Pertinacity in Light Verbs Revisited

Miriam Butt and Aditi Lahiri

University of Konstanz and Oxford University

SALA-37
37th South Asian Languages Analysis Roundtable Workshop Complex Predicates
October 4–7, 2023
Butt and Lahiri (2013) look at V-V complex predicates in Urdu and Bangla and claim that light verbs do not arise as a result of historical change. This claim has been challenged by Slade (2013) and Ittzés (2022). However, a careful scrutiny of the arguments advanced in both Slade (2013) and Ittzés (2022) shows that they in fact provide more evidence for Butt&Lahiri’s original claim.

- Slade with respect to V-V aspectual complex predicates.
- Ittzés with respect to N-V complex predicates.

This talk goes through the arguments and concludes that Butt&Lahiri’s original claim must be upheld.
The Claim: Light Verbs are Diachronically Pertinacious

Butt and Lahiri investigate V-V complex predicates in Urdu and Bangla and make the following claim (Butt and Lahiri 2013, 26–27):

*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.*

Note:

- Both Slade and Ittzés understand this claim in slightly different ways.
- It is important to note that this is a claim about the organization of the lexicon, not about the precise shape or form of different complex predcations.
The Claim: Light Verbs are Diachronically Pertinacious

The Claim:

- Light verbs are not the product of historical change.
- For a small set of core verbs (mostly motion verbs and verbs of putting and placing, ‘do/make’, ‘be, become’), a “lighter” version of the main verb is available synchronically (throughout the ages).

\[\text{Underlying Entry} \quad \text{Light Verb} \quad \text{Main Verb} \quad (\text{Auxiliary via reanalysis})\]

- This light verb combines with a main predicator to modify an event predication: it forms a complex predicate together with another predicator.
  - The argument and event structure is complex and composed.
  - But in terms of dependency structure, the result is a monoclausal predication.
The Urdu Permissive: A Monoclausal Complex Predicate

*Nadya let Yassin [cut the plant].*

- composed a(rgument)-structure:
  
  GIVE/LET < agent goal;  
  CUT < agent;  patient >>

- f(unctional)-structure

```
[  
  [PRED 'let-cut < SUBJ, OBJ<go>, OBJ > ']  
  [SUBJ [PRED 'Nadya' ] ]  
  [OBJ<go> [PRED 'Yassin' ] ]  
  [OBJ [PRED 'plant' ] ]  
  [TNS-ASP [TENSE PAST ] ]  
  ]
```
Types of Complex Predicates

Several different types of complex predicates in Urdu (and South Asian Languages in general).

All of them involve combined argument structures that correspond to a monoclausal dependency structure.

- permissives (Butt 1994, 1995)
- morphological causatives (Butt 1998)
- A-V complex predicates (Butt and King forthcoming)
Butt & Lahiri were curious about where complex predicates originated from and how they were formed diachronically.

Assumption (following Hook (1991)): reanalyzed from main verbs and destined to become auxiliaries via a light verb (vector) stage.

However: we found no diachronic evidence for a light verb moving away from a main verb and then being reanalyzed as an auxiliary.

So we concluded (Butt and Lahiri 2013, 18):

\[
\text{… that the use of light verbs can be traced to a very early stage of the language and that at every stage, the light verb was form-identical to a main verb. This suggests to us that rather than engaging in the type of reanalysis which leads to the formation of auxiliaries or modals, light verbs are diachronically inert and have been an integral part of the language at every stage.}
\]
We did find documentable instances in which a main verb was reanalyzed as an auxiliary.

This happens when a main verb embedding a clause is reanalyzed as a wholly functional element.

This can be illustrated schematically as shown in (1):

(1) Ravi [VP — to Venice fly] goes.

    Ravi to Venice fly.Fut. (=Ravi will fly to Venice)

This type of change has been documented extensively.

But how can one tell if a main verb or a light verb is the source for reanalysis to an auxiliary?

This is a tricky question (and therefore gives rise to controversy.)
Slade (2013) and Ittzés (2022) take issue with Butt & Lahiri’s claim of historical pertinacity of light verbs.

We argue that instead of casting doubt on Butt & Lahiri’s conclusions, both Slade (2013) and Ittzés (2022) actually provide more evidence for them.

They do so from two different perspectives:

- V-V complex predicates (Slade)
- N-V complex predicates (Ittzés).

In what follows, we begin with Slade (2013).
Adducing Evidence from Slade

- Slade (2013) claims to show that light verbs are not diachronically inert in Indo-Aryan.

