Jens Fleischhauer and Claudius Patrick Kihara

Light verbs and aspectual semi-auxiliaries in Gĩkũyũ

1 Introduction

Gĩkũyũ (E51[1]) is a Bantu language spoken in Central Kenya. The language shows many properties typical of Bantu languages: at the clausal level, it is a morphologically rich head-marking language (cf. Nichols, 1986). It can express up to two arguments by bound argument markers prefixed to the verbal stem. The language has a complex tense-aspect system mostly expressed by verbal pre- and suffixes. In addition, the language has a set of verbal suffixes – usually called 'extensions' in the Bantuist literature – expressing voice, i.e., causative, passive, and applicative.

Although the language is morphologically rich, it makes frequent use of periphrastic predicative constructions of the type 'Verb + Noun'. The noun can either be a lexical noun or an infinitive. Although infinitives show some properties of verbs (e.g., they license certain verbal affixes), they also behave like nouns and should be considered deverbal nouns (cf. Mugane, 2003). Infinitives are derived by adding the infinitive marker *kũ-* 'noun class 15' to a verbal stem. In predicative 'noun + verb' constructions, the verb forms a predicational unit with a noun which can be characterized as a complex predicate. Examples are shown in (1).[2] All these constructions have in common that the eventuality, i.e., the state or event denoted by the predicate, is introduced by the noun rather than the verb.[3] The complex predicate in (1a) is an instance of a prospective aspect construction and can be paraphrased as 'it is about to rain'. The raining event is contributed by the verbal noun *kũura* 'rain'. The example in (1b) is a phasal verb construction which denotes the beginning of an

1 'E51' is the so-called Guthrie number, a mainly geographically-based classification system of the Bantu languages proposed by Guthrie (1967) and updated by Maho (2003).

2 Complex predicates in Gĩkũyũ are not restricted to the types exemplified in (1) but due to reasons of space we restricted ourselves to these constructions. The Gĩkũyũ examples are derived from the introspection of the second author and have been validated with the institutions of other native speakers.

3 We use the term 'eventuality' in the sense of Bach (1986) as a cover term for both states and events.

Jens Fleischhauer, Heinrich-Heine-University, Düsseldorf, Germany
Claudius Patrick Kihara, Chuka University, Chuka, Kenya

https://doi.org/10.1515/9783112229330-004

event of singing. Finally, in (1c) the complex predicate 'beat phone' is paraphrased as 'make a call'.

(1) a. *Nĩ=kũ-r-end-a kũ-ur-a.*
FOC-17-PRS-want-FV 15-rain-FV
'It is about to rain.' (lit. It wants to rain.')[4]
b. *Ma-kĩ-rik-a kũ-in-a.*
2-NARR-start-FV 15-sing-FV
'And they started to sing/singing.'
c. *Mũ-tumia nĩ=a-ra-hũr-a thimo.*
1-woman FOC-1-PRS-beat-FV 9.phone
'The woman is making a call.' (lit. The woman is beating the phone)

At first sight, it looks as if the three constructions in (1) have a lot in common. However, we argue below that the construction in (1c) is grammatically closer to regular predicate-argument constructions like in (2) than to the two complex predicates in (1a) and (1b).

(2) *Mũ-tumia nĩ=a-ra-hũr-a ngui.*
1-woman FOC=1-PRS-beat-FV 9.dog
'The woman is beating the dog.'

The example in (1c) is an instance of a light verb construction (LVCs), the two examples in (1a) and (1b) represent semi-auxiliary constructions. On the basis of grammatical data, we argue that LVCs cannot be identified on the basis of their grammatical properties, whereas semi-auxiliary constructions show grammatical restrictions – compared to corresponding regular predicate-argument constructions – which are the result of an ongoing auxiliarization process. The data at hand do not present any evidence for light verbs undergoing a similar process of auxiliarization.

The paper is structured as follows: In section 2, we introduce the relevant grammatical background on Gĩkũyũ simplex verbs. For illustration purposes, we use the verb *hũra* 'beat' which has both a light use as in (1c) as well as a heavy use in (2). Section 3 introduces light verb constructions headed by *hũra*. The focus will be on the question of whether we can identify any grammatical differences between *hũra*'s light and heavy uses. The discussion reveals that the two systematically differ in only

4 There is some debate concerning the morphological as well as functional status of *nĩ*. Functionally, it is either analyzed as being an assertion marker and/or a focus marker (cf. Clements, 1984; Bergvall, 1987; Schwarz, 2003; Kihara, 2016, a.o.). Morphologically, it is either treated as a free particle, an affix or a clitic (cf. Barlow, 1960 and the above-mentioned literature). As the exact analysis of this element is not relevant for the current analysis, we adopt a uniform representation as a focus clitic.

one respect: the nominal element of an LVC cannot be pronominalized by a bound argument marker whereas the object argument of 'heavy' *hũra* can. In section 4, we turn to the prospective aspect construction illustrated in (1a). The central question to be asked is whether we can identify grammatical differences between *enda*'s use as a lexical full verb meaning 'want' and its occurrence within a prospective construction. This time, the comparison reveals stricter differences between the two uses. The phasal verb construction illustrated in (1b) is the subject of section 5. The constructions verbal element *rika* is also used as a lexical full verb meaning 'get into'. A comparison of the two uses reveals the same differences as identified between *enda*'s two uses. Section 6 presents a systematic discussion of the grammatical properties of the verbs and their nominal elements. These differences allow separating light verbs from semi-auxiliaries, as we term the relevant non-full verb uses of *enda* and *rika*. A cross-linguistic comparison of the Gĩkũyũ *hũra*-LVCs with light verb constructions in other languages is undertaken in section 7. The section is split into two subsections. In 7.1 we compare the Gĩkũyũ LVCs with similar LVCs in the genetically related language Swahili. Section 7.2 presents a brief comparison of some specific grammatical properties – direct object properties of the LVC's nominal element – between Gĩkũyũ and the Iranian language Persian. The paper ends with a conclusion in section 8.

