* of the kinds of historical processes that *should* apply.

## 2.4. Light verbs vs. compounds

The V–V constructions in (1) are also sometimes analyzed as *compound verbs* (e.g., [Hook, 1974](#)) thus shifting the burden of verbal composition into the lexicon. However, these V–V constructions exhibit none of the properties usually associated with compounds. If the V–V constructions were compounded, one would expect a tight morphological relationship between the light verb and the main verb. However, this is not so: as (17) shows, the light verb can be separated from the main verb via phenomena such as topicalization and the two verbs can be separated by a focus clitic.

- (17) a. batʃʃa (so)<sub>ω</sub> (ga-ya)<sub>ω</sub>  
 child.M.Nom sleep go-Perf.M.Sg  
 ‘The child has gone to sleep.’ (Urdu)
- b. so to batʃʃa ga-ya  
 sleep Top child.M.Nom go-Perf.M.Sg  
 ‘The child has gone to sleep.’ (Urdu)
- c. batʃʃa so hi ga-ya  
 child.M.Nom sleep even go-Perf.M.Sg  
 ‘The child has even gone to sleep.’ (Urdu)

One would also expect the main verb to be in its stem form (as is the case in nominal compounding, cf. the Level-Ordering Hypothesis ([Kiparsky, 1982a](#))), however, as the examples from Bengali in (2) showed, this is also not the case: the main verb carries perfective or gerundive morphology.<sup>12</sup>

We conclude that the two verbs must therefore be considered to be separate syntactic entities and that the term *compound verb* is a misnomer.

## 2.5. Summary

Light verbs form a distinct syntactic category with distinct distributional properties. We follow Butt and Geuder (2001b) and view light verbs as a subtype of V, rather than as a subtype of an auxiliary (e.g., T or I). That is, we see light verbs as primarily lexical in nature (rather than functional). However, while light verbs are in the lexical domain, they are not main verbs (or serializing verbs), but instead interact with the semantics of a main verb by modifying the main event ([section 4](#)).

## 3. Historical evidence: pertinacity vs. transience

### 3.1. Semantic bleaching

The crosslinguistic existence of light verbs is generally attributed to a process of *semantic bleaching* (e.g., [Hopper and Traugott, 1993](#); [Hook, 1991](#)), which is assumed to be connected to a historical process of grammaticalization, whereby content verbs proceed along the cline shown in (18) on their way to a reanalysis as function words/morphology. The term *vector verb* is used in one tradition of literature on South Asian languages to designate light verbs.<sup>13</sup>

- (18) full verb > (vector verb) > auxiliary > clitic > affix  
 (Grammaticalization Cline, [Hopper and Traugott, 1993](#):108)

Clines as in (18) are generally posited as part of the Grammaticalization approach ([Lehmann, 1982, 1985](#)).<sup>14</sup> The cline makes a historical prediction that light verbs are derived historically from main verbs and that from there they develop further into auxiliaries and finally into verbal morphology.

In particular, [Hook \(1991, 1993\)](#) analyzes the existence of light verbs in South Asian languages in terms of a historical process of aspectogenesis. The light verbs are seen as a class of perfective aspectual auxiliaries which have developed from main verbs via grammaticalization. In a larger and later survey, [Hook \(2001\)](#) is more careful with the claims

<sup>12</sup> The main verb in Urdu/Hindi did as well until it was lost in Old Urdu/Hindi (see [section 3.2.2](#)).

<sup>13</sup> The term *vector verb* is originally due to [Pray \(1970\)](#).

<sup>14</sup> It is noteworthy that this approach is the only analysis within historical linguistics we are currently aware of which makes a clear distinction between light verbs and auxiliaries in that light (vector) verbs are seen as an optional precursor to auxiliary formation.

surrounding perfectivity, noting that the notion “perfective” is not well defined and that light verbs show functions which go beyond the purely aspectual.

In our view, the aspectogenesis hypothesis runs counter to the observable facts in that light verbs do not behave like auxiliaries either synchronically or diachronically. The next section provides evidence for the historical pertinacity of light verbs in Indo-Aryan. This pertinacity is then contrasted with the comparative mutability of auxiliaries in [section 3.3](#).

### 3.2. Diachronic pertinacity of light verbs

It is generally agreed (e.g., [Hook, 1991](#); [Tikkanen, 1987](#); [Hendriksen, 1944](#); [Chatterji, 1926](#)) that the ancestral construction of the modern V–V complex predicate is the Sanskrit “gerund” or “absolutive” in *-tvā(ya)*, or *-ya/yā*. These suffixes served as derivational morphemes which resulted in an indeclinable participle (e.g., [Whitney, 1889:345–360](#)).<sup>15</sup> The use of the *tvā* participles was manifold and varied. [Tikkanen \(1987:7\)](#) uses the following constructed example to illustrate the various possible translations found in the literature.

- (19) indram ārabh-**ya** car-a  
 Indra.ACC grasp-GD go-IMP.2SG  
 a. ‘Having taken hold of Indra, move!’  
 b. ‘Take hold of Indra and move!’  
 c. ‘Take hold of Indra before moving.’  
 (d. ‘Move by taking hold of Indra!’)  
 (e. ‘Go to take hold of Indra!’)  
 f. ‘Keep yourself to Indra!’ (Sanskrit)

Note that while the translational possibilities in *a–d* are truth-conditionally equivalent, the possibilities in *a* and *f* can be contrasted with one another. These two translational possibilities are reminiscent of the pattern found in the modern languages: the usage in *f* is comparable to the modern V–V complex predicate, while *a* is comparable to the biclausal adverbial reading found with the same V–V string (cf. (13)).

#### 3.2.1. Clause chaining

The Sanskrit *tvā* is also sometimes referred to as a *conjunctive participle* (CP), presumably to reflect the clause chaining effect under the adverbial reading. The sequential linking of clauses via the *tvā* participle was very common in Sanskrit. An example is shown in (20).

- (20) ... [up-**tvā** keśaśmaśrūṇi] [nakhāni ni-kr-**tya**] [ājyā{-}] bhyaj-**ya** ...  
 shear-Gd hair-of-head-and-beard.Nom.PI nail.Nom.PI in-do-Gd clarified-butter.Nom.PI smear-Gd  
 ‘... having shaved his hair and beard, cut his nails, smeared his eyes and anointed his body, ...’ (Sanskrit)  
 (Jaiminīyabrāhmaṇa, from [Tikkanen \(1987:187\)](#))

This “clause chaining” phenomenon is still common in the modern languages and employs the same (-e) that is used in complex predicates in Bengali ((13)), and which is descended from the original participial affix ([section 3.2.2](#)). [Chatterji \(1926:1011\)](#) cites an example with as many as 15 subclauses for Bengali. Just three of these 15 are shown in (21).

- (21) b<sup>h</sup>ore ut<sup>h</sup>-e, ſtefane pōuc<sup>h</sup>-e, ſikiṭ kor-e, ... col-e jeo  
 early rise-Gd station reach-Gd ticket buy-Gd walk-Gd go.Imp  
 ‘get up early, get to the station, buy a ticket ... and leave’  
 [‘having gotten up early, having gone to the station, having bought a ticket ... , leave’]  
 (Bengali, [Chatterji, 1926:1011](#))

The final -e affix on ‘walk’ does not signify yet another modificatory adverbial, but marks the main verb of the complex predicate ‘walk go’(=‘leave’). This is the main predication of the sentence. Chatterji’s example nicely illustrates the

<sup>15</sup> This participle replaced a variant with *-tvī* and is thought to be a frozen form of an old instrumental verbal noun in *tu* ([Speijer, 1886, § 379](#), [Chatterji, 1926:1008](#)).

possibilities for interpretative ambiguity between a verb embedded in a 'having'-type clause and the main verb of a complex predicate. This interpretive ambiguity has survived intact in the modern language, but can be resolved prosodically, as was seen in [section 2.2](#).

### 3.2.2. The ancestral morphology

The ancestry of the modern forms in both the clause chaining (adverbial) construction as well as the V–V complex predicates has been reconstructed in some detail by [Chatterji \(1926:1006–1011\)](#), [Kellogg \(1893:341\)](#) and [Beames \(1872–1879, vol. 3, § 73\)](#). We list those forms which pertain directly to Bengali and Urdu/Hindi in (22).<sup>16</sup>

(22)	Sanskrit	Early MIA (Pāli)	Late MIA	Old		Modern Coll.	
				Bengali	Hindi	Bengali	Hindi
	-tvā	-tvā	-ttā, -tā, -cca, -ccā				
	-yā/ya	-ya, -iya	-ia, -i	iā, -iyā	-i	-e	∅

Sanskrit shows evidence for a distributional difference between *tvā*, which was commonly dubbed the Absolute I, and *ya*, the Absolute II. The *ya* was generally used in conjunction with preverbs ([Whitney, 1889, § 989](#), also see [Macdonell, 1917](#) for Vedic). Given that the modern light verbs contribute to the predication of the event in much the same elusive way as preverbs did in Sanskrit, this may be a significant factor in the spread of V–V constructions ([section 5](#)). It might also furnish an explanation for why it was the *ya* and not the *tvā* form which was the ancestor of the modern forms. On the other hand, [Chatterji \(1926:1008\)](#) and [Whitney \(1889, § 990\)](#) assert that a number of exceptions to this complementary distribution can already be found in later Sanskrit and that it appears to have been completely ignored in Middle Indo-Aryan.

