

Shishir Bhattacharja*

On the formation of complex predicates in Bengali

<https://doi.org/10.1515/jsall-2018-0005>

Abstract: In this article we show, in the light of Bengali data, how verbal constructions known as Complex predicates can be handled in grammar. These constructions are generally described as constituted of two items, the former chosen among various categories of words: noun, verbal forms, adjective, preposition, etc., and the latter, a normally inflected verb. We argue that such constructions are words, and it is preferable to handle them exclusively in morphology. We assume, in the light of Whole Word Morphology, that a Word Formation Strategy may become part of the morphological module of a speaker-hearer if her lexica contains a set of semantically related word-pairs based on the same (i) formal contrast and (ii) categorial affiliation. Hence the individual mental lexica of Bengali speaker-hearers contain sets of pairs of words constituted of simple and complex predicates (such as *likhe* ‘he writes’ ~ *likhejay* ‘he continues to write’, etc.). These pairs license particular WFSs (such as $/Xe/v, 3SG\ PRS \leftrightarrow /Xe/v, 3SG\ PRS$) which can be activated as needed, to form, remember or retrieve other complex predicates such as *bolejay* ‘he continues to speak’. Therefore, there is no need to list each one of them separately in a mental lexicon.

Keywords: complex predicates, compound verbs, conjunct verbs, word-formation

1 Introduction

Verbal constructions that are generally described as *Complex predicates* are usually constituted of two items, one chosen among various categories of words: *noun*, *verbal forms*, *adjective*, *preposition*, *adverb*, *onomatopoeic*, etc., and the other, a normally inflected verb. The first item is called a *Pole* and the second one a *Vector* (a.k.a. *explicator*, *operator* or *light verb*) (cf. Hook 1974: 17; Masica 1976: 143; Dasgupta 1977: 68; Butt 2003). These predicates are generally put into two different categories on the basis of the syntactic category of the pole: (i) *Compound verb* (1, 2) that categorially involves a non-inflectional verbal form (e.g. a participle, an absolutive or a past gerund), and (ii) *Conjunct verb* (3, 4) that involves categories

*Corresponding author: Shishir Bhattacharja, Institute of Modern Languages, University of Dhaka, Dhaka, Bangladesh, E-mail: bhattacharjashishir@gmail.com

other than the verb. In other words, a ‘Compound verb’ results from the combination of two different verbs, and a ‘Conjunct verb’, from the combination of a verb and another word that belongs to some other category. The main characteristic of Compound and Conjunct verbs is that they must denote one single action.¹

- (1) Rik [likhe phele]
Rik having.written throw.off.3SG.PRS
‘Rik finishes writing’
- (2) Rik [likhte thake]
Rik to.write.Compl.verbal stay.3SG.PRS
‘Rik continues to write (something).’
- (3) Gargi [bazar kore]
Gargi market do.3SG.PRS
‘Gargi has done shopping.’
- (4) daktar rogi [bhalo kOre]
doctor patient good do.3SG.PRS
‘Doctors cure patients.’

There seems to be no general consensus either about (i) the *domain* or about (ii) the *process* of formation of such predicates (cf. Butt 2003). In this article, we will try to contribute to these two problems in the light of Bengali data constituted of sentences that are acceptable according to the competence of the author in this language.² This article is divided into 6 sections. In Section 2, we briefly present a couple of solutions that have been proposed by authors who prefer to handle complex predicates in syntax, and then we argue why syntax is not the right place to account for them. In Section 3, after a brief description of W(hole) W(ord) M(orphology) → the model we shall use in this article – we show how complex predicates can be handled in morphology. In Section 4, we argue that complex predicates are words. In Section 5, we point out a few problems of handling complex predicates in morphology in the light of atomistic models of word formation, i.e. models that must involve units smaller than the word, before, finally, we draw conclusions in Section 6.

¹ These verbs should not be confused with two other similar constructions known as (i) *Phrasal verbs* (e.g. *put out* or *lie down* in English) and (ii) *Serial verbs* (in Bengali: *potrika kine, phon kore, bajar niye hente chole asho* ‘newspaper having.bought, phone having.done, market having.taken, having.walked having.moved come’ = ‘buy the newspaper, make a phone call, do your shopping and then come on foot’).

² The transliteration does not necessarily reflect the standard pronunciation.

2 Bengali complex predicates: Are they formed in syntax?

In our view (inspired by Singh 1996) one must try to answer the following three questions of *Domain*, *Status* and *Process* in order to account for any linguistic construct.

- (i) In which domain are these constructs formed, for example, in Syntax or in Morphology?
- (ii) What is the status of these constructs, should they be considered *Phrases* (or *Clauses*), or are they *Words*?
- (iii) Through which particular process(es) are they formed?