2 Simplex verbs in Gĩkũyũ

Before we turn to a discussion of the different types of complex predicates mentioned in the introduction, we start with a brief discussion of simplex verbs. The reason is twofold: First, we need to distinguish complex predicates from simplex predicates – which is especially relevant for light verb constructions – and second, the discussion of simplex predicates allows introducing the relevant morphosyntactic background on Gĩkũyũ verbs. In the following discussion, we restrict ourselves to those aspects of verbal morphology which are relevant for our discussion and leave out other aspects, for instance reciprocal marking, as they are irrelevant for the discussion at hand.

Gĩkũyũ is head-marking at the clausal level (Nichols, 1986), i.e., the relation between the verb – the head of the clause – and its arguments – the dependent elements – is indicated by bound argument markers on the verb. The only obligatorily realized bound argument marker is the one realizing the subject argument. It can co-

occur with a coreferential independent subject referential phrase (RP).[5] However, a bound object marker cannot co-occur with an independent object RP within the same clause (Kihara, 2016, 55, Fleischhauer, 2023a, 166).[6] This is illustrated by the example in (3). The third person singular subject is realized by the bound argument marker *a-*, an independent pronoun or subject RP is not necessary. The only information given is that the subject referent belongs to noun class 1, e.g., the referent introduced by the noun *mũtumia* 'woman' in the preceding discourse (cf. example (2)). The object is realized by the bound argument marker *mĩ-* indicating an object argument of noun class 9. The optional independent RP *ngui* 'dog' is separated by comma intonation form the preceding sentence. Thus, it is right-dislocated and positionally realized outside of the clause (cf. Lambrecht, 2001 for the notion of dislocation).

(3) *Nĩ=a-ra-mĩ-hũr-a, ngui.*
FOC-1-PRS-9-beat-FV 9.dog
'S/he is beating it, the dog.'

The bound object marker is in complementary distribution to the reflexive prefix *ĩ-* (4). This marker is invariant and does not change depending on the subject argument's noun class.

(4) *Nĩ=a-ra-ĩ-hũr-a.*
FOC=1-PRS-REFL-beat-FV
'S/he is beating herself/himself.'

Gĩkũyũ has an elaborate tense and aspect system. In the temporal system, the language has a basic 'past – present – future' distinction which is accompanied by so-called remoteness distinctions in the past and future tense (cf. Cable, 2013). Whereas tense is realized by prefixes, aspect is realized by suffixes. The aspectual system is less elaborate than the tense system, however the language possesses three aspectual suffixes: *-ir* 'perfective', *-ag* 'imperfective' and *-ĩt* 'perfect'.

Finally, Gĩkũyũ has a rich inventory of voice affixes which operate on a verb's argument structure. Applicative and causative marking add an additional argument, passive and medium reduce one argument. Passive voice is expressed by adding *-o*

5 The notion of a referential phrase (RP) is taken form Role and Reference Grammar (RRG, Van Valin, 2008) and compensates the notion of a nominal phrase (NP) or determiner phrase (DP) used in other syntactic frameworks.

6 We adopt Van Valin's (2023, 67) analysis of bound argument markers in head-marking languages and assume that markers which can co-occur with a coreferential RP function as pronominal anaphora, whereas does which cannot function as pronouns.

as a final suffix, replacing the verb's final vowel (5a). Applicative (5b) and causative (5c) are realized by the (non-final) suffixes *-ĩr* and *-ithi* respectively.

(5) a. *Ngui nĩ=ĩ-ra-hũr-ir-wo nĩ mũ-tumia.*
9.dog FOC=9-PST-beat-PFV-PASS by 1-woman
'The dog was beaten by the woman.'
b. *Mũ-tumia nĩ=a-ra-hũr-ĩr-a nyina ngui.*
1-woman FOC=1-PRS-beat-APPL-FV 1.mother 9.dog
'The woman is beating the dog for the mother.'
c. *Mũ-thuri nĩ=a-ra-hũr-ithi-a mũ-tumia ngui.*
1-man FOC=1-PRS-beat-CAUS-FV 1-woman 9.dog
'The man caused the woman to beat the dog/ The man helped the woman the beat the dog.'

Although Gĩkũyũ licenses maximally one bound argument marker, this marker can refer to any argument. It can be one of the verb's direct arguments – as in (3) – or an applied argument, i.e., an additional argument introduced by an applicative marker, as in (6). Thus, Gĩkũyũ is a 'symmetric object language' as any non-subject argument can show direct object properties (Fleischhauer, 2023a, 166).[7]

(6) *Mũ-tumia nĩ=a-ra-mũ-hũr-ĩr-a ngui.*
1-woman FOC=1-PRS-1-beat-APPL-FV 9.dog
'The woman is beating the dog for him/her.'

In the remainder, we use to properties for identifying (direct) object arguments in Gĩkũyũ.[8] First, a direct object argument can be realized by a bound object marker (cf. (3) and (6)), and, second, a direct object argument can become the subject under passivization (5a). We will show below that these two properties are crucial for distinguishing LVCs from (semi-)auxiliary constructions.

3 Light verb constructions

As already mentioned in section 1, the verb *hũra* 'beat' is used both as a lexically full verb as well as a light verb. In its light use, *hũra* combines with different nouns to form a complex predicate. A few examples are listed in (7). Similar light verb

7 For the distinction between symmetric and asymmetric object languages, see, for instance, (e.g. Bresnan and Moshi, 1990).

8 The standard criteria for identifying object arguments in Bantu have been introduced by Hyman and Duranti (1982).

constructions are attested in, for instance, Swahili (cf. Krifka, 1995, 1401, Olejarnik, 2009, 2011; Martin, 2019) and Digo (Nicolle, 2023, 125). We present a brief comparison of Gĩkũyũ LVCs headed by *hũra* with Swahili *piga*-LVCs in section 7.1.