The *-ya* form in early Middle Indo-Aryan already showed signs of assimilation ([Chatterji, 1926](#)), with the *-i-* of the *-iya* form coming from the preceding material. This short vowel can be found in the writings of medieval (Old) Hindi, but has been altogether lost in modern Urdu/Hindi. For Bengali, on the other hand, [Chatterji \(1926:1010\)](#) assumes that the short vowel was strengthened by either a “definitive” form *ā*, or by the *ā* from an extended and related form of the participle: *iyāṅā*, which became *-iā*. The literary form of Bengali retains the Old Bengali *iā*, *-iyā*, the colloquial form exhibits the short vowel *-e*.

### 3.2.3. Ancestral usage

We thus see that the ancestral “absolute” or gerund in *-ya* has persisted in the modern languages and that an unbroken chain of developments can be reconstructed for the form of the New Indo-Aryan affixes. Beyond this morphological continuity, our contention is furthermore that the languages always allowed for the use of light verbs and that in every case where a light verb reading is probable, the light verb is in fact form-identical to a main verb in the language.

We begin with classical Sanskrit and Vedic as discussed by [Tikkanen \(1987\)](#), who notes that the constructions formed with *-tvā/ya* displayed some curious syntactic and semantic properties. Syntactically, there is evidence for a kind of “clause union” in the sense that negation and wh-operators appear to be able to take scope over all the verbs in the clause ([Tikkanen, 1987:16](#)). In addition, while the default interpretation of constructions involving the *tvā/ya* is the adverbial ‘having done X’ (cf. (20)), this default interpretation is often unsatisfactory [Tikkanen \(1987\)](#).

One example of an awkward interpretation under the default ‘having’ reading is the Vedic instantiation in (23) of the constructed Sanskrit example in (19). The translation shown here is from [Tikkanen](#), who bases himself on [Delbrück’s \(1888:406\)](#) German rendition: *die wir wandern, nachdem wir dich ergriffen, zum Stützpunkt genommen haben* ‘we who wander, after we grasped you, took you as our base’ or *die wir uns immer an dich halten* ‘those who always keep themselves to you’.

(23)	ime	ta	indra	te
	Pron.Dual	Dem.Pron.3.Sg	Indra.Voc.Sg	Pron.2.Sg.Gen
	vayam	puruṣṭuta		
	Pron.1.PI.Nom	much.praise.PP.Voc.Sg		

<sup>16</sup> Other forms include formations with *-tvī* or *-tu* and the later addition of a *-na*.

ye tvā-rabh-ya car-āmasi  
 Rel.Pron Pron.2.Sg.Acc-grasp-Gd go-Pres.1.Pl  
 ‘We here are yours, O ever-praised Indra, who wander about having taken hold of you/who constantly keep ourselves to you.’ (Vedic)  
 (R̥gveda I.57.4; Tikkanen, 1987:175)

Grasmann (1872:437) points out for *car* ‘go, move’ that in conjunction with a participle its meaning is weakened to such an extent that it only seems to express the duration of the action denoted by the participle. This is the case in (23), where any analysis in terms of an embedded adverbial yields a strange interpretation. On the other hand, if one analyzes the ‘go’ as a light verb and the ‘grasp’ as a main verb, the more plausible reading “we hold on to you (constantly)” (=‘we constantly keep ourselves to you’) becomes possible.

However, we need to admit at this point that the above pieces of data cannot be seen as unassailable proof of complex predication in Sanskrit. Rather, the irritation of scholars like Delbrück and the lengthy attempts to deal with the rather puzzling range of interpretations and structural properties associated with the *-tvā/ya* participle in both Sanskrit (see Tikkanen, 1987 for a summary) and Pāli (Hendriksen, 1944), must be taken as indications that the construction is not well understood.

This sense of profound puzzlement is a hallmark of complex predicates in many modern South Asian grammars as well (where they are variously treated as compound verbs, idioms, adverbials, etc.). For the modern languages, however, various linguistic tests are available (e.g., Aissen and Perlmutter, 1983; Neeleman, 1994; Mohanan, 1994; Butt, 1995; Alsina, 1996; Alsina et al., 1997) so that complex predicate constructions can be distinguished from other V–V combinations.

The Sanskrit and Vedic forms examined above are drawn from different time spans. There is no precise dating of the times when the different languages were spoken, but the oldest attested form is thought to go back to 1200 BCE. Vedic is generally dated until about 600 BCE. Epic and Classical Sanskrit fall into the time from 600 BCE to 200 CE. Together with Vedic, these are referred to as Old Indo-Aryan. Middle Indo-Aryan includes Pāli (mainly preserved in Buddhist texts), several Prākṛit languages (which include non-standard dialects of Sanskrit), Apabhramśa and inscriptions of the Emperor Aśoka (270–232 BCE). The Middle Indo-Aryan period stretches from about 200 BCE to 1100 CE. The languages of the period from then on are commonly referred to as New Indo-Aryan. As of 1100 CE distinct ancestors of the modern languages such as Old Hindi, Old Bengali or Old Marathi have emerged.

Complex or joint predication becomes much easier to identify in Middle Indo-Aryan (Hook 1993). The Pāli examples in (24) both involve the verb ‘give’ as a finite verb which combines with the participle of ‘make’. For both these sentences it would be strange to assume that the meaning should be rendered as: “having led her to the hermitage, having made a fire, he gave (it) (to her)”. Rather, the complex predicate benefactive reading in the glosses seems more appropriate.

- (24) a. ... assamapadam āne-tvā aggim **ka-tvā a-dāsi**  
 hermitage.Acc lead-Gd fire.Acc.Sg make-Gd Aug-give.Impf.3.Sg  
 ‘... brought her to his hermitage and made a fire for her’ (Pāli)  
 [‘having brought (her) to the hermitage, made a fire (for her)’]  
 Jātaka Tales I.296.10, Sri Lanka (Hendriksen, 1944:134)
- b. daruni āhari-tvā aggim **ka-tvā dassati**  
 sticks bring-Gd fire.Acc.Sg make-Gd give.Fut.3.Sg  
 ‘Bringing wood he’ll make a fire (benefactive use).’ (Pāli)  
 (Trenckner 1879:77, cited by Hook, 1993:97)

Hendriksen (1944) provides a detailed discussion of the interpretational difficulties associated with the gerund in Pāli, some of which echoes the discussion for Sanskrit found in Tikkanen (1987). Compare the Pāli to the modern Urdu in (25), where the finite light verb ‘give’ also allows for a benefactive interpretation.

- (25) ... apne g<sup>h</sup>ar le kar roṭi **bana d-i**  
 self.Obl house.M.Sg.Nom bring having bread.F.Sg.Nom make give-Perf.F.Sg  
 ‘... having brought her home, (he/she) made food (for her/him)’ (Urdu)

In New Indo-Aryan complex predicates can be identified even more readily than in Middle Indo-Aryan. We show some examples for Old Bengali in (26) from the Caryapad (950–1550 CE), which consists of 46 complete songs and one incomplete song of 6 lines by 24 different poets. Here the finite verbs ‘take’ ((26a)) and ‘give’ ((26b)) cannot be taken in

their main verb sense, but must rather be interpreted as light verbs which signal completion, much as is done in the modern language. Also note that ‘give’ has a benefactive sense, while ‘take’ indicates a self-directed action, a pattern which is typical of light verbs crosslinguistically and which is found in several of the modern South Asian languages. Finally, note also the separability of the main and light verb in (26b). As was seen in [section 2.4](#), this property remains a feature in the modern languages.