The unmarked stand vis-a-vis the domain of the formation of complex predicates is that they are formed in syntax. Grimshaw and Mester (1988) describe Japanese complex predicates as phrases whereas Kageyama (2001) describes them as Word + (word plus), an intermediate category between the word and the phrase. For Matsumoto (1996) Compound verbs are biclausal syntactic structures.³ Authors who claim that words may be formed in syntax propose a couple of processes: (i) A(rgument) L(inking) P(rinciple) (cf. Lieber 1983; Lieber 1992, 2004 among others)⁴ or (ii) A(bstract) I(ncorporation)

³ According to Matsumoto (1996) Japanese conjunct verbs are bi-clausal constructions in which one predicate functions as the complement of the other while the subject of one predicate controls or binds (+ c-command, + co-indexation) the covert subject (Pro) of the other predicate. Hence, in (a) the flexional verb *suru* and the verbal noun *keikoku* are two independent predicates, where the subject (John) of *suru* controls the covert subject (Pro) of *keikoku*.

- (a) [John-wa_i [murabito-ni ookami-ga kuru-to-no
 John-TOP_i villager-DAT wolf-NOM come-COMP-GEN
 (PRO)_i keikoku-o] suru]
 warn-ACC does
 'John warns the villagers that the wolf is coming.'

⁴ In Lieber (1983, 1992, 2004), compound formation is constrained by the A(rgument) L(inking) P(rinciple) according to which a verb or a preposition must be able to link its internal arguments with she (1983: 257) defines as follows: "all obligatory (i.e. lexically specified) arguments with the exception of the subject are internal." For instance, the verbs *kOre* 'he does' links its theme *bazar* and *phon* in the Conjunct verb *bazar kOre* 'go shopping' and *phon kOre* 'make phone calls' respectively.

(cf. Kageyama 1991),⁵ both based on the semantic relation between the argument noun appearing as the pole and the verb appearing as the vector. According to Grimshaw and Mester (1988) and Matsumoto (1996) Japanese complex predicates are formed through syntactic processes which usually form phrases and clauses.

It is unclear how word-formation processes like ALP or AI can account for Bengali Compound verbs which are constituted of two different verbs (1, 2) or for Bengali Conjunct verbs in which the pole is represented by a word belonging to a category other than the noun (4) (because such poles cannot appear as the argument of a predicate). Unlike Compound verbs (1, 2), Conjunct verbs (3, 4) are not bi-clausal constructions and it is unclear how they can be handled in frameworks like Grimshaw and Mester (1988) or Matsumoto (1996). However, one can always argue that at least some Conjunct verbs, such as (3), can be formed through ALP or AI and some others with syntactic processes.

The question of status (ii) and that of process (iii) are both in fact redundant, in that it is the domain that automatically determines the status of the construct in question. For example, linguists generally do not disagree with the fact that syntactic constructs such as *phrases* and *clauses* are semantically more transparent and formally less cohesive as compared to words and idioms. If this is in fact the case, it would be reasonable to handle them in some domain other than the syntax. Processes, on the other hand, can determine neither the domain nor the status of the output because, logically, the same (type of) processes may output different things in different domains. For example, as we have seen above, linguists claim to account for the formation of words in syntax with processes that are otherwise used to form phrases and clauses.

In the following, we argue, on the basis of empirical data, why syntax is not the right domain for handling complex predicates in Bengali.

i. Unlike phrases, Complex predicates are generally semantically opaque. For example, the Compound verb *likhe pheleche* ‘he has finished writing’ (1) cannot be decomposed into two different verbs: (a) *likhe* ‘having written’, and (b) *pheleche* ‘he has thrown off’, because, as we can see, the basic meaning of the verb *pheleche* is non-existent in that Compound verb.

⁵ Kageyama (1991: 196), following Baker (1988), claims that Japanese Conjunct verbs (such as *keikoku suru* ‘warning do’ = ‘to warn’) are formed through a process called A(bstract) I(ncorporation) which he defines as “an instance of Incorporation that does not manifest any physical sign of movement but only gives abstract co-indexes to an incorporating host and an incorporated element”. Hence the verbal noun head (e.g. *keikoku* ‘warning’) appears, in an underlying syntactic structure, as the argument of the verb head *suru*, and then, the latter incorporates the former. According to Kageyama (1991: 196) AI is “basically a word-formation process, though it takes place in the syntax.”

ii. Like words, Compound and Complex verbs have specific meaning. For example, the Conjunct verb *bazar koreche* in (3) can only mean ‘he has done shopping’, and cannot have a meaning like, for example, ‘he has built a market’. Equally a Conjunct verb like *phon kOra* means ‘to make a phone call’ and not ‘to fabricate a phone set’. Again, the conjunct verb *bari kOra* ‘house do’ means ‘to make/build a house; to be/become owner of a house’, whereas the verb phrase *bari banano* ‘house build’, means ‘to build/construct a house (preparing it for habitation)’.

iii. Unlike phrases (5), Compound (6) and Conjunct (7) verbs resist deletion of one of their constituents, which points to the fact that they have some sort of formal integrity similar to words.