(7) *hũra nguo* 'to wash clothes' (lit. beat clothes), *hũra njĩra* 'to walk on a route/ path' (lit. beat path), *hũra mbica* 'to take a photo' (lit. beat photo), *hũra riboti* 'to report an incident' (lit. beat report), *hũra bathi* 'to iron clothes' (lit. beat iron box), *hũra thimu* 'to make a call' (lit. beat phone), *hũra kanua* 'to click' (lit. beat mouth), *hũra ndore* 'to fart loudly' (lit. beat fart), *hũra kiisi* 'to kiss' (lit. beat kiss)

Similar to its heavy use, light *hũra* requires the realization of two referential phrases: the subject argument and the phrasal element contributing the main predicational content. A telling example is the LVC *hũra thimũ* which literally means 'beat phone'. Although a literal interpretation is possible, the LVC is interpreted as 'to phone'. The noun refers to an artefact – a phone – which gives rise to event inference on the basis of the noun's meaning. Artefacts are designed for specific purposes and in the case of the noun 'phone' – and its correspondent in Gĩkũyũ – the referent is designed to call someone. Therefore, the noun licenses inference to an event of calling someone/phoning. At this stage, we do not go further into the semantics of the light verb construction but we would like to highlight that the nominal element contributes to the complex predicate's meaning. Crucially, the noun does not refer to an object but contributes the denoted eventuality (for further discussion of this issue see the discussion of non-eventive nouns in the introduction to the volume). Although the nominal element forms a semantically complex predicate with *hũra* 'beat', the light verb constructions does not show any grammatical differences to the regular predicate-argument construction discussed in the previous section.

As the examples in (8) show, it is possible to have a bound argument marker coreferential with the noun *thimũ*. Crucially, the noun has to be right-dislocated as discussed above. The form of the noun does not change depending on number but the number distinction is reflected in the choice of the bound argument marker. As the contrast between (8a) and (8b) indicates, nominal number affects the interpretation. The plural noun introduces a pluralic event (many calls), whereas the singular noun allows either for a habitual interpretation or an excessive one.

(8) a. *Mũ-tumia nĩ=a-ra-i-hũr-a, thimũ.*
1.woman FOC=1-PRS-10-beat-FV 10.phone
'The woman is making many calls./ The woman is calling many numbers.'

b. *Nĩ=a-ra-mĩ-hũr-a, thimũ.*
FOC=1-PRS-9-beat-FV 9.phone
'S/he is (really) calling the phone regularly./ He is calling extensively.'

Since the light verb forms a tight semantic unit with the nominal element, the noun cannot be replaced by a reflexive marker. Reflexive marking presupposes applicative marking, the reflexive marker instantiates the applied argument. But this is the only grammatical restriction on LVCs observed in our study.

(9) *Mũ-tumia nĩ=a-ra-ĩ-hũr-ĩr-a thimũ.*
1.woman FOC=1-PRS-REFL-beat-APPL-FV 9.phone
'The woman is calling herself/ dialing her own number. / The woman is calling other people herself.'

As Gĩkũyũ is a symmetrical object language, a bound object marker can also be licensed by an applied argument. This is shown in (10) which illustrates the combination of the LVC with the applicative marker *-ĩr*. The applicative marker introduces the recipient of the call, i.e., the person who is called, or a beneficient of the action. It could be realized with an independent RP or, as in (10), by a bound object marker.

(10) *Mũ-tumia nĩ=a-ra-mũ-hũr-ĩr-a thimũ.*
1.woman FOC=1-PRS-1-beat-APPL-FV 9.phone
'The woman is calling him/her. / The woman is calling on his/her behalf.'

The LVC *hũra thimũ* 'to call' only allows the applied object to become the subject under passivization (cf. 11). Thus, in this specific case does passivization presupposes applicative marking.

(11) *Mũ-tumia nĩ=a-ra-hũr-ir-wo thimũ nĩ nyina.*
1.woman FOC=1-PRS-beat-APPL-PASS 9.phone by 1.mother
'The woman is being called by her mother.'

We also regularly find a possibly partially lexicalized usage, in which the recipient of the call can be realized by a bound argument marker, while the noun *thimũ* 'phone' is not realized (12).

(12) *Nĩ=n-gũ-kũ-hũr-ĩr-a.*
FOC=1-FUT-1-beat-APPL-FV
'I will call you.'

However, *hũra thimũ* is not representative for all *hũra*-LVCs as the examples in (13) show. The examples illustrate passivization of the LVC *hũra baĩni* 'to fine' (lit. beat fine). In (13a), an active use of the LVC is shown. The sentence contains two non-subject RPs: *baĩni* 'fine' and *mũndũ* 'person'. Thus, the syntactic valency is – compared to *hũra*'s use as a lexical full verb – increased by one. Both postverbal RPs can become subject under passivization, as the examples in (13b) and (13c) show.

(13) a. *I-goti nĩ=ri-a-hũr-ir-e mũ-ndũ baĩni.*
5-court FOC=5-PST-beat-PFV-FV 1-person 9.fine
'The court fined the person.'
b. *Mũ-ndũ nĩ=a-a-hũr-ir-wo baĩni nĩ i-goti.*
1-person FOC=1-PST-beat-PFV-PASS 9.fine by 5-court
'The person was fined by the court.'
c. *Baĩni y-a-hũr-ir-wo mũ-ndũ nĩ i-goti.*
9.fine 9-PST-beat-PFV-PASS 1-person by 5-court
'The fine was charged on the man by the court.'

Although *baĩni* has one object property, it cannot be pronominalized by a bound argument marker. *Mũndũ*, on the other hand, can be pronominalized as (14) shows.

(14) *I-goti nĩ=ri-a-mũ-hũr-ir-e baĩni.*
5-court FOC=5-PST-1-beat-PFV-FV 9.fine
'The court fined him/her.'

This brief comparison of the two LVCs revealed LVC-specific restrictions concerning passivization. However, it also proofs that the nominal element shows (at least some) direct object properties. In addition, we observe that the LVCs syntactic valency does not (necessarily) coincide with *hũra*'s valency. As a lexical full verb, *hũra* takes two arguments – one subject and one object argument (e.g. (2)) – but there are two non-subject RPs in (13) without the verb undergoing any valency increasing operation. Although the question how the argument structure of an LVC is created is a highly relevant grammatical issue, we will not address it in the current paper as our focus is on other grammatical properties.