- (26) a. cauṣat<sup>hi</sup> koṭ<sup>ha</sup> **guṇ-iā** **lehu**  
 sixty-four rooms count-Gd take  
 ‘count sixty-four rooms (for yourself)’  
 (Caryapad 12, [Mojunder 1973:248](#)) (Old Bengali)
- b. bājule **di-la** moha-kak<sup>hu</sup> **b<sup>h</sup>aṇ-iā**  
 Bajula.Obl give-Past.3.Sg rooms of illusion count-Gd  
 ‘Bajula counted the rooms of illusion (for his disciple).’  
 (Caryapad 35, [Mojunder 1973:248](#)) (Old Bengali)

We close this section with a few examples of Old Hindi which date to the middle ages (Tulsī Dās’ *Rāmāyaṇa* of about 1575 CE contained V–V formations to such an extent that a dissertation on the subject was warranted, see [Meiṣner, 1964](#)). Example (26) and the examples in (27) of Indrajit of Orchā’s Braj Bhāṣā prose (ca. 1600 CE) are very close to the modern forms. [McGregor \(1968:209–213\)](#) explicitly notes that they were used much like complex predicates in modern Hindi.

- (27) a. ... **cor-i** **le-tu** hai  
 steal-Gd take-Impf be.Pres.3.Sg  
 ‘... (he) steals’ (Old Hindi)
- b. **kāḍh-i** **le-i**  
 pull out-Gd take-Perf  
 ‘(he) pulled out (with effort)’ (Old Hindi)
- c. samudrahiṃ **nāṣ-i** **jā-ta** haim  
 ocean.Obj cross-Gd go-Impf be.Pres.1/3.PI  
 ‘(They/We) cross oceans (completely).’ (Old Hindi)

In summary, this section has shown that the use of light verbs can be traced to a very early stage of the language and that at every stage, the light verb was form-identical to a main verb. This suggests to us that rather than engaging in the type of reanalysis which leads to the formation of auxiliaries or modals, light verbs are diachronically inert and have been an integral part of the language at every stage.

### 3.3. Auxiliaries and diachronic change

Given that the availability and use of light verbs in V–V complex predicates can be traced back over the millenia, one might speculate that the diachronic pattern presented above could be due to a general structural inertness in Indo-Aryan. However, while the V–V complex predicate construction appears to have been available in basically the same structural form as far back as the written record extends, other structural properties in Indo-Aryan have undergone a significant amount of change. A relevant contrast is furnished by the tense/aspect system. The heavily inflectional system of Old Indo-Aryan has been replaced by a mixed periphrastic and inflectional system in New Indo-Aryan with most of the old inflectional morphology having been lost over the millenia.

In this section, we look at two verbs, ‘be’ and ‘go’, which are also used as light verbs in V–V complex predicates. We show that while these verbs were used as main verbs as well as light verbs, they also simultaneously gave rise to new periphrastic auxiliary constructions and were then further reanalyzed as inflectional tense/aspect morphology. Despite the rise of auxiliaries, the main and light verb uses were not lost, as would be predicted by the grammaticalization cline in (18).

Otherwise, the type of historical development found with these verbs is nothing special in crosslinguistic terms: the path of change is a type of reanalysis that has already been well established crosslinguistically (e.g., [Harris and Ramat, 1987](#); [Harris and Campbell, 1995](#)).

### 3.3.1. The verb 'be'

The following examples show 'be' in Bengali and Urdu in a main verb ((28)), a light verb ((29)), an auxiliary use ((30)) and examples where the original 'be' is realized as a clitic ((31a)) and a morphological affix ((31b)).

- (28) a. **Main Verb**  
ami b<sup>h</sup>alo aʃ<sup>h</sup>i  
I.Nom well be.Pres.1  
'I am well.' (Bengali)
- b. **Main Verb**  
mē ʃ<sup>h</sup>ik hū  
I.Nom well be.Pres.1.Sg  
'I am well.' (Urdu)
- (29) a. **Light Verb**  
amar mone aʃ<sup>h</sup>e  
I.Gen mind.Loc be.Pres.3  
'I remember.' (Bengali)
- b. **Light Verb**  
moj<sup>h</sup>ə yad hε  
I.Obl.Dat memory.M.Sg.Nom be.Pres.3.Sg  
'I remember.' (Urdu)
- (30) a. **Auxiliary**  
nadya=ko xat mil-e hε  
Nadya.F=Dat letter.M.Nom receive-Perf.M.PI be.Pres.3.Sg  
'Nadya has received letters.' (Urdu)
- b. **Auxiliary**  
nadya=ko xat mil-t-e hε̃  
Nadya.F=Dat letter.M.Nom receive-Impf-M.PI be.Pres-M.PI  
'Nadya used to receive letters.' (Urdu)
- (31) a. **Clitic**  
ram ʃ<sup>h</sup>i pe-(y)e=ʃ<sup>h</sup>-ilo  
Ram.Nom letter.Nom receive-Perf=be-Past.3  
'Ram had received letters.' (Bengali)
- b. **Morpheme**  
ram ʃ<sup>h</sup>i pa-ʃ<sup>h</sup>-ilo  
Ram.Nom letter.Nom receive-be-Past.3  
'Ram was receiving letters.' (Bengali)

In the Urdu examples, the 'be' is immediately recognizable because it is instantiated as a lexical item that is form-identical with the main verb use of 'be'. In Bengali, on the other hand, it is not immediately obvious that the forms glossed as 'be' in (31) are related to the 'be' of (28) and (29). This is because the verb aʃ<sup>h</sup> 'be' has given rise to an inflectional morpheme via a stage in which 'be' formed a periphrastic auxiliary with a former participle (Chatterji, 1926; Lahiri, 2000).

Bengali is particularly interesting since there is not only a 'be' auxiliary and an inflectional morpheme that is derived from a 'be' auxiliary, but also a clitic version. The perfect in Bengali is expressed with the clitic version of 'be' (31a), the progressive with a morpheme derived from 'be' (31b). As discussed in the remainder of this section, the evidence for this analysis is intricate, but solid.

We begin with a close investigation of the surface forms of the progressive and the perfect in (31). This reveals several differences. For one, the root vowel changes in the perfect (the perfect of *pa* 'receive' in (30a) is *pe(y)eʃ<sup>h</sup>*), but not in the progressive. Additionally, the progressive suffix has two allomorphs: /ʃ<sup>h</sup>/ and /ʃ<sup>h</sup>ʃ<sup>h</sup>/. The choice of allomorph depends on the shape of the verb root. For instance, in the CVC- verb *par* 'to pick from a higher level, to be able to', the /ʃ<sup>h</sup>/ does not geminate, while there is gemination with the CV- verb *pa* (*par-ʃ<sup>h</sup>-e* vs. *pa-ʃ<sup>h</sup>ʃ<sup>h</sup>-e*).

Another difference is brought out by the distribution of the emphatic clitic *o* 'also' (Fitzpatrick-Cole, 1996; Lahiri and Fitzpatrick-Cole, 1999). As shown in (32), this clitic can be introduced between the stem and the -ʃ<sup>h</sup> in the perfect, but not in the progressive. The suffix /-e/ in the perfect is a true suffix. This is the morphophonological remainder of the Old

Indo-Aryan past participle *-ta* (Chatterji, 1926:660). The clitic *o* cannot intrude between a stem and an affix. It can, however, intrude between the stem and the /tʃʰ/ used in the perfect. This indicates that the 'be' in the perfect has the status of a cliticized auxiliary. In (30b), on the other hand, the 'be' has been reanalyzed as a piece of verbal inflection which signifies the progressive.

(32)	Perfect	Progressive
pa	pe-e-tʃʰ-e=o	pa-tʃʰ-e=o
'receive'	pe-e=o-tʃʰ-e *pe=o-e-tʃʰ-e	*pa=o-tʃʰ-e
par	per-e-tʃʰ-e=o	par-tʃʰ-e=o
'be able'	per-e=o-tʃʰ-e *per=o-e-tʃʰ-e	*par=o-tʃʰ-e

We posit the reanalysis in (33) for the progressive and the perfect. Since the progressive behaves like a suffix rather than a clitic, the root + suffix combination plus the person suffixes must be analyzed as a single prosodic word in modern Bengali. The combination of 'be' plus a perfect stem, on the other hand, is analyzed as two separate prosodic words, which are then realized as one via the incorporation of the clitic (e.g., Zec and Inkelas, 1990; Bayer and Lahiri, 1990; Lahiri, 2000).