- (5) *ami dupure [bhat khai] rate ruti*
 1SG noon.LOC rice eat.1SG.PRS night.LOC bread
 ‘I eat rice in lunch and bread at dinner’

- (6) **ami [ghumiye thaki] o [boshe]*
 1SG having.slept remain.1SG.PRS 3SG sit.PFV.VERBAL
 ‘I sleep and he remains seated’

- (7) ²*ami shokale [bazar kori] rate [ranna]*
 1SG morning.LOC market do.1SG.PRS night.LOC cooking
 ‘I go for shopping in the morning and cook at night’

iv. As we can see in Table 1, not all vector verbs can combine with all verbs or all verbals.⁶ According to Dasgupta (1977: 70) the incompatibility between particular vector verbs and pole verbs is due to their (sub)categorial mismatch or ‘transitivity harmony’. Although Dasgupta’s rule of thumb works in a good number of cases, exceptions abound. For example, in the Compound verb *shune jawa* ‘continue to listen’ a transitive verb *shuna* ‘listen’ combines with an intransitive verb *jawa* ‘to go’. Equally, *ghumiye neya* ‘to have slept for a while’ is a combination of the intransitive verb *ghumano* ‘to sleep’ and the ditransitive verb *neya* ‘to take’.

v. As in other South Asian languages (cf. Masica 1976: 143) the meaning of complex predicates in Bengali is varied and manifold. Depending on the verbs

⁶ Bhattacharja (2007) classifies verbals in Bengali into four different types: (i) *Perfective verbals* which denote an accomplished action; (ii) *Argument verbals* which are used as subject noun or complement of verb or a preposition; (iii) *Complement verbals* which generally appear as complement of modal verbs; (iv) *Conditional verbals* which have an adverbial use of *if...then* type.

Table 1: Vector verbs and different verbs and verbals appearing as their pole: a selection.

Argument verbal	Pole		Vector
	Perfective verbal	Complement verbal	
*kora 'to do'	kore	korte	dewa 'to give', e.g. <i>O korte dey</i> '(S)he allows one to do.'
*ghumano 'to sleep'	*ghumiye	ghumate	dewa 'to give', e.g. <i>O ghumate dey</i> '(S)he allows one to sleep.'
*kora	kore/ ghumiye	korte/ *ghumate	newa 'to take', e.g. <i>O ghumiye ney</i> '(S)he sleeps for a while.'
*shuna 'to listen'	*shune	shunte	pawa 'gets', e.g. <i>Ami shunte pai</i> 'I can hear (despite hindrances).'
*kora	kore/ ghumiye	korte	thaka 'to remain', e.g. <i>O kore thake</i> '(S)he usually does.'
*kora	*kore	korte	howa 'to be', e.g. <i>Or korte hoy</i> '(S)he has to do'
*bosha 'to sit'	boshe	*boshte	roy 'to remain/exist', e.g. <i>O boshe roy</i> '(S)he remains seated.'
bosha	boshe/ ghumiye	?boshte	ache 'exists' e.g. <i>O boshe ache</i> '(S)he has been sitting for a while.'
kora	kore/ ghumiye	?korte	chola 'to continue' e.g. <i>O kore chole</i> '(S)he does something incessantly.'
kora	kore/ ghumiye	korte	jawa 'to go' e.g. <i>O kore jay</i> '(S)he does incessantly.'
*kora	kore/ ghumiye	?korte	dekha 'to see' e.g. <i>O kore dekhe</i> '(S)he does something to see what happens.'
*kora	kore/? ghumiye	*korte	phela 'to throw away' e.g. <i>O kore phele</i> '(S)he does something all of a sudden and repents thereafter.'
*kora	kore/? ghumiye	*korte	rakha 'to keep' e.g. <i>O kore rakhe</i> '(S)he does something beforehand.'
*kora	kore/? ghumiye	*korte	chara 'to leave' e.g. <i>O kore chare</i> '(S)he does something anyway.'
*kora	kore/ *ghumiye	*korte	ana 'to bring' e.g. <i>Tumi kore ano</i> 'You do something beforehand and bring it in.'
*kora	kore/? ghumiye	*korte	berano 'to roam around' e.g. <i>O kore beray</i> '(S)he keeps on doing something here and there.'
*haNtano 'to make (someone) walk'	haNtiye	*haNtate	mara 'to kill' e.g. <i>O haNtiye mare</i> '(S)he obliges to walk and makes one tired.'
*haNta 'to walk'	heNte	*haNte	mora 'to die' e.g. <i>O heNte more</i> '(S)he walks for nothing and becomes tired.'
*kora	kore	korte	bosha 'to sit' e.g. <i>O kore boshe</i> '(S)he does something all of a sudden and repents thenafter.'

involved and on the circumstances in which they are used, they can connote different “manners of action” (not to confuse with aspect) such as Completion (1), Forethought (8), Continuation (9), Benefaction (10), Directionality (11), Stubbornness (12), Regret (13), etc. (see Masica *ibid.* for a more exhaustive list of such manners).