- The evidence includes:
  1. The contemporaneous existence of a gerund interpretation.
  2. Differences between Hindi and Nepali.
  3. Morphosyntactic changes over time.
  4. The fact that a new progressive *auxiliary* derived from *rah* ‘stay/remain’
  5. Discussion of the *modal* *sak* ‘can/be able to’.
Age Old Ambiguity with the Gerund/Absolutive

- Slade focuses on V-V complex predicates of the type in (1).
- His argumentation is with respect to Hindi and Nepali, but the construction is also found in Bangla (which Butt&Lahiri look at).
- These complex predicates are known to be in principle ambiguous between a complex predicate reading and clausal adjunct reading.

(1) শামোলী দোর হু-এ প্লেল-ল-ও
Shyamoli.Nom shawl.Nom wash-Perf throw-Past-3
Reading 1: ‘Shyamoli washed (completely) the shawl.’ (complex predicate)
Reading 2: ‘Having washed the shawl, Shyamoli threw it.’
(Lahiri and Fitzpatrick-Cole 1999, 125)  

Historical note:

- The morpheme -e glossed ‘Perf’ in (1) is a remnant of the Sanskrit “gerund” or “absolutive” in -tvā(ya), or -ya/yā.
- This morpheme is realized as -i in Nepali but has been eroded to null in Urdu/Hindi.
V-V sequences are not always equally likely, it depends on the contextual situation and that some sequences are more entrenched in language use as either complex predicates or adjuncts.

In (2), it is highly unlikely that Shyamoli first beat the boys and then threw them somewhere.

(2) Shyamoli.Nom boy-Pl.Nom beat-Perf throw-Perf-Pres-3
Reading 1: ‘Shyamoli beat the boys (to death).’  
Reading 2: ‘Having beat the boys, Shyamoli threw them.’

(Lahiri and Fitzpatrick-Cole 1999, 136–137)
Similarly, Slade shows that not all V.Gerund-V sequences that can be found in the diachronic record allow for the adjunct reading (Slade 2013, 540).

(14) yathā sūcyā vāsah samāndhad iyād evam evāī as needle.SG.INSTR clothing.SG.NOM mend.pres_ptcp.sg.neu.nom go.pres.opt.3sg thus part
'tābhīr yajñasya chidram samāndhad eti these.pl.instr sacrifice.sg.gen defect.sg.nom mend.pres_ptcp.sg.neu.nom go.pres.3sg

a. “Just as one would [habitually] mend a garment with a needle, so with these one [habitually] mends [any] defect of the sacrifice.” (complex predicate reading)
b. “Just as one would move mending a garment with a needle, so with these one moves mending [any defect of] the sacrifice.” (literal present participle reading)

(Aitareya Brāhmaṇa 3,18,6; cf. Whitney 1879/1889. 1st/2nd edn: §1075a)

In effect, Slade thus provides more data that complex predication can indeed be found in all stages of Indo-Aryan.

All of his examples involve a light verb that is synchronically form-identical to a main verb.
Slade’s argument against Butt&Lahiri revolves around the morphosyntactic form of such examples.

For example:

- He points out that the morphology involved (present participle in the mending example) is not what is involved in New Indo-Aryan (NIA) languages.
- He points to differences in distribution and function of the V-V sequences.
- This includes a discussion of Nepali, which has V-V complex predicates, but whose precise morphosyntactic form and distribution differs from Hindi.

However: Butt&Lahiri made no claims as to the morphosyntactic form and distribution of light verbs.

It is already well-established that V-V complex predicates increased in frequency (presumably taking up the semantic space and function of the Sanskrit preverbs as those were lost), see the detailed work by Hook (1993, 2001) and in particular, Hook and Pardeshi (2009).
Existence of Light Verbs in the Lexicon

- Recall: 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. (Butt and Lahiri 2013, 18)

- The statement is about the existence of a light verb use that is form-identical to a main verb.
- The claim can be disproven if one finds productively used light verbs that are not form-identical to a main verb.
- But the precise morphosyntactic form and distribution of the complex predicate constructions is not at issue.
A New Progressive

- Slade points to Urdu/Hindi and Nepali *rah* ‘stay/remain’ as an instance of a light verb being reanalyzed as an auxiliary.

- In both Urdu/Hindi and Nepali the *rah* ‘stay/remain’ is clearly an auxiliary (much distributional evidence, see data in Slade (2013) and Butt and Geuder (2001)).