- (33) **Progressive:**  
 [Root + Suffix<sub>PART</sub>]<sub>ω</sub> + [be + Suffix<sub>PERS/NUM</sub>]<sub>ω</sub>  
 > [[Root + Suffix<sub>PART</sub>] + [be + Suffix<sub>PERS/NUM</sub>CLITIC]]<sub>ω</sub>  
 > [[Root + Suffix<sub>PART</sub> + be<sub>AFF</sub>] + Suffix<sub>PERS/NUM</sub>]<sub>ω</sub>  
 > [[Root + be<sub>AFF</sub>] + Suffix<sub>PERS</sub>]<sub>ω</sub>
- Perfect:**  
 [Root + Suffix<sub>PERF</sub>]<sub>ω</sub> + [be + Suffix<sub>PERS/NUM</sub>]<sub>ω</sub>  
 > [[Root + Suffix<sub>PERF</sub>] + [be + Suffix<sub>PERS</sub>CLITIC]]<sub>ω</sub>

As shown in (33), the progressive was originally formed via a periphrastic construction which consisted of an auxiliary and a participle (attested in Middle Bengali). The participle was formed via a suffix *i* whose precise origin is unclear.<sup>17</sup> The perfect was also formed via a periphrastic construction, but with /-iya/, the descendent of the Old Indo-Aryan past participle *-ta*. This /-iya/ coalesced to give modern Bengali *-e*. Both participles took a form of 'be' which inflected for person and number. Number marking was lost in the course of historical change for unrelated reasons and so modern Bengali only displays person agreement. The abstract schema in (33) can be rendered more transparent by considering an actual example with the root *pa* 'receive'.

- (34) **Progressive:**  
 [pa + i]<sub>ω</sub> + [atʃʰ/ + Suffix<sub>PERS/NUM</sub>]<sub>ω</sub>  
 > [[pa + i] + [tʃʰ + Suffix<sub>PERS/NUM</sub>CLITIC]]<sub>ω</sub>  
 > [[pa + i + tʃʰ<sub>AFF</sub>] + Suffix<sub>PERS/NUM</sub>]<sub>ω</sub>  
 > [[pa + tʃʰ<sub>AFF</sub>] + e]<sub>ω</sub>
- Perfect:**  
 [pa + iya]<sub>ω</sub> + [atʃʰ/ + Suffix<sub>PERS/NUM</sub>]<sub>ω</sub>  
 > [pa + e]<sub>ω</sub> + [atʃʰ/ + Suffix<sub>PERS/NUM</sub>]<sub>ω</sub>  
 > [[pa + e] + [tʃʰ + e]CLITIC]<sub>ω</sub>

The observant reader will note that the forms in (34) do not quite reflect the actual forms found in (32). In particular, the reason for the gemination in the progressive is not clear from what has been said so far. Under the assumption of Lexical Phonology (Kiparsky, 1982b, 1985), a summary of the analysis needed to account for the modern surface forms is shown in (35). A full discussion can be found in Lahiri (2000).

<sup>17</sup> In Literary Bengali, this participle is formed with *ite*, a descendent of the Sanskrit present participle affix *-ant*.

	Perfect		Progressive	
Root	pa	par	pa	par
Deriv. Morph.	pa+iya	par+iya	pa+i	par+i
Coalescence	pa+e	par+e		
V Harmony	pe-e	per-e		
Aux. Clitic	(pe-e tʃ <sup>h</sup> e) <sub>ω</sub> (per-e tʃ <sup>h</sup> e) <sub>ω</sub>		(pa+i tʃ <sup>h</sup> e) <sub>ω</sub> (par+i tʃ <sup>h</sup> e) <sub>ω</sub>	
Reanalysis			[pa+i+tʃ <sup>h</sup> ]+e [par+i+tʃ <sup>h</sup> ]+e	
Med. <i>i</i> Loss			[pa+tʃ <sup>h</sup> ]+e [par+tʃ <sup>h</sup> ]+e	
Gemination			pa+tʃ <sup>h</sup> tʃ <sup>h</sup> –	
Infl. Morph.			patʃ <sup>h</sup> tʃ <sup>h</sup> +e partʃ <sup>h</sup> +e	

As can be seen, participles are first formed via derivational morphology. /a/ goes to /e/ in Bengali via vowel harmony in the synchronic grammar. This accounts for the surface form of the Bengali perfect. The auxiliary cliticizes to the perfect and nothing more need to be said about that. In the progressive, however, the clitic (without the person/number ending) was reanalyzed as a tense/aspect affix. This was followed by the loss of the original participle marker *i*. The loss of this medial *i* freed a mora in the CV roots, which led to gemination. In the CVC roots no mora is free, so no gemination occurs.<sup>18</sup>

Thus, Bengali shows evidence of progressive reanalysis with respect to the verb ‘be’, which can be used as a main verb, a light verb and as an auxiliary. But there is a crucial difference between the light verb use and the auxiliary. While the light verb remains form-identical to the main verb, the auxiliary has cliticized in the perfect and has been reanalyzed as a piece of inflectional morphology in the progressive. This reanalysis has taken place in the space of a few centuries while the light verbs have remained form-identical to the main verb for millenia.

### 3.3.2. The verb ‘go’

In this section, we present another case in which reanalysis from main verb to auxiliary to inflectional morpheme took place in the space of a few centuries. Just as with ‘be’, the verb ‘go’ can be used as a main verb (36) and a light verb (37) in Bengali and Urdu. Additionally, it is used as a passive auxiliary in Urdu (38) and gave rise to today’s future morphology -g- (39).

- (36) a. **Main Verb**  
ami ge-l-am  
I.Nom go-Past-1  
‘I went.’ (Bengali)
- b. **Main Verb**  
mē ga-yi  
I.Nom go-Perf.F.Sg  
‘I went.’ (Urdu)
- (37) a. **Light Verb**  
batʃʃa por-e gæ-tʃ<sup>h</sup>-e  
child.M.Nom fall-Gd go-Perf-Sg  
‘The child fell (down)’ (Bengali)
- b. **Light Verb**  
batʃʃa gir ga-ya  
child.M.Nom fall go-Perf.M.Sg  
‘The child fell (down).’ (Urdu)

<sup>18</sup> Note that there is further evidence that the synchronic progressive marker in modern Bengali is the geminated form /-tʃʃ<sup>h</sup>/ (Lahiri, 2000), and that the CVC roots are synchronically formed by degemination.

- (38) **Auxiliary**  
 tʃor (pʊlɪs=se) pɑkr-a gɑ-ya  
 thief.M.Sg.Nom police=Inst catch-Perf.M.Sg go-Perf.M.Sg  
 ‘The thief was caught by the police.’ (Urdu)

- (39) **Morpheme**  
 pʊlɪs tʃor=ko pɑkr-e-g-i  
 police.F.Sg.Nom thief.M.Sg=Acc catch-3.Sg-Fut-F.Sg  
 ‘The police will catch the thief.’ (Urdu)

The grammars of the last century agree that the verb ‘go’ gave rise to the Urdu future morpheme *-g-*. The inflectional s-future of Old Indo-Aryan was lost and replaced by auxiliaries or alternative morphemes in Urdu/Hindi and Bengali, as well as in many other NIA languages.<sup>19</sup> The consensus in the available literature is that the future *-g-* morpheme is derived from a Sanskrit participle of the verb *gā* ‘go’ (Kellogg, 1893:231; Beg, 1988:191; McGregor, 1968). The gender and number agreement morphology (*a/i/e*) exhibited by the future is regular synchronically in that exactly this agreement morphology is also found on the perfect, imperfect and progressive forms, all descended from participles. The appearance of this morphology follows unproblematically if the *-g-* is indeed associated with an old participle of ‘go’.

The person/number inflection of the future paradigm is very likely a continuation of old present morphology, which could also be used as an immediate future or with subjunctive force (McGregor, 1968:176). We thus posit the stages of reanalysis shown in (40) for the development of the modern Urdu future.

- (40) **Future:**  
 [Stem + Suffix<sub>PERS/NUM</sub>]<sub>ω</sub> [go + Suffix]<sub>ω</sub>  
 > [[Stem + Suffix<sub>PERS/NUM</sub>] + [go + Suffix]<sub>CLITIC</sub>]<sub>ω</sub>  
 > [Stem + Suffix<sub>PERS/NUM</sub> + go + Suffix]<sub>ω</sub>

Evidence for the clitic stage is provided by Kellogg (1893, § 399), who notes that in the Hindi translation of the play *Shakuntala* (originally written in Sanskrit by Kalidasa between 100 BCE and 400–500 CE) several examples can be found in which the emphatic particle *hi* intrudes between the *-g-* suffix and the rest of the verb, as shown in (41).

- (41) a. kah-ũ=hi=ga  
 say-1.Sg-Emph-Fut.M.Sg  
 ‘I will say (it), of course.’ (Hindi, from Kellogg, 1893, § 399)
- b. man-e=hĩ=gi  
 heed-3.Sg-Emph-Fut.F.Sg  
 ‘She will (have to) see reason.’ (Hindi, from Kellogg, 1893, § 399)

The emphatic clitic *hi* in Urdu/Hindi is comparable to the emphatic clitic *o* discussed in connection with the Bengali progressive in section 3.3.1. As in Bengali, this clitic attaches to prosodic words. The existence of examples as in (41) supports the hypothesis that the modern future morphology arose through the reanalysis of an originally periphrastic construction involving a participle of ‘go’ via a stage of cliticization.