- (8) *Rik [likhe rakhe]*
 Rik write.PFV.VERBAL keep.3SG.PRS
 ‘Rik writes beforehand.’
- (9) *Rik [likhe jay]*
 Rik write.PFV.VERBAL go.3SG.PRS
 ‘Rik continues to write.’
- (10) *Rik [likhe dey]*
 Rik write.PFV.VERBAL give.3SG.PRS
 ‘Rik writes for someone else.’
- (11) *Rik [likhe ane]*
 Rik write.PFV.VERBAL bring.3SG.PRS
 ‘Rik writes somewhere else and brings the document.’
- (12) *Rik [likhe chare]*
 Rik write.PFV.VERBAL leave.away.3SG.PRS
 ‘Rik does not stop until he finishes writing.’
- (13) *Rik [likhe boshe]*
 Rik having.written sit.3SG.PRS
 ‘All of a sudden Rik writes, but then realizes that he should not have done so.’

vi. The meaning of Complex predicates may also depend on pragmatics. For example, the verb *boshe achi* ‘having.sat I.am’ may mean either ‘I am sitting’ or ‘I am waiting’ depending on the circumstances in which the verb is used.

As a whole, much idiosyncrasy and connotation is involved in complex predicates and syntax does not seem to be the right place to handle such things. Due to their idiosyncratic character and numerous connotations, it is difficult to consider Complex predicates as syntactic constructs (such as sentences and phrases), which are usually semantically transparent and compositional. If complex predicates cannot be handled satisfactorily in syntax, then one should

try to account for them in morphology. If that also fails, then each one of these predicates has to be listed in the lexicon, like idioms, for example.

3 Complex predicates: Are they formed in morphology?

Models of morphology can be put into two different categories: *Atomistic* and *Holistic*. By ‘atomistic’ we refer to models that must involve units smaller than the word, such as *affix* or *stem*, whereas a true holistic model involves none of these units in the process of word-formation. In what follows, after a brief description of WWM (elaborated in Ford et al. 1997; Singh 2006), a holistic model of word-formation, we will try to demonstrate how complex predicates can be handled in the light of this model.

According to Singh (2006: 578):

All that needs to be said about word structure in any language (of any type whatsoever) can and must be said by instantiations of the schema in (S1). These instantiations are referred to as Word Formation Strategies (WFSs) because, as generalizations drawn from known particular facts, they can be activated in the production and understanding of new words. WFSs must be formulated as generally as possible, but - and this is crucial - only as generally as the facts of the matter permit.

S1. $/X/_{\alpha} \leftrightarrow /X'/_{\beta}$ where

1. $/X/_{\alpha}$ and $/X'/_{\beta}$ are words and X and X' are abbreviations of the forms of classes of words belonging to categories α and β (with which specific words belonging to the right category can be unified or on to which they can be mapped).
2. ' represents (all the) form-related differences between $/X/$ and $/X'/_$ that fall outside of automatic phonology.
3. α and β are categories that may be represented as feature bundles.
4. The $\leftrightarrow$ represents a bidirectional implication (if $/X/$ then $/X'/_$, and if $/X'/_$, then $/X/$).
5. The interpretation of $/X/_{\alpha}$ is a semantic function of $/X'/_{\beta}$ and vice versa.
6. ' can be null iff $\alpha \neq \beta$.

Singh (2006: 578) states that

S1 offers a unified account of what have sometimes been seen as different types of morphologies and encapsulates the rejection of multipartite analysis of words into ‘roots’, ‘affixes’, ‘stems’, and so on, entries that are hard to define and harder to tell apart.

For WWM, words have no internal (non-phonological) hierarchical structure. According to this model (cf. Singh 2006: 578): “Morphological complexity is a matter of analyzability (# segmentability) of a word into a variable and a constant component with respect to a WFS.”

According to Singh (2006: 578) WWM sees “morphology” “not as a combinatorics of morphs or morphemes but as a system of generalized and abstract bidirectional correspondence among patterns instantiated by sets of whole words that exploit the same contrast.

For example, on the basis of morphologically related sets of words like (14) and (15), one can obviously set up a WFS like (16) which is licensed by a set of semantically related word-pairs that manifest the same (i) formal contrast: X/Xness on the one hand, and (ii) categorial affiliation: Adjective/Noun on the other. According to (16) *kind*, *bold*, *bright*, etc. provide the differing values for the variable X in (14), (15) while the phonemic sequence *ness*, which remains constant throughout (15), provides the particular value of the prime (') in the schema (S1).⁷

(14) *kind*, *bold*, *bright*, etc.

(15) *kindness*, *boldness*, *brightness*, etc.

(16) /X/_{Adj} ↔ /Xness/_N

Subcomponents can be represented by any phonic element: single phoneme, meaningless sound cluster, simple or complex word, discontinuous and inseparable segmental as well as supra-segmental means like stress and tone (variables, however, cannot be exclusively supra-segmental). For example, if the Hebrew word *hagdala* ‘enlargement’ is mapped onto (17), the *variable* will be represented by the discontinuous sequence of consonants: /h/-/gd/-/l/ and the *constant* by the discontinuous sequence of vowels: /a/-/a/-/a/.

⁷ WWM is, to some extent, one of the modern avatars of the model of morphology based on “proportional analogy” or “four-term proportion” which Saussure (1915: 253) called “the new school” in contrast to the old one which analyzes words into smaller units.