(3)

a. अन्जुम अदनान=को मर रह-ि हे
Anjum.F.Sg.Nom Adnan.M.Sg=Acc hit stay-Perf.F.Sg be.Pres.3.Sg
‘Anjum is hitting Adnan.’ (Butt and Rizvi 2010) (Urdu/Hindi)

b. कुट्टा भोक-ि रह-ि
dog.M.Sg.Nom bark-Impf-M.Sg stay-Perf.M.Sg
‘The dog kept on barking.’ (Butt and Rizvi 2010) (Urdu/Hindi)

c. माँ मन्दिर-ि गई- रह-ेको चु
I temple-Loc go-Abs remain-Perf_Ptcp.M.Sg be.1.Pres
‘I am going to the temple.’/‘I have been going to the temple.’ (Slade 2013, 564) (Nepali)

- But was it really a light verb before?
It is often tricky to distinguish between auxiliaries, modals and light verbs (see, e.g., Butt (2010) for some diagnostics)

Thinking about their respective contribution in terms of event semantics allows for a clear distinction between auxiliaries/modals and light verbs.

- Light verbs **contribute** to an independently existing event predication at the subevental level.
- Auxiliaries **situate** an event in **time**. They do not modify the basic event predication.
- Modals **situate** an event with respect to **possible worlds**. They do not modify the basic event predication.

Auxiliaries and modals do not modify the primary event predication
- they do **not** form complex predicates
- and are subject to diachronic reanalysis
Slade asserts that *rah* ‘stay/remain’ originated as a light verb, but does not demonstrate this.

Deo (2006) assumes it is an auxiliary in a periphrastic construction.

Previously (up until 19th century Hindi) the imperfective morphology could also denote progressives (Deo 2006, 176).

\[
\begin{align*}
(18) & \quad a. \text{ niśā } \quad \text{ roṭi } \quad \text{ banā } \text{ rah-i } \quad \text{ hai} \\
& \quad \text{N.NOM bread.NOM make-PROG.F.SG be-PR.3.SG} \\
& \quad \text{niśā is making bread. not *niśā makes bread.} \\
& \quad b. \text{ niśā } \quad \text{ roṭi } \quad \text{ banā-ti } \quad \text{ hai} \\
& \quad \text{N.NOM bread.NOM make-IMPF.F.SG be-PR.3.SG} \\
& \quad \text{niśā makes bread. not *niśā is making bread.}
\end{align*}
\]

As part of a general reorganization of the Indo-Aryan tense-aspect system, *rah* comes to denote specifically a progressive reading.

→ The introduction of *rah* as a progressive is a very recent phenomenon.
The Progressive

- Slade points to Kellogg (1893) for the classification of *rah* as a light verb.
- Indeed, Kellogg lists *rah* among a list of verbs that would generally agree to be light verbs (he calls them ‘intensives’), with the familiar range of extra semantic dimensions of completion/result and directionality.
- However, his discussion indicates an auxiliary status rather than a light verb.
  - The semantics is continuative (related to progressive): *beth rahna* ‘to sit still’ (emphasis Kellogg’s). (Kellogg 1893, §427)
  - It is the only verb in the list that has this type of meaning.
  - Kellogg further notes that:
    
    *Compounds with rahna are common, but for the most part are used in the tenses of the past participle. As above remarked, these compounds exhibit the action of the verb, emphatically, as continuing or permanent. (Kellogg 1893, §428.e)*

- This points to a selective use of forms — this is characteristic of an auxiliary.
Deo (2006) shows that the imperfective morphology in Indo-Aryan:
- goes from being able to denote both progressive and non-progressive (present, habitual, generic) situations
- to only expressing non-progressives

Progressives are expressed via the addition of tense auxiliaries ('be').

In Urdu/Hindi (and Nepali) additionally rах ‘stay/remain' was recruited.

This is a crosslinguistically well-established type of change: recruiting verbs of standing/remaining for continuative/progressive meanings, e.g., see Bybee et al. (1994).

The available data does not exclude the possibility of rах having been reanalyzed from a light verb version.

But it also does not conclusively establish rах as a light verb.
The most likely scenario is the crosslinguistically well-established reanalysis of a main verb as a progressive.

This is what happened in Bangla (see Lahiri (2000), Butt and Lahiri (2013)) with the new progressive formed from the main verb version of *atʃ* ‘be’.

The schema below provides a sketch with the root *pa* 'receive' (Butt and Lahiri 2013, 20).

Note: the -i on *pa* 'receive' is a descendent of *tvā/yā*.