With respect to the other instances of verbal morphology – tense, aspect and causative marking – light *hũra* does not differ from its heavy use. We will not illustrate this for the complete paradigm of tense and aspect forms but restrict ourselves to the one example in (15a). In contrast to the preceding examples, the verb bears future tense marking. The example in (15b) shows the morphological causative *-ithi*[9] on the light verb *hũra*. The causative marker introduces a new argument – the causer –

9 The causative suffix *-ithi* is bimorphic as evidenced by the fact that it is not necessarily realized adjacently, i.e., the perfective aspect suffix is positioned between *ith* and *i*. Diachronically, *-ithi* is a combination of causative *-ith* and the transitivizer *-i* (Good, 2005, 10–11, Good, 2007, 208–209. Synchronically, *-i* does not make an independent semantic contribution in examples like (10). However, it functions as a transitivizer – also called 'short causative' – in other contexts. In examples like (15b), we simply gloss the discontinuous part of the morpheme as 'DC' – discontinuous element – and conceive it (synchronically) as being a part of final part of the bimorphic ('long') causative *-ithi*.

and demotes the previous subject argument to direct object status. In (15), the agent – the one calling – is expressed by a bound object marker.

(15) a. *Mũ-tumia nĩ=a-kũ-hũr-a thimũ.*
1-woman FOC=1-FUT-beat-FV 9.phone
'The woman will call.'

b. *A-tũ-hũr-ith-ir-i-e thimũ ci-a tũhũ.*
1-2-beat-CAUS-PFV-DC-FV 10.phone 10-ASSOC useless
'S/He had made us make useless calls.'

To summarize: Although the empirical basis is rather limited, we do not find any general grammatical differences between a verb's heavy use and its use as a light verb. The most promising property – which requires further empirical validation on the basis of a broader data sample – is the that the nominal element does not always show direct object properties.

4 Prospective constructions

Gĩkũyũ – like many other languages (cf. Heine, 1994) – uses periphrastic constructions of the type 'verb + noun' for the expression of prospective aspect.[10] Prospective aspect relates the subject referent's current state to a subsequent eventuality (e.g., Comrie, 1976, 52). In (16), for instance, the referent of *ngaari* 'car' is in a state preceding a crash. The actual realization of this eventuality is – as discussed above – not entailed.

Gĩkũyũ expresses prospective aspect by a combination of the verb *enda* 'want' and an infinitive. Infinitives are – as already mentioned above – derived by adding a marker of noun class 15 to a verbal stem. The verbal stem minimally consists of the lexical root – *-gũ-* – in (16) and the final vowel *-a*. However, the infinitive can bear further verbal markers like the bound object marker or voice markers. The marker for class 15 is a regular bound subject marker and the infinitives behave in all relevant aspects like nouns (cf. Mugane, 2003). The prospective construction in (16) resembles the LVCs discussed in the previous section as the verb and its non-subject argument form a predicational unit.[11]

10 For an analysis of (aspectual) auxiliary + verb constructions in the three Bantu languages Swahili, Rangi, and SiSwati, see Gibson and Marten (2016).

11 The German prospective '*stehen vor* NP'-constructions (lit. stand in front of NP) similarly resembles LVCs within the languages. The resemblance is even greater than in Gĩkũyũ as *stehen* 'stand' is also attested as a light verb in other LVCs (cf. Fleischhauer, 2023c). However, desiderative verbs seem

(16) *Ngaari ĩ-kũ-end-ag-a kũ-gũ-a.*
9.car 9-PST-want-IMPF-FV 15-fall-FV
'The car was about to crash.' (lit. The car wanted to fall.)

The verb *enda* is also used as a regular full verb expressing "a desire that the complement proposition be realized" (Noonan, 2007, 132). In its use as a desiderative predicate – illustrated in (17) –, *enda* translates as 'want'. We refer to this use as the 'desiderative use', resp. 'desiderative interpretation' of *enda*. As a full verb, *enda* can take a verbal complement. *Ũine*, in (17), is a verb bearing pronominal subject marking. The verb's final vowel is *-e* which is, among other things, used in subjunctive contexts. However, the final vowel cannot be analyzed as a subjunctive marker as it occurs in other contexts as well, i.e., perfective aspect forces the final vowel to be *-e* rather than *-a*.

(17) *Nĩ=nd-ĩr-end-a ũ-in-e.*
FOC-1SG-PRS-want-FV 2SG-sing-FV
'I want you to sing (right now).'

(Wittke, 2015, 80)

The example in (17) has unambiguously a desiderative interpretation: the speaker wants the addressee to sing. Other sentences, like the one in (18), are ambiguous between a desiderative and a prospective interpretation. The sentence can either mean that the subject referent wants to go to work or that the subject referent is about to go to work. In both cases, it does not follow that the subject referent will go to work. In the first case, it is just the subject referents wish to go to work. Whereas in the second case, it might be expected that – if nothing intervenes – the subject referent is close to going to work. It is not contradictory to say: I was about going to work but I did not go since I broke my leg while putting on my shoes.

(18) *Tũ-r-end-a gũ-thiĩ wĩra.*
2-PRS-want-FV 15-go 9.work
(i.) 'They want to go to work.'
(ii.) 'They are about going to work.'

Evidence that we are dealing an aspectual prospective construction[12] rather than a periphrastic (immediate) future tense construction is that fact that the finite verb

to be more frequently attested as the source of a prospective construction than spatial expressions (cf. Heine, 1994, 44).

12 We adopt Dik's (1997, 221) division between 'aspect proper' and 'relational aspect'. 'Aspect proper' covers perfective as well as imperfective aspect, whereas 'relational aspect' comprises the two notions of 'perfect' and 'prospective'.

can be marked for past tense to refer to a prospective situation in the past. In (19), it is expressed that at some point in the past, it was about to rain. Note that the noun class marker for class 15 and 17 are homophonous; whereas noun class 15 realized nominalized infinitives, 17 is a locational noun class deriving nouns referring to a location 'here'.