Kellogg (1893, § 399) notes that the development of the *-g-* future is of relatively recent origin: “the addition of this *gā* (from Sk. *gatah* ‘going’) to form from the contingent an absolute future appears to have been one of the latest developments of the language.” This confirms the central point of the paper: the formation of auxiliaries from main verbs generally feeds into a rapid historical process of reanalysis by which the resulting inflectional morphology loses its original form-identity to the main verb. This stands in stark contrast to the comparative diachronic inertness or pertinacity of light verbs.

#### 4. Lexical semantic representation

As already noted above, one common analysis of light verbs is that they are the result of the *semantic bleaching* of a main verb. That is, one assumes a lexical entry for a verb like ‘take’ which has a full-blown argument structure and event

<sup>19</sup> The modern Bengali future morphology derives from the Old Indo-Aryan participle *-tavya* (Chatterji, 1926:676).

semantics. From this full version, a light version is derived via a historical process. This light verb retains only some of the semantic content of the full verb, the other bits having been “bleached” away. From there, further bleaching results in an auxiliary which can then be reanalyzed further as a piece of inflectional morphology.

However, this proposed path of change does not do justice to the data adduced in this paper. In the next section, we therefore propose that light verbs be viewed in a different light. Namely, we propose a model which connects the light verb to the main verb version via one common underlying lexical entry.

#### 4.1. A single underlying entry for light and main verbs

Light verbs are unlike main verbs in that they are dependent on another predicative element. That is, they modulate or further structure a given predication, but do not supply their own main predication. The modulation can be in terms of providing more information about the typical parts of an event: who did the causation, what the result was, whether the event was bounded or whether it was benefactive, sudden, agentive/volitional, accidental, etc.

Light verbs are also unlike auxiliaries. The functional work done by auxiliaries is to situate the main event predication with respect to temporal or aspectual information, that is, with respect to speech and reference time (cf. [Kamp and Reyle, 1993](#)). This is not the functional work carried by light verbs, which modify the main predication in some (usually subtle) manner, but do not situate it with respect to speech and reference times.

In general, light verbs *add* information to an event predication. This also stands in contrast to phenomena such as passives, reflexives or middles, which operate on an existing argument structure in such a way as to provide a different perspective on the participants of the events, i.e., express *voice*.

Thus, light verbs must be differentiated from main verbs and auxiliaries in functional terms and this functional difference goes hand in hand with differences in synchronic morphosyntactic realization and diachronic developments, as seen above. However, as also emphasized, light verbs are always form-identical to a corresponding main verb.<sup>20</sup> We further note that the notion of *semantic bleaching* commonly associated with light verbs is intuitive to some degree as light verbs generally appear to contain a subset of the lexical semantic information of the corresponding main verb. However, while they clearly do not predicate as a main verb, they are also not necessarily confined to predicating just a proper subset of the main verb’s lexical semantic content.

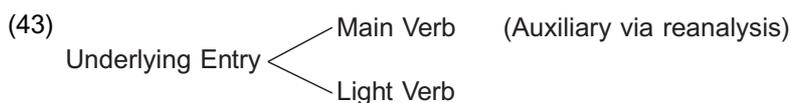
Consider the example in (42). This is a sentence uttered by a man in a context where the woman he has been promised to by his parents (arranged marriage) releases him from his obligation. As she explains that she realizes he does not love her (while she will always love him), he is moved almost to tears and he begs her to stop talking because otherwise he will have to cry.

- (42) mē ro d-ũ-g-a  
 I.Nom cry give-1.Sg-Fut-M.Sg  
 ‘I will cry.’ (from the movie *Kabhi Khushi Kabhi Gham*)

Our hero uses the light verb *de* ‘give’ to express his emotional state. Usually, this light verb carries a sense of benefaction with it. However, in this case there is no beneficiary of the crying act. Rather, the use of ‘give’ here appears to indicate control over the action (cf. [Butt and Geuder, 2001](#)) and is used to override the default assumption that crying is more of an involuntary, rather than a controlled action (our hero is manfully controlling his crying, but might still decide to cry anyway if the woman keeps talking).

It appears that light verbs predicate a subset of lexical semantic information associated with the main verb as well as being able to add shades of meaning not necessarily immediately associated with the lexical semantics of the main verb. We propose to explain this property of light verbs by positing just one underlying lexical entry for both main verbs and light verbs. This is illustrated schematically in (43). Note that we assume that the diachronic development of auxiliaries is based on the main verb version and not the light verb instantiation of the underlying lexical entry.

<sup>20</sup> Reviewers have expressed the worry that our discussion of light verbs is somewhat circular in the sense that a verb is only considered to be a light verb if it is form-identical to a main verb. We acknowledge the danger of circularity, but would like to emphasize that form-identity is not the only criterion for light verb status. Rather, it must have the characteristics of being in a syntactically monoclausal construction that corresponds to a biclausal argument structure and its functional/semantic work must be event modulation and structuring, but not temporal/aspectual event situation. It is true that our definition of a light verb is narrower than what one may encounter otherwise in the literature, but that is part of the point we are trying to make: the literature has to date not been very careful about distinguishing light verbs from auxiliaries, verbs in serial verb constructions or verbs found in lexical compounds. This leads not only to terminological, but also to analytical confusion, a situation we are attempting to remedy.



Exactly how the lexical semantic information in this single underlying entry should be coded is a tricky question, just as most research into the appropriate lexical semantic representations remains tricky. Most approaches to lexical semantics assume at least three dimensions (e.g., [Jackendoff, 1990](#); [Hale and Keyser, 1993](#); [Levin and Rappaport, 2005](#); [Ramchand, 2008](#)): (1) a thematic dimension encoding the semantic content of the verb; (2) some information about the subevents the predication includes (e.g., causation, result); and (3) information about the number and type of arguments that are licensed in the syntax (often referred to as thematic roles). The first two dimensions are generally seen to play a crucial role in determining the third one – this is the subject of Linking Theories (see [Butt, 2006](#) for an overview, [Randall, 2009](#) for a recent new proposal), which continue to be worked out in various ways.

In a landmark paper, [Dowty \(1991\)](#) introduced the notion of lexical entailments in order to allow for a more differentiated approach to argument realization and linking than the notion of thematic roles could provide for. We propose the following. Underlying lexical entries for verbs as in (43) consist of a set of lexical entailments in Dowty’s sense. That is, the lexical semantic content of verb is made up of a number of entailments detailing whether this is a volitional event, whether there is a beneficiary/recipient, whether there is a change of state, etc. In addition to this type of information, the underlying entry is also associated with information that is usually associated more with world knowledge. That is, what a falling event usually entails, namely that it is involitional, that it is sudden, that it is downward. Or what a giving event usually entails: that it is usually for the benefit of somebody (but not necessarily) and that it generally is a considered action (weighing the pros and cons). These further pieces of information may influence the argument structure in terms of what kinds of arguments are realized in the syntax. However, they are also likely to provide that extra bit of semantic predication which is the hallmark of light verbs, namely the information whether a given action is sudden, benefactive or the responsibility of the actor. And since it is information coming out of our world knowledge, it is also defeasible, i.e., not every predication with the light verb ‘give’, for example, will necessarily always have the same range of semantic connotations.

But information about the subevental make up of the verbal predication or the syntactic argument structure is not encoded at the common underlying lexical representation. This information only comes in once a predicational choice has to be made: whether the set of lexical entailments will flow into a main verb predication or a light verb. When the verb enters the syntax as a main verb, it predicates as a full event with a full range of argument participants. When the verb enters the syntax as a light verb, i.e., is slotted into the distributional space for light verbs in a language, then its lexical semantic content must combine with a full event predication. Thus, depending on the syntactic role of the verb, the information contained in it is deployed differently.

Note that this idea shares many features with the ideas articulated in [Marantz \(1997\)](#), [Borer \(2005\)](#) and [Ramchand \(2008\)](#) by which the syntactic type and predicational power of a lexical item is only determined once it has been inserted into syntax.