**Formation of New Progressive:**

\[ \text{[pa} + i]_ω + [/atʃ^h/ + \text{Suffix}_{\text{PERS/NUM}}]_ω \]

\[ > \text{[[pa} + i]+[ tʃ^h + \text{Suffix}_{\text{PERS/NUM}}]_{\text{CLITIC}}]_ω \]

\[ > \text{[[pa} + i + tʃ^h_{\text{AFF}}] + \text{Suffix}_{\text{PERS/NUM}}]_ω \]

\[ > \text{[[pa} + tʃtʃ^h_{\text{AFF}}] + e]_ω \]
‘can’ and other arguments

- Slade invokes a number of other examples to argue for a light verb to auxiliary reanalysis. Among them:
  - Nepali honorifics
  - Nepali *topalnu* ‘to pretend’
  - The modal *sak* ‘can’

- In no case does he conclusively establish that the items are light verbs.

- The Nepali honorific is originally borrowed from Persian and is used in a formulaic construction.

- Both the verbs ‘pretend’ and ‘can’ are not light verbs: they embed a full VP with denotes an independent event.

- Slade assumes that ‘pretend’ and ‘can’ must be light verbs because the *surface form* of the V-V sequence looks like that of V-V complex predicates.
But we know that a given surface form does not necessarily correspond to exactly one structure (see the ambiguity with the \(-tvā/ya\) gerund)

With ‘pretend’ and ’can’ we can see clearly that the overt string might look similar to V-V complex predicates, but the meaning is very different.

(4) a. \(u\) gā-i- ṭopal-dai- cha
   he/she sing-Abs- pretend-Impf_Ptcp- be.Pres.3.Sg
   ‘He is pretending to sing.’ (Pokharel 1991, 195) (Nepali)

b. vo ga sāk-t-a ʰe
   Pron.3.Sg.Nom sing can-Impf-M.Sg be.Pres.3.Sg
   ‘He can sing.’ (based on Slade 2013, 560) (Urdu/Hindi)

These are clearly not light verbs, but main verbs embedding another main verb — there are two very separate events involved.

And neither was ‘can’ a light verb in Sanskrit, rather it was a raising verb (cf. John Lowe’s invited SALA-37 talk & Lowe et al. (2021)).

(1) a. rājāno rāmaṁ hantum na śaknuvantī
   kings.NOM.PL.M R.ACC.SG.M slay.INF not can.3PL
   ‘The kings cannot slay Rāma.’

b. rāmo rājabhir hantum na śakyate
   R.NOM.SG.M kings.INS.PL.M slay.INF not can.PASS.3SG
   ‘Rāma cannot be slain by the kings.’
Interim Summary

- Slade makes a number of very interesting points with respect to whether light verbs may be subject to further reanalysis.
- However: he mistakes Butt&Lahiri’s claim to be about the construction rather than the light verb.
  - Butt&Lahiri are quite clear on changes with respect to the construction and link the rise in use of V-V complex predicates to the demise of the preverbs.
    
    *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). (Butt and Lahiri 2013, 16)*

- Slade furthermore adduces evidence that:
  - Assumes light verb status rather than proving it.
  - Fails to look beyond the surface string to the underlying structure.

→ In the end there is no evidence that goes against Butt&Lahiri.
N-V in Old Indo-Aryan

- Ittzés (2022) looks at the Old Indo-Aryan (OIA) use of alternative ways of forming the perfect.
- Since this involves the verbs ‘do’, ‘be’ and ‘become’ he couches this in a larger discussion of grammaticalization.
- He posits that the OIA ‘do’ serves as a light verb and the N-am+V construction gives rise to an alternative periphrastic perfect (which then is primarily realized with ‘be’).
- But there is no evidence that ‘do’ itself turned into an auxiliary.
From the perspective of Urdu/Hindi, the discussion in Ittzés (2022) is very interesting as the patterns described are very similar to what is found with N-V complex predicates (Mohanan 1994).

In particular, the three major light verbs ‘do’, ‘be’, and ‘become’ exhibit differences in frequency and constraints on combinatory possibilities.

Ahmed and Butt (2011) set out to understand some of this distribution for Urdu via a corpus study.
(5) shows the same noun *yad* ‘memory’ with the three different major light verbs (more are possible).