- (17) /CaCCaCa/_N ↔ /CiCCiC/_V
 /hagdala/ ‘enlargement’ ↔ /higdil/ ‘enlarge’
 /haxtaba/ ‘dictation’ ↔ /hixtib/ ‘dictate’
 (Booij 2005: 38)

Similarly, if *ts^hu³⁵* ‘a plough’ is mapped onto (18), a WFS of Chinese, the constant will be represented by the rising tone and the variable with the sequence of segments. Quite coincidentally, when a subcomponent is represented by a continuous sequence of phonemes (*ness* for example) it can have phonic resemblance to word parts which some grammatical traditions would label as *morpheme*, *root*, *stem* or *affix*.

- (18) /CV¹¹/_V ↔ /CV³⁵/_N
 /mu¹¹/ ‘to grind’ ↔ /mu³⁵/ ‘a grind’
 /ts^hu¹¹/ ‘to plough’ ↔ /ts^hu³⁵/ ‘a plough’
 (Yu 2007: 191)

We shall now move onto demonstrate how WWM can account for complex predicates. We assume that the lexicon of a Bengali speaker contain sets of word-pairs constituted of verbals like *likhe* or *bOle* and also verbs like *likhejay* and *bOlejey*. Such pairs license (19) or (20) which can be activated, if needed, to form, analyze and retrieve other Compound verbs like *shunejay* ‘he continues to listen’ or *dekhejay* ‘he continues to see’ by mapping verbs like *shune* ‘having listened to’ or *dekhe* ‘having seen’ onto the WFS. Each pair in (19, 20) manifests semantic relatedness between the pair-mates (between *likhe* and *likhejay* for example), each pair-mate belongs to a different morphological category, and finally, in both of the pairs we can see the same formal contrast represented by the sequence *jay*. The Compound verb *likhejay* can be analyzed into the variable *likh* and the constant *ejay* if we map it onto (19). However, the same verb can be analyzed into the variable *likhe* and the constant *jay* if we map it onto (20).

- (19) /XC/_{V, Imp} ↔ /Xe_{jay}/_{V, 3rd, Ind, Pres}
 /likh/ ‘write’ ↔ /likhejay/ ‘he writes non-stop/he continues to write’
 /bOl/ ‘say’ ↔ /bOlejey/ ‘he speaks non-stop/he continues to speak’
- (20) /Xe/_{Verbal} ↔ /Xe_{jay}/_{V, 3rd, Ind, Pres}
 /likhe/ ‘having written’ ↔ /likhejay/ ‘he writes non-stop/he continues to write’
 /bOle/ ‘having said’ ↔ /bOlejey/ ‘he speaks non-stop/he continues to speak’

Equally, the Conjunct verb *bazarkore* can be analyzed into the variable *bazar* and the constant *kore* respectively with (21). As we mentioned above, the sub-components (variable and constant) are represented by a mere sequence of phonemes and the fact that *likh* is isophonic with the regular word *likh* ‘write’ or that the constant *ejay* is not isophonic with any regular word, is irrelevant for this model of morphology.

- (21) /X/_N ↔ /XkOre/_{V, 3rd, Ind, pres}
 /bazar/ ‘market’ ↔ /bazarkOre/ ‘he does shopping’
 /phon/ ‘telephone’ ↔ /phonkOre/ ‘he makes phone calls’

4 Bengali complex predicates: Are they words?

One may argue that Bengali complex predicates are not words but phrasal constructions because a large number of them (e.g. 22, 23) lack formal cohesion in the sense that their pole and the vector part can be interrupted with some other words, and hence they lack formal integrity, one of the characteristics of words.

- (22) [bose] to Rik onekkhon dhorei
 sit.PFV.VERBAL EMPH Rik many-seconds during.EMPH
 [ache]
 exist.3SG.PRS
 ‘Rik has been sitting for a while’

- (23) [bazar] to Rik kobei [koreche]
 market EMPH Rik long.time.ago.EMPH do.3SG.PRP
 ‘Rik finished shopping long time ago’

Dasgupta (1977), however, claims that the constituting elements of Compound verbs cannot be separated from each other. His rule-of-thumb (1977: 75) is that if it is possible to insert an adverb between the pole and the vector, as in (24), then we are dealing with a sequence of poles (in other words, a serial verb), and if the particular construction shuns such insertions, then it is a Compound verb (25).