#### 4.2. Verbal *passepartouts*

Not all verbs allow for a light verb use. Indeed, the same set of verbs seems to be implicated over and over again crosslinguistically. Some common examples are the verbs for ‘come’, ‘go’, ‘take’, ‘give’, ‘hit’, ‘throw’, ‘rise’, ‘fall’ and ‘do/make’. We assume that it is the set of verbs which contain comparatively less lexically specified information that allow their wide range of application. One can think of this set of verbs as *passepartouts*<sup>21</sup>: their lexical semantic specifications are so general that they can be used in multitude of contexts, that is, they “fit” many constellations. Therefore, it is precisely these verbs which are able to enter into complex predicate constructions, as well as form the basis for auxiliary or modal uses. The idea is not that there is a main, fully specified use of these verbs from which the “lighter” versions are derived, but that there is one underlying entry which allows for both full and light uses. The full verb meanings are what give rise to auxiliary formation. Because the precise interpretation of these *passepartouts* must rely on further contextual information, they are ideal for basic predications in which a verb denoting physical movement comes to denote movement in a temporal space (e.g., situating an event with respect to speech and reference times). Other, lexically more specified verbs do not lend themselves to such predications (e.g., *go* vs. *perambulate*, *make* vs. *manufacture*, *have* vs. *own*, etc.). The

<sup>21</sup> One of the definitions of a *passepartout* is “Something that passes everywhere or provides a universal means of passage.” (Oxford English Dictionary).

very general nature of the semantics of this handful of verbs thus allows them to interact with the grammar in a variety of ways.

## 5. Discussion

At least three questions arise with respect to the material presented so far: (1) are the patterns observed for Indo-Aryan confined to that language family?; (2) can light verbs be reanalyzed further, even if they do not give rise to auxiliaries?; and (3) what role does frequency play? Each of these questions is addressed briefly in the next three sections.

### 5.1. Evidence from other language families

Our general conclusions are confirmed by evidence from other language families. Steever (1988) discusses both synchronic and diachronic uses of various multiverb constructions in Dravidian, one of the other major language families in South Asia. While Steever labels the constructions “serial verbs”, a subset of the multiverb constructions he discusses would seem to be more properly analyzed as instances of V–V complex predicates. A more in-depth study of the Dravidian structures is not viable within the scope of this paper, but it is interesting to note Steever’s final conclusion that “The SVF [Serial Verb Formation] is therefore an integral part of Proto-Dravidian Syntax.” (Steever, 1988:116). That is, instances of constructions that bear the hallmark of complex predication can also be found in the older stages of Dravidian. Indeed, Steever proposes that the complex predicate construction entered the Indo-Aryan languages via a Dravidian substrate. This idea has been proposed by other researchers as well, however, Hook (2001) argues against this view by showing that the areas with the highest synchronic occurrence of complex predicates “lies far from the present-day Dravidian heartland” (Hook, 2001:104).

Discussions in Brinton and Akimoto (1999) examine the origin and use of English N–V complex predicates like *take a bath* or *give a shiver* and thus further confirm the view that light verbs are a generally available linguistic device. On the basis of the available evidence, Traugott (1999) concludes that English N–V complex predicates have been native to the language at least since Old English (the furthest one can look back in terms of textual evidence). Though the middle ages saw a marked increase in their use, no other concomitant signs of grammaticalization can be identified: the light verbs do not diminish in form (e.g., become auxiliaries or affixes) and they do not lead to the development of functional categories (Traugott, 1999:257). Indeed, the overall system appears to be quite stable given that the number and type of light verbs involved remains relatively constant, and given that their ranking in terms of frequency of use remains stable with respect to each other (e.g., *give* consistently comes third in terms of frequency of use since early modern English). As is crosslinguistically typical, the light verbs are shown to contribute aspectual nuances as well as other types of semantic information. For example, the use of *give an answer* as opposed to the simple verb *answer* appears to signal that the action was done deliberately (Traugott, 1999:253).

A further piece of diachronic evidence which suggests that we are correct in positing such a tight connection between full and light verbs is Iglesias-Rábade’s (2001) study of the uses of *nimen/taken* ‘take’ in English. In the history of English, the verb *taken* gradually replaced the older form *nimen*. Iglesias-Rábade goes through the N–V constructions described in Brinton and Akimoto (1999) for Middle English and does not find a difference in usage between the two verbs, indicating that *taken* took over all the functions of *nimen*, including light verb functions. This is not surprising under our view, given that the light and full versions of a verb are tied to the same lexical entry. It is, however, surprising under a view which assumes that light verbs are derived from main verbs because this analysis allows for the possibility of *nimen* continuing on as a light verb side-by-side with a main verb *taken*.

### 5.2. Possible reanalysis of light verbs

Besides positing a tight connection between light verbs and main verbs via a single underlying entry, we have also made the case that auxiliaries are based on the main verb entry, not the light verb. Again, this appears to be a claim that is crosslinguistically valid. Bowerman (2008, 174) conducts a comprehensive survey of the relevant literature and crosslinguistic data and concludes that “I have found no evidence for light verbs being reanalyzed as auxiliaries.”

A very strong claim, and one espoused in previous versions of this paper, would be that light verbs are historical dead ends – they are diachronically inert and do not participate in historical change due to reanalysis. However, Bowerman (2008) claims that one does find further change crosslinguistically: the use of constructions can become more restricted, light verbs may meld onto their hosts (Bowerman calls this univerbation) and the word class of the light verb may change. Bowerman does not state whether the main verb use is lost when this happens, something that we would expect, given the tight connection between main and light verb usage that we posit.

Hook and Pardeshi (2006) address the question whether light verbs live an eternal life and provide examples from Tamil, a Dravidian language, that light verbs may undergo phonological change, may be “orphans” in the sense that no corresponding main verb exists and may be used to furnish a new conjunctive participle in the language. Their data is interesting and will need to be investigated further. Primarily what needs to be done is to establish that the items in question are indeed light verbs in the sense used here and not aspectual auxiliaries of some kind.

However, we do acknowledge that light verbs are not completely diachronically inert. Instances of lexicalization do occur – there is evidence for this in the history of Hindi/Urdu. For example, it is quite likely that the modern verb *lana* ‘bring’ is a lexicalization of a V–V complex predicate *le ana* ‘take come’. Lexicalization would explain the odd syntax associated with *lana*, as it acts like an intransitive verb in terms of subject case assignment. Furthermore, one finds verbs like *camakna* ‘to shine/sparkle’, where the *-k-* harks back to an old form of *kar* ‘do’. The modern verb *camakna* is thus the result of a lexicalization of an original light verb construction.

Despite such instances of historical change, we do want to claim that the historical evidence shows that light verbs are diachronically pertinacious in comparison with other parts of the language, such as auxiliaries. That is light verbs generally **resist** language change, while auxiliaries are amenable to them. This has to do with the semantic and functional role played by light verbs vs. auxiliaries: light verbs crucially contribute a subtle but substantial range of lexical semantic entailments that are closely linked to the main verb usage. In contrast, auxiliaries usually fulfill a semantically very specialized role. Just one or two aspects of the original main verb’s meaning are used and the reanalysis tends to be based on a specific use that then becomes reinterpreted, so that the auxiliary can easily develop away from the original main verb in both form and semantic content.

### 5.3. Light verbs and frequency of use

We now turn to the role of frequency. There are several pertinent observations that bear on the question of the diachronic pertinacity of light verbs that we argue for here.

One is that while light verbs are a frequent and immediately observable phenomenon in the modern languages, their use was undoubtedly rarer in the older languages. Hook (1991, 1993) and Hook and Pardeshi (2001) carefully document an increase in light verb usage as part of V–V complex predicates since the 1500s in the modern Indo-Aryan languages Hindi, Marwari and Marathi. Hook (2001) extends these studies to provide as complete a picture of the South Asian situation as the available data will allow.

Similarly, fewer gerunds/absolutives (i.e., the *tvā* construction which is the ancestor of today’s V–V complex predicate construction) can be found in Vedic than in Sanskrit. Whitney (1889:358) points out that while the later Nala and Bhagavad-Gītā have only one tenth as many verb forms as the R̥gveda, more than three times as many examples of the gerund can be found in these (also see Tikkanen, 1987, for more facts and figures along these lines). The reason for this could either be attributed to the differing genres (Vedic is represented mainly by hymns), or it could be the result of language change in that the gerund gained greater ground as time went on. The historical process could be seen as having begun within the oldest attested stage of the language, namely Vedic, and as continuing on down to the modern languages, where complex predicates based on the Old Indo-Aryan gerund are an integral part of the system.

Hook (1991, 1993, 2001) takes this increase in frequency to be indicative of grammaticalization in progress, by which main verbs are on their way to becoming aspectual auxiliaries via a light verb stage. That is, he sees the existence of light verbs as an instance of aspectogenesis, which arises out of the progressive grammaticalization and concomitant semantic bleaching of some main verbs. However, as we have shown, the light verb usage must be clearly delimited from that of auxiliaries. It can easily be shown that the light verbs do not situate events with respect to speech and reference times. Aspectogenesis is thus not what is going on in Indo-Aryan. However, the increase in frequency over the millenia and centuries is indisputable and we do not as yet have a good take on it. One possibility we are currently exploring is that the rise in frequency since the middle ages may be due to language contact with Turkic languages, which was brought in via the Moghuls. This would also explain the north-south divide which Hook and Pardeshi document – the Moghuls invaded from the north and extended their dominion into the south of India over the course of several centuries.