(5)  

a. nadya=ne kahani yad k-i  
Nadya.F.Sg=Erg story.F.Sg.Nom memory do-Perf.F.Sg  
‘Nadya remembered a/the story.’ (lit.: ‘Nadya did memory of the story.’)

b. nadya=ko kahani yad hε  
Nadya.F.Sg=Dat story.F.Sg.Nom memory be.Pres.3.Sg  
‘Nadya remembers/knows a/the story.’  
(lit.: ‘Memory of the story is at Nadya.’)

c. nadya=ko kahani yad hu-i  
Nadya.F.Sg=Dat story.F.Sg.Nom memory be.Part-Perf.F.Sg  
‘Nadya came to remember a/the story.’  
’ (lit.: ‘Memory of the story became to be at Nadya.’)
We know these are complex predicates because each of the predicational elements contributes arguments to a joint prediction:

- *kahani* ‘story’ is only licensed by *yad* ‘memory’, not ‘do’.
- But *kahani* is not in the genitive, as it would be as an argument of a noun.
- Instead, it is in the nominative and acts as an object of the main clause (can be passivized, verb agrees with it, etc.)

\[(6) \quad \text{nadya}=\text{ne} \quad \text{kahani} \quad \text{yad} \quad \text{k-i} \]

\[
\text{Nadya.F.Sg=}\text{Erg} \quad \text{story.F.Sg.Nom} \quad \text{memory} \quad \text{do-Perf.F.Sg}
\]

‘Nadya remembered a/the story.’ (lit.: ‘Nadya did memory of the story.’)
The verbs are **light verbs** because:

- they contribute an argument to the joint predication (effect can be seen on the case marking on the subject).
- they augment the predication of the main element (‘memory’)
- but do not situate it in time (as an auxiliary would do) or with respect to possible worlds (as a modal would do).

\[(7)\]

(a) \textit{nadya=ne kahani yad k-i} \\
Nadya.F.Sg=Erg story.F.Sg.Nom memory do-Perf.F.Sg \\
‘Nadya remembered a/the story.’ (lit.: ‘Nadya did memory of the story.’)

(b) \textit{nadya=ko kahani yad he} \\
Nadya.F.Sg=Dat story.F.Sg.Nom memory be.Pres.3.Sg \\
‘Nadya remembers/knows a/the story.’
For example, we get the composed transitive predicate ‘do-memory’ with an agent (‘Nadya’) and a theme (‘story’) from the pieces below.

```
do < agent %Pred >
  |
  memory < theme >
```

- %Pred is a variable that indicates a slot to be filled by another predicate.
- This 'incomplete predication' is what characterizes light verbs according to Butt (1995) and Alsina (1996).
While the basic structure is the same for all N-V complex predicates, the light verbs differ as to what nouns they will combine with.

The constraints on the combinations are difficult to work out.

Ahmed and Butt (2011) did a corpus study with ‘do’, ‘be’ and ‘become’ and found three major classes.

- Corpus: 100,000 words from the Penn Treebank translated into Urdu and POS-tagged (built in Lahore)
- 45 nouns
- 4 of 45 allowed for the full range (all three light verbs): these are all psych nouns (like ‘memory’ and ‘belief’)
- 38 of the 45 allowed only for ‘do’. These are all agentive event nouns such as ‘signal’, ‘construction’ or ‘refusal’.
- 3 of the 45 disallowed ‘become’: taslim ‘acceptance’, bardaf ‘tolerance’ and intizar ‘wait’.
Ahmed and Butt (2011) did not conclude that there is any form of grammaticalization going on in Urdu.

The differences in frequency and combinatory possibilities are instead attributed to different (semantic) classes of nouns.

**Question:** Could that work for Old Indo-Aryan as well? (or is it all morphosyntactically conditioned, cf. Grestenberger and Ittzés?).

**Upshot**

Ittzés (2022) as well as Ittzés’ SALA-37 talk seem to provide more evidence for Butt & Lahiri’s claim that light verbs were already available within the overall language structure in Old Indo-Aryan.
Conclusions

- Butt&Lahiri worked on V-V complex predicates.
- Slade takes this same domain and tries to argue that:
  - The V-ya V sequence did not contain light verbs and so there is no direct connection back from the modern languages to OIA.
  - Light verbs can be reanalyzed into auxiliaries
- But:
  - The evidence for reanalysis of light verbs does not hold up.
  - Butt&Lahiri made no deep claims about particular constructions, just about the existence of complex predicate formation.
- Ittzés works on a different domain: N-V sequences.
- His data seems to suggest OIA already had N-V complex predicates along roughly the same lines as modern Urdu/Hindi.
- This bolsters Butt&Lahiri’s original claim.
Thanks

We would like to thank the organizers of this workshop, Eystein Dahl and Beatrice Grieco, for conceiving this idea and bringing together so many people looking at the question of complex verbs in Old Indo Aryan!


Conclusions

References II