- (24) Rik [likhe] tarpor [jay]
 Rik write.PFV.VERBAL then go.3SG.PRS
 ‘Rik finishes writing and then he goes’

- (25) **Rik* [*likhe*] *tarpor* [*jay*]
 Rik write.PFV.VERBAL then go.3SG.PRS
 ‘*Rik* continues to write’

Although words are expected to have both (i) formal integrity and (ii) semantic opacity, some languages have been reported to have words that allow insertion of lexical and grammatical elements (cf. Sadock 1998). For example, in (26) the adjective *bhalo* can be separated from the rest of the subcomponents of the word *bhalobashe* ‘he loves’. Despite this separability, *bhalobasha* remains one single word because the sequence *basha* is completely devoid of meaning as a separate word in Modern Bengali.⁸ Such examples point to the fact that words may lack formal integrity. On the other hand, linguists generally do not disagree with the Frege-Wittgensteinian idea that words lack semantic compositionality. Hence, semantic opacity seems to be a more reliable and universal criterion for word-hood as compared to formal integrity. In other words, words are expected to have formal integrity but semantic opacity is a requirement for word-hood (which does not mean that other linguistic constructs that manifest semantic opacity, for example *idioms*, are words). As complex predicates categorially lack semantic compositionality, they are words, and like any other word, they are better handled in morphology.

- (26) [*bhalo*] *Rik* *Gargike* [*bashe* *na*]
 good *Rik* *Gargi*.ACC ??-3SG.PRS NEG
 ‘*Rik* does not love *Gargi*’

Although some Compound and Conjunct verbs happen to lack formal integrity, it seems that most of them must acquire it in the course of time. In some dialects of Bengali there are examples of Compound verbs (27, 28) that have been fused into simple verbs. (Conjunct verbs have not been noticed to have undergone such cohesion however.) However, a large number of Compound verbs are resistant to structural disruption (6, 7 and 25). It is possible that this resistance develops in order to prepare the way towards the formal cohesion that is finally achieved by some phonological change, and then by internal *sandhi*, at some point of development of the language.

Chittagonian

- (27) *di phalai* ‘give I.threw.off’ > *dialai* ‘I already gave’

⁸ *-basha* cannot be considered a suffix either, because no other word in Bengali is formed with this submorphemic unit.

Dialect of Kishorganj

(28) *diya phalaichi* ‘having.given I.threw.off’ > *dialchi* ‘I already gave’

Only a small subset of regular verbs (we have listed about 20 of them in Table 1) appear as vector in complex predicates in Bengali and also cross-linguistically (see Butt 2003), which points to the fact that verbs that are used as vectors have a different distribution compared to other regular verbs that are used only as poles. Complex predicates are not different from the words known as *compounds*, except for the fact that the latter categorially manifest formal integrity. The vector part of complex predicates are like *house* in *doghouse* (29), which has a different distribution as compared to *dog* which can be replaced by any noun denoting animal whereas *house* cannot be replaced by most nouns denoting ‘place of living’ such as *room*, *home*, *village*, etc. Hence *tiger-house*, *snake-house*, etc. are acceptable, but **dog-room*, **dog-home*, etc. are not. In the same way, only some of the twenty vector verbs (*phele* ‘gives’, *rakhe* ‘keeps’, etc.) of Table 1 can replace the vector *boshe* ‘has sat’ in *likheboshe* (30) whereas many verbs can replace the pole *likhe* ‘having written’.

(29) $/X/_N \leftrightarrow /Xhouse/_N$
dog $\leftrightarrow$ *doghouse*; *bird* $\leftrightarrow$ *birdhouse*

(30) $/Xe/_V, 3SG, PRS \leftrightarrow /Xeboshe/_V, 3SG, PRS$
/likhe/ ‘he writes’ $\leftrightarrow$ */likheboshe/* ‘all of a sudden he writes (something) but then he realizes that he should not have done that’
/bole/ ‘he says’ $\leftrightarrow$ */boleboshe/* ‘all of a sudden he says (something) but then he realizes that he should not have done that’

We claim, on the basis of (27, 28), that verbs that appear as vectors in complex predicates have already stepped into the process of grammaticalization and are heading towards becoming something that atomistic models generally describe as *affixes* – like *dom* in *kingdom* and *hood* in *boyhood* which were regular words at some point of diachrony. (This however does not mean that all vector verbs will cease to be used as regular verbs in the long run.) The process of grammaticalization involves, usually in consecutive stages, the loss of wordhood, loss of meaning, formal change and fusion with other subcomponents.

The Bengali verb/noun *bhalobashe* ‘he loves’ is a good example of what the process of grammaticalization may do to a word. In Middle Bengali *bashe* ‘he

feels' was used as a regular verb and also as a vector (see Chatterji 1970 [1926]: 899 for various uses of *bash*). In Modern Bengali, however, the regular verb use of *basha* is obsolete; it is devoid of any meaning (or "use" in the Wittgensteinian sense) and it appears exclusively in the verb/noun *bhalobasha* 'to love/love' and its paradigms.⁹

5 Problems with possible atomistic accounts of complex predicates

In this section, we will see whether atomistic models can satisfactorily handle all types of complex predicates. We have mentioned above that the vector part of the complex predicates may be considered as (becoming) suffixes or suffixoids that denotes manner (# aspect) of the action referred to. One of the options would be to list vector verbs and their regular verb counterparts as separate lexical entries. Vector verbs functioning as *suffixoids* are attached to the pole through some process after Aronoff (1976) or Lieber (1992). This is in keeping with Butt's (2003) claim that vectors evolve into auxiliaries and then into *clitics* before becoming affixes.