(19) *Nĩ=kũ-r-end-ag-a kũ-ur-a ira.*
FOC=17-PST-want-IMPF-FV 15-rain-FV yesterday
'It was about to rain yesterday.'

The sentences in (16) and (19) are unambiguous examples of a prospective interpretation of *enda* as an inanimate subject referent cannot have any wishes. Whereas (19) has a locational subject argument – the referent of the bound subject marker is interpreted as 'here' (noun class 17) –, the one in (16) has a concrete subject referent. Thus, a desiderative interpretation for *enda* is definitely ruled out in these cases (exceptions are fictional contexts like, for instance, in the Pixar movie *Cars*).

As a verbal component of the prospective construction, *enda* exhibits stronger restrictions regarding its complement than in its use as a full verb. In the prospective construction, the subject argument of *enda* has to be coreferential with the (unrealized) subject of its infinitival complement (cf. Fleischhauer and Gamerschlag, 2019, 146). This is different for the desiderative use of *enda* as the example in (17) shows. The subject of *enda* is different from the subject of its complement clause. A consequence of the coreferentiality restriction is that in its prospective function, *enda* only combines with infinitives but not – unlike *enda*'s desiderative interpretation (cf. (17)) – with finite verbs. However, desiderative *enda* does not obligatorily take complement clauses but can also take RP complements (20). As (20b) shows, the object of the verb can be realized by a bound object marker (the RP *ngaari ĩno* 'this car' is left-dislocated and separated by an intonation break from the remaining clause).

(20) a. *Ndĩ-r-end-a ngaari ĩ-no.*
1SG-PRS-want-FV 9.car 9-DEM
'I want this car.'
b. *Ngaari ĩ-no, nĩ=ndĩ-ra-mĩ-end-a.*
9.car 9-DEM FOC=1SG-PRS-9-want-FV
'This car, I want it.'

Pronominalization indicates that the non-subject argument bears direct object properties. It is therefore not surprising that passivization is possible as well (21). In this case, passivization does not reduce syntactic valency but affects the semantic role of the subject argument. Whereas in the active sentence the subject of *enda* is the one

having a wish, passivization makes the subject argument the logical object of the wishing, i.e., the one of which it is wished that s/he is singing.

(21) *Ndĩ-r-end-wo ny-in-e.*
1SG-PRS-want-PASS 1SG-sing-FV
'I am expected to sing.' (lit. I am wanted to sing)

Passivization is even possible if desiderative *enda* takes an infinitival complement (22). The bound object marker *kw-* is coreferential with the right-dislocated infinitival clause *kũina* 'singing'.

(22) *Kũ-in-a n-di-kw-end-ĩt-e.*
15-sing-FV 1-NEG-15-want-PERF-FV
'Singing, I don't want/like it.'

Whereas *nyine* in (21) and *kũina* in (22) qualify as clausal complements of (desiderative) *enda*, *kũgũa* 'to fall' – in the prospective construction in (19) – does not . Noonan gives the following characterization of 'complementation': "By complementation, we mean the syntactic situation that arises when a notional sentence or predication is an argument of a predicate. For our purposes, a predication can be viewed as an argument of a predicate if it functions as the subject or object of that predicate" (Noonan, 2007, 52). *Kũgũa* 'to fall' (in (16)) does not bear any object properties, especially it cannot be pronominalized by a bound object marker (23). Thus, the combination of *enda* and *kũgũa* is not an instance of complementation.

(23) **Kũ-gũ-a, ngaari ĩ-gũ-kũ-end-ag-a.*
15-fall-FV 9.car 9-PST-15-want-IMPF-FV
Intended: 'The crash, the car was about to do it.'

Furthermore, the prospective interpretation is lost under passivization as the example in (24) shows. *Ngaari* 'car' becomes subject of the sentence, *gũa* 'fall' is not realized as an infinitive but bears a bound subject marker for noun class 9 which is coreferential with *ngaari*. Crucially, the sentence does not have a prospective interpretation but is interpreted as 'the subject is expected to do' which is that the same interpretation we observed for the desiderative interpretation of *enda* in (21). We conclude that the prospective construction is incompatible with passive voice in Gĩkũyũ.

(24) *Ngaari ĩ-kũ-end-ag-wo ĩ-gũ-e.*
9.car 9-PST-want-IMPF-PASS 9-fall-FV
'The car was expected/required to crash.'

When it comes to other voice morphology – causative, applicative – we observe that neither desiderative nor prospective *enda* license these markers. Therefore, we cannot say that the two uses of *enda* differ with respect to this property.

The prospective construction shows an interesting interaction with aspect. As already mentioned, prospective aspect does not entail the actual realization of the prospective eventuality. Taking (25) as an example, it is possible that it actually rained. Thus, can be continued with *na nĩ=kũraurire* 'and (indeed) it rained'.

(25) *Ira nĩ=kũ-r-end-ag-a kũ-ur-a.*
yesterday FOC=15-PST-want-IMPF-FV 15-rain-FV
'Yesterday, it was about to rain.'

If we switch aspect from the imperfective to perfective (26a) or perfect (26b), adding *na nĩ=kũraurire* 'and it rained' results in a contradiction. Thus, perfect(ive) enforces a non-realizational interpretation of the construction.

(26) a. *Ira nĩ=kũ-r-end-ir-e kũ-ur-a.*
yesterday FOC=15-PST-want-PFV-FV 15-rain-FV
'Yesterday, it was about to rain.'
b. *Ira nĩ=kũ-r-end-ĩt-e kũ-ur-a.*
yesterday FOC=15-PST-want-PERF-FV 15-rain-FV
'Yesterday, it was about to rain.'