## 6. Conclusion

Grammars change and grammatical systems undergo change – this is an established truism. However, we have shown with respect to light verbs that certain linguistic structures are comparatively stable and unlikely to be subject to reanalysis or restructuring. We propose that there is a very tight connection between a light verb and its corresponding main verb, and that this connection differs markedly from the relationship an auxiliary bears to the main verb it is derived from. We depart from the received view that the existence of a light verb is due to a historical process of *semantic*

*bleaching* and instead propose that there is a single underlying lexical entry, which can account not only for the simultaneous synchronic uses of light and main verbs, but also for the available historical data.

We believe that light forms of verbs exist for those verbs which denote very basic relations and which can thus function as *passepourtout*: they are handy in a wide variety of structures, leading to a wide variety of uses. One of these uses is to add to the predication of a clause by interacting with the event semantics of the main verb, rather than heading an independent predication. The existence of the light verbs may be dormant in some languages and may only be activated when necessary (through other processes of historical change). Once in use, however, light verbs remain stable and are closely linked to the main verb use via a single underlying entry. Any change affecting the verb will affect both the main and the light verb. Auxiliaries, on the other hand, are derived from main verbs by various processes, but are not linked synchronically to the main verb by a single representation. The use of the light form of a verb may get lost, but unlike auxiliaries, they resist becoming reanalyzed as grammatical morphemes. The stability of the light verb is thus predicted by its semantic function.

## Acknowledgements

Versions of this paper have been presented at various places over the years. In each case, the audiences were extremely helpful in providing feedback. We would therefore like to thank the Department of Linguistics at Stanford University (March 1998), the members in the audience at DIGS in York (May 1998), the Workshop on New Reflections on Grammaticalization in Potsdam (June 1999), the members of the audience in Düsseldorf (December 2000), the Workshop on Grammaticalization vs. Lexicalization in Konstanz (February 2001), the Workshop on Morphosyntactic Change at Freudental (June 2001), the Light Verb Workshop in Harvard (November 2002), the members of the audience in Tromsø (May 2005) and the members of the SFB 471 at Konstanz. The writing of this paper was made possible by the financial support obtained from the DFG (Deutsche Forschungsgemeinschaft) via the SFB 471 at the University of Konstanz and via the Leibniz Prize awarded to Aditi Lahiri. Individuals whom we would like to thank for particularly detailed feedback are Ashwini Deo, Wilhelm Geuder, Peter Hook, Brian Joseph, Tracy Holloway King, Paul Kiparsky, Gillian Ramchand, Dieter Wunderlich, and several anonymous reviewers. Finally, we would like to thank Karin Schunk for her help with Vedic, Pāli and Sanskrit.

## References

- Abbi, A., 1991. Reduplication in South Asian Languages. An Areal, Topological and Historical Study. Allied, New Delhi.
- Abeillé, A., Godard, D., Sag, I.A., 1998. Two kinds of composition in French complex predicates. In: Hinrichs, E., Kathol, A., Nakazawa, T. (Eds.), *Complex Predicates in Nondérivational Syntax, Syntax and Semantics*, vol. 30. Academic Press, San Diego.
- Adger, D., 2003. *Core Syntax: A Minimalist Approach*. Oxford University Press, Oxford.
- Aissen, J., Perlmutter, D., 1983. Clause reduction in Spanish. In: Perlmutter, D. (Ed.), *Studies in Relational Grammar 1*. The University of Chicago Press, Chicago, pp. 360–403.
- Alsina, A., 1996. *The Role of Argument Structure in Grammar*. CSLI Publications, Stanford, CA.
- Alsina, A., Bresnan, J., Sells, P., 1997. Complex predicates: structure and theory. In: Alsina, A., Bresnan, J., Sells, P. (Eds.), *Complex Predicates*. CSLI Publications, Stanford, CA, pp. 1–12.
- Ameka, F.K., 2006. Ewe serial verb constructions in their grammatical context. In: Aikhenvald, A.Y., Dixon, R. (Eds.), *Serial Verb Constructions: A Cross-linguistic Typology*. Oxford University Press, Oxford, pp. 124–143.
- Baker, M., 1989. Object sharing and projection in serial verb constructions. *Linguistic Inquiry* 20, 513–553.
- Bayer, J., Lahiri, A., 1990. Bengali emphatic clitics in the lexicon–syntax interface. In: Dressler, W., Luschützky, H., Pfeiffer, O., Rennison, J. (Eds.), *Contemporary Morphology*. Mouton de Gruyter, Berlin, pp. 3–16.
- Beames, J., 1872–1879. *A Comparative Grammar of the Modern Aryan Languages of India*. Munshiram Manoharlal, Delhi, Republished 1966.
- Beg, M.K.A., 1988. *Urdu Grammar: History and Structure*. Bahri Publications, New Delhi.
- Borer, H., 2005. *Structuring Sense*. Oxford University Press, Oxford.
- Bowern, C., 2008. The diachrony of complex predicates. *Diachronica* 25, 161–185.
- Brinton, L.J., Akimoto, M. (Eds.), 1999. *Collocational and Idiomatic Aspects of Composite Predicates in the History of English*. John Benjamins, Amsterdam.
- Bruce, L., 1988. Serialization: from syntax to lexicon. *Studies in Language* 12, 19–49.
- Butt, M., 1995. *The Structure of Complex Predicates in Urdu*. CSLI Publications, Stanford, CA.
- Butt, M., 1997. Complex predicates in Urdu. In: Alsina, A., Bresnan, J., Sells, P. (Eds.), *Complex Predicates*. CSLI Publications, Stanford, pp. 107–149.
- Butt, M., 2003. The light verb jungle. In: Aygen, G., Bowern, C., Quinn, C. (Eds.), *Harvard Working Papers in Linguistics*, vol. 9. Papers from the GSAS/Dudley House Workshop on Light Verbs, pp. 1–49.
- Butt, M., 2006. *Theories of Case*. Cambridge University Press, Cambridge.
- Butt, M., 2010. The light verb jungle: still hacking away. In: Amberber, M., Harvey, M., Baker, B. (Eds.), *Complex Predicates in Cross-linguistic Perspective*. Cambridge University Press, Cambridge, pp. 48–78.
- Butt, M., Geuder, W., 2001. On the (semi)lexical status of light verbs. In: Corver, N., van Riemsdijk, H. (Eds.), *Semi-lexical Categories: On the Content of Function Words and the Function of Content Words*. Mouton de Gruyter, Berlin, pp. 323–370.