Although this seems to work in a large number of cases, there are examples like *bhalobasha* 'to love', *Ter pawa* 'to be aware of', *khabi khawa* 'to gasp (as a drowning person does)' which are difficult to handle in these models. In *bhalobasha* the sequence *bhalo* is isophonic with the regular word *bhalo* 'good', but the rest of the sequence: *basha* (or *bashi* in *bhalobashi* 'I love') is neither a word nor a stem, because, as we mentioned above, it is devoid of meaning.¹⁰ It cannot be a suffix either because no other word is formed with this submorphemic unit. We note that, contrary to Butt's (2003) claim, the vector in the verb *bhalobasha* as well as the vector in the complex predicates in (27, 28) have not evolved into some affix but was rather fused with the pole.

⁹ In the traditional approaches, the Bengali verb *bhalobasha* 'to love' would be considered a Conjoint verb because it results from the combination of the adjective *bhalo* 'good' and a verb. But as *basha* does not exist as a separate word, we do not see how one can consider *bhalobasha* a complex predicate in a synchronic analysis.

¹⁰ This justifies Butt and Tantos's (2004) claim that there is an intimate connection between a vector and its regular verb counterpart to the extent that if one ceases to exist, then the other also disappears simultaneously. For example, as *bash* appears with no other pole but *bhalo* 'good', we cannot consider it a genuine vector any more.

Equally, in *Ter pawa* or *khabī khawa* the sequences *pawa* and *khawa* are isophonetic with regular verbs, but the sequence *Ter* or *khabī* are devoid of meaning. These constructions can be considered neither as an *idiom* nor as a *phrase* because both of these categories usually require each of their constituting elements to have a meaning. Be that as it may, things like *basha*, *Ter* or *khabī* should be problematic for atomistic models which must give a name to, and/or list each and every word-internal part.

In a WWM description these are simply sequence of phonemes, and the fact that we find them in the verb *bhalobasha*, *Ter pawa* or *khabī khawa* is a mere coincidence. These verbs can be output with WFSs like (31, 32), which, as we can see, may be licensed, among others, by pairs constituted of simple predicates.¹¹ We note that both simple and complex predicates can be formed with WFSs instantiating the schema $/X/_{\text{a}} \leftrightarrow /X'/_{\text{b}}$, and therefore, from a morphological point of view, the idea of the complexity of predicates turns out to be an exclusively (atomistic) theory-internal matter.

- (31) $/X\text{Ce}/_{\text{V, 3SG, PRS}} \leftrightarrow /X\text{Ca}/_{\text{V, Inf}}$
/bhalobashe/ ‘he loves’ $\leftrightarrow$ */bhalobasha/* ‘love’
/kOre/ ‘he does’ $\leftrightarrow$ */kOra/* ‘to do’

- (32) $/X\text{Cai}/_{\text{V, 3SG, PRS}} \leftrightarrow /X\text{Cawa}/_{\text{V, Inf}}$
/Terpai/ ‘I am aware of’ $\leftrightarrow$ */Terpawa/* ‘to be aware of’
/bhOypai/ ‘I am afraid’ $\leftrightarrow$ */bhOypawa/* ‘to be afraid’
/khai/ ‘I eat’ $\leftrightarrow$ */khawa/* ‘to eat’

Although we have used traditional terms like *pole*, *vector*, *complex predicate*, *Compound verb*, *Conjunct verb*, etc. throughout this article to ease discussion, there will in fact be no need of using them if we adopt the WWM framework. Complex predicates are verbs, and like any other verb (or any other word), they can be formed, analyzed or retrieved with relevant WFSs, and therefore, neither they nor their subcomponents require different names.

¹¹ If some speaker-hearer maps the verb *Terpawa* onto the following WFS he will get the sequence *Ter* which is not a word.

$/X\text{pawa}/_{\text{V, Inf}} \leftrightarrow /X/_{\text{N}}$
/bhOypawa/ ‘to be afraid’ $\leftrightarrow$ */bhOy/* ‘fear’
/bethapawa/ ‘to feel pain’ $\leftrightarrow$ */betha/* ‘pain’
 **/Terpawa/* ‘to be aware of’ $\leftrightarrow$ */Ter/* ‘?’

However, as many other righthand outputs of the WFS, for example, *bhOy* ‘fear’, *betha* ‘pain’, etc. are words, our speaker-hearer may have the wrong impression that *Ter* has some meaning, or, it is also a word!