More specifically, in the context of past perfect(ive), the prospective construction has an avertive interpretation. Kuteva et al. (2019, 858) dub a linguistic construction 'avertive' if is refers to "a verb situation which was on the verge of taking place but did not take place" (similarly Kuteva, 2001, 78). In Gĩkũyũ, an avertive interpretation results from a specific combination of tense (past) and aspect (perfect(ive)) and is, interestingly, not restricted to the prospective construction. An avertive interpretation also arises with desiderative *enda* as (27) shows. The examples show a conjunction of two main clauses. The first clause states that the subject referent wanted to buy a car, whereas the second conjunct expresses that s/he finally did so. The conjunction of the two sentences is only licit if *enda* bears imperfective aspect marking (27a). If *enda* is marked for perfect aspect, adding the sentence *na nĩ=ndagũrire* 'and I bought it' is odd (27b). Thus, realization of the eventuality expressed by the infinitive serving as the complement of *enda* – irrespective whether the verb is used as a verb of desire or within a prospective construction – is prohibited in a past perfect(ive) context.

(27) a. *Nd-e-end-ag-a kũ-gũr-a ngaari na nĩ=nd-a-gũr-ir-e.*
1-PST-want-IMPF-FV 15-buy-FV 9.car with FOC=1-PST-buy-PFV-FV
'I wanted to buy a car and I bought one.'

b. *Nd-e-end-ĩt-e kũ-gũr-a ngaari #na nĩ=nd-a-gũr-ir-e.*
1-PST-want-PERF-FV 15-buy-FV 9.car with FOC=1-PST-buy-PFV-FV
'I wanted to buy a car and I bought it.'

The avertive interpretation is not a specific property of *enda*'s use as a prospective (semi-)auxiliary but it attested with desiderative *enda* as well. Therefore, we do not assume that prospective *enda* shows specific aspectual restrictions which are not attested with desiderative *enda* as well. However, the two differ in a number of respects as we have shown in this section.

Although the two uses of *enda* clearly differ with respect to their grammatical behavior, we retained the glossing 'want' for both uses. The reasons are simply that we assume, first, *enda* has not developed into a full-fledged auxiliary yet and, second, it is not obvious what the (lexical) meaning of prospective *enda* is. However, we will adopt a different strategy for the verb *rika* to which we turn in the next section.

5 Phasal verbs

Longacre defines 'phasals' as an expression which "indicates whether an action is beginning, continuing, or ending" (Longacre, 1976, 238). Phasal verbs are a specific subtype of phasals and lexically encode whether an eventuality is conceived of as, for instance, beginning. In the literature on phasals – sometimes also referred to as 'phasal aspect' – three subtypes are commonly distinguished: 'inceptive', 'durative' and 'terminative', referring to the beginning, continuation and ending of an eventuality (e.g., Longacre, 1976, 238; Noonan, 2007, 139–140; Croft, 2022, 559–560).

In this section, we focus on the Gĩkũyũ verb *rika* which – as a lexical full verb – means 'to get into' (28a).[13] However, the verb is also used for the expression of inceptive aspect focusing on the beginning of an eventuality (28b). As we can identify two clear semantic contributions in the two different uses, we provide to different glosses for the verb 'get.into' and 'begin', respectively.

(28) a. *Nĩ=ma-a-rik-ir-e mũ-taro.*
FOC=2PL-RMPST-get.into-PFV-FV 3-trench
'They got into the trench.'
b. *Ma-a-rik-ir-e kũ-in-a orĩo.*
2PL-RMPST-begin-FV 15-sing-FV immediately
'They began to sing immediately.'

13 *Rika* is not the only phrasal verb in Gĩkũyũ; other verbs in this group include, *anjia/ambia* 'start/begin', and *rĩkia* 'complete/finish', among others.

Benson (1964, 386) argues in favor of a polysemy analysis of *rika* and attributes the two distinct meanings to the same verbal stem. We depart from this analysis and propose that inceptive *rika* is a semi-auxiliary. Thus, it is still on its way of becoming a grammaticalized auxiliary. However, inceptive *rika* differs from the lexical full verb in a number of grammatical respects very similar to the differences observed between the two uses of *enda* in the preceding section.

Grammatical differences between the two uses of *rika* are again found with respect to the status of the construction's obligatory nominal element. As a lexical full verb, *rika*'s non-subject argument bears direct object properties. It can be pronominalized by a bound argument marker (29a) and can also become subject under passivization (29b). Different from *enda*, *rika* does not take sentential complements.

(29) a. *Nĩ=ma-a-ũ-rik-ir-e.*
FOC=3PL-RMPST-3-get.into-PFV-FV
'They got into it (e.g., the trench).'
b. *Nĩ=ũ-a-rik-ir-wo.*
FOC=3-RMPST-get.into-PFV-PASS
'It was dipped into/ It was got into.'

In the phasal verb construction, the infinitive – *kũina* 'to sing' in (28a) – does not show any object properties. It can neither be realized by a bound argument marker nor can it become the subject of a passive sentence. However, the infinitive can take an object argument (30a). This object can also be realized by a bound object marker (30b) and it becomes subject under passivization (30c). Although passive is marked on the infinitive, the new subject is expressed by a bound subject marker on *rika*.

(30) a. *Nĩ=ma-a-rik-ir-e kũ-in-a rw-ĩmbo.*
FOC=3PL-RMPST-begin-PFV-FV 15-sing-FV 11-song
'They began to sing a song.'
b. *Nĩ=ma-a-rik-ir-e kũ-rũ-in-a.*
FOC=3PL-RMPST-begin-PFV-FV 15-11-sing-FV
'They began to sing it.'
c. *Nĩ=rũ-ra-rik-ir-e kũ-in-wo nĩ mũ-tumia.*
FOC=11-RMPST-begin-PFV-FV 15-sing-PASS by 1-woman
'It was begun to be sung by the woman.'

There is an interesting contrast between phasal verb constructions and prospective aspect constructions when it comes to passivization. Whereas the phasal verb construction realizes the passive morpheme at the infinitive, it is realized on the finite verb in the case of *enda* (31). However, as already shown in the preceding section, the prospective construction is incompatible with passive marking. The addition of the

passive marker results in the desiderative interpretation of *enda* (31b). Interestingly, the desiderative interpretation also obtains if the passive marker is realized on the infinitive (31c) or together on the infinitive and the finite verb (31d).