- Butt, M., Ramchand, G., 2005. Complex aspectual structure in Hindi/Urdu. In: Erteschik-Shir, N., Rappaport, T. (Eds.), *The Syntax of Aspect*. Oxford University Press, Oxford, pp. 117–153.
- Butt, M., Rizvi, J., 2010. Tense and aspect in Urdu. In: Cabredo-Hofherr, P., Laca, B. (Eds.), *Layers of Aspect*. CSLI Publications, Stanford, pp. 43–66.
- Bybee, J.L., Perkins, R.D., Pagliuca, W., 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. The University of Chicago Press, Chicago.
- Chatterji, S.K., 1926. *The Origin and Development of the Bengali Language*, vol. II, D, 1975 ed. Mehra, Rupa & Co, Calcutta.
- Choi, I., Wechsler, S., 2001. Mixed categories and argument transfer in the Korean light verb construction. In: *Proceedings of the HPSG-01*, Trondheim, Norway, August.
- Chomsky, N., 1957. *Syntactic Structures*. Mouton de Gruyter, The Hague.
- Crowley, T., 1987. Serial verbs in Paamese. *Studies in Language* 11, 35–84.
- Delbrück, B., 1888. *Altindische Syntax*. Verlag der Buchhandlung des Waisenhauses, Halle.
- Dowty, D., 1991. Thematic proto-roles and argument selection. *Language* 67, 547–619.
- Durie, M., 1997. Grammatical structures in verb serialization. In: Alsina, A., Bresnan, J., Sells, P. (Eds.), *Complex Predicates*. CSLI Publications, Stanford, CA, pp. 289–354.
- Fitzpatrick-Cole, J., 1994. *The Prosodic Domain Hierarchy in Reduplication*, Ph.D. thesis, Stanford University.
- Fitzpatrick-Cole, J., 1996. Reduplication meets the phonological phrase in Bengali. *The Linguistic Review* 13, 305–356.
- Grassmann, H., 1872. *Wörterbuch zum Rig-Veda*, 4th ed. Otto Harrassowitz, Wiesbaden, Reprinted in 1964.
- Hale, K., Keyser, J., 1993. Argument structure. In: Hale, K., Keyser, J. (Eds.), *The View from Building 20*. The MIT Press, Cambridge, MA, pp. 53–109.
- Harris, A.C., Campbell, L., 1995. *Historical Syntax in Cross-linguistic Perspective*. Cambridge University Press, Cambridge.
- Harris, M., Ramat, P. (Eds.), 1987. *Historical Development of Auxiliaries*. Mouton de Gruyter, Berlin.
- Hayes, B., Lahiri, A., 1991. Bengali intonational phonology. *Natural Language and Linguistic Theory* 9, 47–96.
- Hendriksen, H., 1944. *Syntax of the Infinite Verb Forms of Pali*. Einar Munksgaard, Copenhagen.
- Hook, P.E., 1974. *The Compound Verb in Hindi*. The University of Michigan, Center for South and Southeast Asian Studies.
- Hook, P.E., 1991. The emergence of perfective aspect in Indo-Aryan languages. In: Traugott, E., Heine, B. (Eds.), *Approaches to Grammaticalization*. Benjamins, Amsterdam, pp. 59–89.
- Hook, P.E., 1993. Aspectogenesis and the compound verb in Indo-Aryan. In: Verma, M. (Ed.), *Complex Predicates in South Asian Languages*. Manohar Publishers & Distributors, Delhi, pp. 97–113.
- Hook, P.E., 2001. Where do compound verbs come from? (And where are they going?). In: Bhaskararao, P., Subbarao, K. (Eds.), *The Yearbook of South Asian Languages and Linguistics*. Sage Publications, New Delhi, pp. 101–130.
- Hook, P.E., Pardeshi, P., 2001. Inflation in the Marathi Compound Verb. Ms. University of Kobe, Japan.
- Hook, P.E., Pardeshi, P., 2006. Are Vector Verbs Eternal? Intertheoretical Approaches to Complex Verb Constructions. In: Talk given at the Eleventh Biennial Symposium. Rice University.
- Hopper, P.J., Traugott, E.C., 1993. *Grammaticalization*. Cambridge University Press, Cambridge.
- Iglesias-Rábade, L., 2001. Composite predicates in Middle English with the verbs *nimen* and *taken*. *Studia Neophilologica* 73, 143–163.
- Jackendoff, R., 1990. *Semantic Structures*. The MIT Press, Cambridge, MA.
- Jespersen, O., 1965. *A Modern English Grammar on Historical Principles*, Part VI, Morphology. George Allen and Unwin Ltd., London.
- Julien, M., 2002. *Syntactic Heads and Word Formation: A Study of Verbal Inflection*. Oxford University Press, Oxford.
- Kamp, H., Reyle, U., 1993. *From Discourse to Logic*. Kluwer Academic Publishers, Dordrecht.
- Kellogg, S.H., 1893. *Grammar of the Hindi Language*, 2nd ed. Munshiram Manoharlal Publishers Pvt. Ltd., Delhi, reprinted 1990.
- Kiparsky, P., 1982a. Lexical morphology and phonology. In: Yang, I.S. (Ed.), *Linguistics in the Morning Calm*. Hanshin, Seoul. Linguistics Society of Korea, pp. 3–91.
- Kiparsky, P., 1982b. Lexical morphology and phonology. In: Korea, T.L.S. (Ed.), *Linguistics in the Morning Calm*. Hanshin Publishing Company, Seoul, pp. 3–91.
- Kiparsky, P., 1985. Some consequences of lexical phonology. *Phonology Yearbook* 2, 85–138.
- Lahiri, A., 2000. Hierarchical restructuring in the creation of verbal morphology in Bengali and Germanic: Evidence from phonology. In: Lahiri, A. (Ed.), *Analogy, Levelling, Markedness: Principles of Change in Phonology and Morphology*. Mouton de Gruyter, Berlin, pp. 71–123.
- Lahiri, A., Fitzpatrick-Cole, J., 1999. Emphatic clitics and focus intonation in Bengali. In: Kager, R., Zonneveld, W. (Eds.), *Phrasal Phonology*. Foris Publications, Dordrecht, pp. 119–144.
- Lehmann, C., 1982. Thoughts on Grammaticalization: A Programmatic Sketch, vol. 1. *Arbeiten des Kölner Universalien-Projekts* 48. Universität zu Köln.
- Lehmann, C., 1985. Grammaticalization: synchronic variation and diachronic change. *Lingua e Stile* 20, 303–318.
- Levin, B., Rappaport, M., 2005. *Argument Realization*. Cambridge University Press, Cambridge.
- Lord, C., 1993. Historical Change in Serial Verb Constructions. *Typological Studies in Language*, vol. 26. John Benjamins Publishing Company, Amsterdam.
- Macdonell, A.A., 1917. *A Vedic Reader for Students*. Motilal Banarsidass Publishers, Delhi, Reprinted 1995.
- Marantz, P.A., 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In: Siegel, L., Dimitriadis, A. (Eds.), *UPenn Working Papers in Linguistics*, vol. 4.2. University of Pennsylvania, pp. 201–225.
- Masica, C., 1976. *Defining a Linguistic Area: South Asia*. The University of Chicago Press, Chicago.
- McGregor, R., 1968. *The Language of Indrajit of Orchā*. Cambridge University Press, Cambridge.
- Meißner, H.K., 1964. Simplex und Verbalkompositum in Tulsī Dās' Rāāyaṇa: cale jānā – cali ānā – uṭhi dhānā, Ph.D. thesis. Phillip-Universität zu Marburg.
- Mohanan, T., 1994. *Argument Structure in Hindi*. CSLI Publications, Stanford, CA.
- Mojunder, A., 1973. *Bengali Language Historical Grammar*, Part II. K.L. Mukhopadhyay, Calcutta.
- Neeleman, A., 1994. *Complex Predicates*, Ph.D. thesis. OTS Dissertation Series, Utrecht.

- Osam, E., 2003. Introduction to the Structure of Akan: Its Verbal and Multiverbal System. Unpublished manuscript, University of Ghana.
- Pray, B., 1970. Topics in Hindi-Urdu Grammar. Center for South and Southeast Asia Studies, Berkeley, California.
- Ramchand, G., 1990. Complex predicate formation in Bengal. In: Halpern, A. (Ed.), Proceedings of the Ninth West Coast Conference on Formal Linguistics. pp. 443–458.
- Ramchand, G., 2008. Verb Meaning and the Lexicon: A First Phase Syntax. Cambridge University Press, Cambridge.
- Randall, J., 2009. Linking: The Geometry of Argument Structure. Springer, Berlin.
- Roberts, I., Roussou, A. Syntactic Change: A Minimalist Approach to Grammaticalisation. Cambridge University Press, Cambridge, in press.
- Rosen, S., 1989. Argument Structure and Complex Predicates, Ph.D. thesis. Brandeis University.
- Saksena, A., 1982. Topics in the Analysis of Causatives with an Account of Hindi Paradigms. University of California Press, Los Angeles.
- Sebba, M., 1987. The Syntax of Serial Verbs. John Benjamins, Amsterdam.
- Singh, M., 1994. Perfectivity, Definiteness, and Specificity: A Classification of Verbal Predicates Hindi, Ph.D. thesis. University of Texas, Austin.
- Speijer, J.S., 1886. Sanskrit Syntax. Motilal Banarsidas, Delhi, Republished 1973.
- Steever, S.B., 1988. The Serial Verb Formation in the Dravidian Languages. Motilal Banarsidass, Delhi.
- Tantos, A., 2008. Computing Events in Discourse: A Case Study Involving Light “have”, Ph.D. thesis. University of Konstanz.
- Tikkanen, B., 1987. The Sanskrit Gerund: A Synchronic, Diachronic and Typological Analysis. Finnish Oriental Society, Helsinki, *Studia Orientalia*.
- Traugott, E.C., 1999. A historical overview of complex predicates. In: Brinton, L.J., Akimoto, M. (Eds.), Collocational and Idiomatic Aspects of Composite Predicates in the History of English. John Benjamins, Amsterdam, pp. 239–260.
- Whitney, W.D., 1889. Sanskrit Grammar. Harvard University Press, Cambridge, Massachusetts., Ninth Issue of the Second Edition (1960).
- Zbavitel, D., 1970. Non-finite Verbal Forms in Bengali. Oriental Institute in the Publishing House of the Czechoslovak Academy of Sciences, Prague.
- Zec, D., Inkelas, S., 1990. Prosodically constrained syntax. In: Inkelas, S., Zec, D. (Eds.), The Phonology-Syntax Connection. CSLI Publications, Stanford, CA, pp. 365–378.