6 Conclusions

In this article, we have claimed that constructions known as ‘complex predicates’ are in fact words. They can be satisfactorily handled within morphology, and therefore, there is no need to treat them in syntax, or list each one of them in the lexicon, or put them in some special category like Word + after Kageyama (2001). The so-called vector verbs and their regular verb counterparts do not need to be listed separately either. However, depending on the speaker-hearer a number of these complex predicates must be part of his individual lexicon, which together with other words form adequate pairs that license WFSs like (19, 21) or (30, 31). Once such WFSs are part of his morphological module, the speaker-hearer can activate them, as needed, to output, understand or retrieve other complex predicates

Abbreviations

1: 1st person, 2: 2nd person, 3: 3rd person, ACC: Accusative, ADJ: Adjective, ALP : Argument Linking Principle, AI : Abstract Incorporation, ARG: Argument, COMPL: Complement, COMP: Complementizer, CONT: Continuous, DAT: Dative, EMPH: Emphatic element, GEN: Genitive, IMP: Imperative, INCL: Inclusive, IND: Indicative, INF: Infinitive, LOC: Locative, N: Noun, NEG: Negative, NOM: Nominative, PRS: Simple Present, PRP: Present Perfect, PFV: Perfective, PRES: Present, PRO: Pronoun, SG: Singular, TOP: Topic, V: Verb, WFS: Word Formation Strategy, WWM: Whole Word Morphology

Acknowledgements: I am grateful to late Professor Rajendra Singh whose thought provoking comments and suggestions had been very helpful. I must thank Professor Taro Kageyama, my Post-doc research supervisor in Japan and Professor Peter Hook for their helpful comments. I specially thank Professor Dr Dil Afroze Quader who read several versions of this paper and tightened up my prose. I also thank Professor Dr Shaila Sultana and Dr Nick Nicholas for reading the penultimate version and weeding out the last typos. For financial support, I gratefully acknowledge the beneficence of the Japan Society for the Promotion of Science (JSPS) and University of Dhaka, Bangladesh, while for logistic support, I am indebted to the Kwansei Gakuin University and the National Institute for Japanese Language and Linguistics, Tokyo, Japan. The usual disclaimers apply.

References

- Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.
- Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: The University of Chicago Press.
- Bhattacharja, Shishir. 2007. *Word formation in Bengali: A Whole Word Morphological description and its theoretical implications*. Munich: Lincom Europa.
- Booij, Geert. 2005. *The grammar of words*. Oxford: Oxford University Press.
- Butt, Miriam. 2003. The light verb jungle. *Harvard Working Papers in Linguistics* 9: Papers from the GSAS/Dudley House workshop on light verbs. 1–49.
- Butt, Miriam & Alexandros Tantos. 2004. Verbal Semantics via Petri Nets. *Proceedings of the LFG04 Conference*, University of Canterbury. <http://csli-publications.stanford.edu/>
- Chatterji, Sunitikumar. 1970 [1926]. *The origin and development of the Bengali language* (2 vols). London: George Allen and Unwin Ltd.
- Dasgupta, Probal. 1977. The internal grammar of compound verbs in Bangla. *Indian Linguistics* 38. 68–77.
- Ford, Alan, Rajendra Singh & Gita Martohardjono. 1997. *Pace Pāṇini: towards a word-based theory of morphology*. New York: Peter Lang.
- Grimshaw, Jane & Armin Mester. 1988. Light verbs and θ -marking. *Linguistic Inquiry* 19(2). 205–232.
- Hook, Peter E. 1974. *The compound verb in Hindi*. Michigan: Center for South and Southeast Asian Studies, University of Michigan.
- Kageyama, Taro. 1991. Light verb constructions and the syntax–Morphology interface. In H. Nakajima (ed.), *Current English linguistics in Japan*, 169–203. Berlin: Mouton de Gruyter.
- Kageyama, Taro. 2001. Word Plus: The interaction of words and phrase. In J. Weijer & T. Nishihara (eds.), *Issues in Japanese Phonology and Morphology*, 245–270. Berlin: Mouton de Gruyter.
- Lieber, Rochelle. 1983. Argument linking and compounds in English. *Linguistic Inquiry* 14(2). 251–285.
- Lieber, Rochelle. 1992. *Deconstructing morphology: Word formation in syntactic theory*. Chicago: The University of Chicago Press.
- Lieber, Rochelle. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press.
- Masica, Colin P. 1976. *Defining a linguistic area: South Asia*. Chicago: University of Chicago Press.
- Matsumoto, Yo. 1996. *Complex predicates in Japanese, a syntactic and semantic study of the notion ‘word’*. Stanford: CSLI Publications & Tokyo: Kurosio Publishers.
- Sadock, Jerald M. 1998. On the autonomy of compounding morphology. In S.G. Lapointe, D.K. Brentari & P.M. Farrell (eds.), *Morphology and its relation to phonology and syntax*, 161–187. Stanford: CSLI Publications.
- Saussure Ferdinand D. 1915 [1962]. *Cours de Linguistique Générale*. Paris: Payot.
- Singh, Rajendra. 1996. Explorations in phonology and morphology. In S.K. Verma (ed.), *Perspectives on language in society: Papers in memory of Prof. Srivastava*, 123–141. New Delhi: Kalina Publications.
- Singh, Rajendra. 2006. Whole word morphology. In K. Brown (ed.), *Encyclopedia of language and linguistics*, 2nd ed. Vol. 13, 578–579. Amsterdam etc.: Elsevier.
- Yu, A. C. L. 2007. Understanding near mergers: The case of morphological tone in Cantonese. *Phonology* 24(1). 187–214.