(31) a. *Mw-aki ũ-r-end-a kũ-hĩhi-a rũkũ.*
3-fire 3-PRS-want-FV 15-burn-FV 11-firewood
'The fire is about to burn the firewood.'
b. *Mw-aki ũ-r-end-wo ũ-hĩhi-e rũkũ.*
3-fire 3-PRS-want-PASS 3-burn-FV 11-firewood
'The fire is expected to burn the firewood.'
c. *Rũkũ rũ-r-end-a rũ-hĩhi-o nĩ mw-aki.*
11.firewood 11-PRS-want-FV 11-burn-PASS by 3-fire
'The firewood is expected to be burned by the fire.'
d. *Rũkũ rũ-r-end-wo rũ-hĩhi-o nĩ mw-aki.*
11.firewood 11-PRS-want-PASS 11-burn-PASS by 3-fire
'The firewood is expected to be burned by the fire.'

The crucial difference between the two constructions is that whereas the prospective *enda*-construction does not allow passivization at all, the phasal *rika*-construction only allows passivization of the embedded infinitive. Thus, the phasal verb itself cannot be passivized. Similar restrictions obtain with respect to causativization and applicativization. In its full verb use, *rika* can, for instance, causativized as shown in (32a). The phasal verb construction, on the other hand, cannot causativized itself. Like passivization, causativization is restricted to the embedded infinitive (32b). The causer is realized by a bound subject marker on *rika* showing that *rika kũinithia* 'start to conduct' has a joint argument structure.

(32) a. *Ma-ma-rik-i-ir-i-e mũ-taro.*
2-2-get.into-CAUS-PFV-DC-FV 3-trench
'They made others to get into the trench.'
b. *Ma-a-rik-ir-e kũ-in-ithi-a rw-imbo.*
2-RMPST-begin-PFV-FV 15-sing-CAUS-FV 11-song
'They began conducting the song.'

The discussion has shown that *rika* and *enda* exhibit very similar restrictions in their semi-auxiliary usage, which we will discuss in detail in the next section.

6 Comparison of the three constructions

The preceding sections focused on a comparison of the three verbs *hũra*, *enda* and *rika* with their respective lexical full verb uses. In this section, we compare the three verbs with respect to their non-full verb uses. We already introduced the label 'semi-auxiliary' for *enda*'s and *rika*'s uses within the prospective aspect, resp. phasal verb construction, whereas we reserved the term 'light verb' for *hũra*'s use in constructions like *hũra thimũ* 'to call'. This distinction between the three verbs is justified on the basis of their shared grammatical properties. The relevant properties can be separated into: (i) properties pertaining to the complex predicate's nominal element and (ii) properties pertaining to the complex predicate's verbal element.

Starting with the nominal elements, the discussion revealed three relevant grammatical properties. First, can the nominal element be pronominalized by a bound argument marker? Second, can the nominal element become subject under passivization? And, third, is the nominal element restricted to a particular nominal class? The results are summarized in Table 1.

Tab. 1: Distribution of the noun-related properties among the three complex predicates.

	***hũra*-LVCs**	***enda*-prospectives**	***rika*-phasals**
(i) pronominalization	yes*	no	no
(ii) passivization	yes*	no	no
(iii) restriction to a particular NC	no	yes	yes

Light *hũra* differs from the two semi-auxiliaries with respect to all three properties. The two first features are related to presence versus absence of direct object properties. As we have shown, the nominal elements of the two semi-auxiliary constructions lack any direct object properties. *Hũra*-LVCs, on the other hand, show a non-uniform behavior. Some LVCs do not allow their nominal element to become subject under passivization (*hũra thimũ*), others do (*hũra baĩni*). Similarly, some LVCs allow their nominal element to be pronominalized by a bound object marker (*hũra thimũ*), others do not (*hũra baĩni*). The interesting question is why the two LVCs show a complementary distribution of these two properties. At the current stage of research, we cannot answer this question but require a larger data set of Gĩkũyũ LVCs to derive sound generalization. However, it is obvious that the nominal elements of the *hũra*-LVC show at least some direct object properties. For this reason, we concluded that the nominal elements can be pronominalized and can become

the subject under passivization but – which we indicate by the asterisk in Table 1 – not necessarily in every single instance.

The three constructions diverge on one further relevant property, namely the restriction of the nominal element to noun class 15. Noun class 15 contains, among some other nouns, infinitives. Infinitives are deverbal nouns which can bear verbal morphology, e.g., a bound object marker as well as passive marking. In difference to other nouns, infinitives do not participate in the nominal number system. Thus, there is no number opposition in case of these nouns. The LVCs under discussion combine with nouns of various noun classes which are still capable of number marking. In case of *thimũ* 'phone', plural is not evident on the noun but is only visible in the agreement system. The quantity expression *-ingĩ* takes the agreement marker for noun class 10 *ny-*, the plural for class 9. The one for noun class 9 – the singular class of *thimũ* – would have been *ĩ-*. The LVCs interpretation is that the woman is not making just one call but many calls (33).

(33) *Mũ-tumia nĩ=a-ra-hũr-a thimũ ny-ingĩ.*
1-woman FOC-1-PRS-beat-FV 10.phone 10-much
'The woman is making many calls.'

There is an additional difference between the nouns occurring in the LVCs and those occurring in the semi-auxiliary constructions. Since the semi-auxiliary constructions take an infinitive, the nominal element is of deverbal origin and therefore denotes an eventuality. The nouns occurring within the light verb constructions are not derived and do not denote any eventuality. In contrast, *thimũ* 'phone' refers to an artifact. The occurrence of artifact nouns within light verb constructions is attested in other languages (e.g., German and Persian) as well (cf. Fleischhauer, 2023b for German and Fleischhauer, 2021 for Persian; but also see the references in the introduction to this volume). Artifacts are objects that have a specific function. This function is a meaning component of the corresponding nouns and allows for the inference of an event in which the artifact, typically as an instrument, is embedded (e.g., Nichols, 2008; Grimm and Levin, 2017; Levin et al., 2019). In the case of *thimũ* 'phone', the inferred event is a calling event, which is the (at least original) primary function of phones.[14] However, due to reasons of space we cannot go into this issue more deeply but we leave the details open for future work.

Turning now to the grammatical properties of the verbs, we identified four relevant properties. First, does the verb license a bound non-subject marker?, Second, does the verb license passive marking? And, third, does the verb license other voice

14 The inferred event in which the artifact is involved as an instrument corresponds to the event associated with Pustejovksky's (1991; 1995) Telic Role.

morphology, i.e., causative and applicative marking? If adding one of the mentioned affixes results in a meaning shift – e.g., from prospective to desiderative – , the verb does not bear the respective property. The results for the verb-related properties are summarized in Table 2:

Tab. 2: Distribution of the verb-related properties among the three complex predicates.

	***hũra*-LVCs**	***enda*-prospectives**	***rika*-phasals**
(i) bound non-subject marker	yes	no	no
(ii) passive morphology	yes	no	no
(iii) other voice morphology	yes	—	no

The light verb *hũra* differs in its grammatical properties from the two semi-auxiliaries. Whereas the semi-auxiliaries are syntactically intransitive, *hũra* is not. This is evidenced by the fact that *hũra* generally licenses bound non-subject markers as well as passive morphology but the semi-auxiliaries do not. In addition, the semi-auxiliaries do not license any argument increasing operations (e.g., causativization, applicativization), whereas *hũra* does. As we observe the same restriction for desiderative *enda*, we cannot attribute this restriction to prospective *enda*'s status as a semi-auxiliary which we indicate by '—' in the respective cell in Table 2.

The mentioned differences might result from the different functions of the constructions. Both the prospective construction as well as the phasal verb construction operate on a (nominalized) verbal concept and can be subsumed under the umbrella term 'periphrastic aspect construction'.[15] Light verb constructions, on the other hand, often fill lexical gaps or extend the lexicon (e.g., Wichmann and Wohlgemuth, 2008, Olejarnik, 2011, 141). Thus, they serve a lexical rather than grammatical function. For instance, Gĩkũyũ does not have a simplex verb meaning 'to call'. The lack of this predicate is compensated by the use of the LVC *hũra thimũ*. The semi-auxiliaries do not function as lexical (complex) predicates but as periphrastic aspectual constructions.

There is a debate whether light verbs are the outcome of a desemantization process or not (e.g., Hopper and Traugott, 1993, 112, Allerton, 2003, 7, Butt and Lahiri, 2013). Sometimes it is even proposed that light verb represent an intermediarey step in the auxiliarization process. Since our analysis is based on limited data, we cannot make a definitive contribution to this debate. However, it seems that we can draw two tentative conclusions from the data. First, both the nominal element as well as the

15 We are using the term 'aspect' here in a very broad sense covering both different types of grammatical aspect (cf. section 4) as well 'phasal aspect'.

verbal element within a semi-auxiliary construction lose some of their independent grammatical properties, i.e., both the verb and the noun show restrictions compared to the corresponding elements in a regular predicate-argument construction. Second, LVCs do not show any general grammatical restrictions compared to corresponding predicate-argument constructions. Whereas semi-auxiliary constructions can be distinguished from corresponding predicate-argument constructions on the basis of their grammatical properties, LVCs differ from corresponding predicate-argument constructions just semantically. The main criterion to distinguish LVCs from regular predicate-argument constructions is: which element –– the noun in the case of the LVC or the finite verb in the case of a regular predicate-argument construction – provides the main predicational content. Thus, we propose that the distinction between LVCs, on the one hand, and regular predicate-argument constructions, on the other hand, is essentially in terms of semantics rather than grammar (a similar claim is presented in Fleischhauer, 2021 for Persian LVCs). Obviously, this is different for the distinction between regular predicate-argument constructions and LVCs, on the one hand, and semi-auxiliary construction on the other. There are indeed grammatical differences allowing to distinguish semi-auxiliary constructions from the mentioned other two predicative construction types.

7 Gĩkũyũ LVCs from a cross-linguistic perspective

In this section, we compare Gĩkũyũ LVCs headed by *hũra* with LVCs in other languages. We compare it to Swahili LVC headed by *piga* 'beat' in section 7.1 and turn to a comparison with LVCs in the Indo-European language Persian in section 7.2.

7.1 Comparison with Swahili LVCs

Similarly to Gĩkũyũ, Swahili possesses LVCs headed by a light verb meaning 'beat, hit' which is *piga*. Olejarnik (2011, 146) states that Swahili LVCs license voice marking. However, concerning the nominal element within such constructions, Krifka (1995, 1401) states that it cannot be a direct object since it can neither be pronominalized nor can it become subject under passivization.[16] Martin (2019, 29) mentions that the

16 Krifka (1995, 1401) proposes that the nominal element of an LVC is (pseudo) incorporated into the verb. Similar claims have been put forward for Persian LVCs (e.g., Saeddi, 2016) but see Fleischhauer (2021) as well as Fleischhauer and Neisani (2020) for arguments against this claim based on Persian language data. These arguments, however, similarly apply to the Bantu languages as well.

nominal element of the LVC cannot be taken up by a bound argument marker. This is shown for the LVC *piga faini* 'to fine' (lit. beat fine) in (34a). However, the noun *raia* 'citizen' – denoting the individual being fined – can be pronominalized by a bound object marker (34b). Crucially, there is no applicative marker on the verb which shows that *raia* is licensed by the complex predicate but not by voice morphology.

(34) a. **Mahakama hiyo i-li-i-pig-a faini raia.*
9.court 9.DEM 9-PST-9-beat-FV 9.fine 1.citizen
'That court fined the citizen.'
b. *Mahakama hiyo i-li-m-pig-a faini raia.*
9.court 9.DEM 9-PST-1-beat-FV 9.fine 1.citizen
'That court fined the citizen.'

(Martin, 2019, 29)

Swahili differs in one relevant syntactic property from Gĩkũyũ. Whereas Gĩkũyũ does not allow the co-occurrence of a bound non-subject marker and coreferential RP within the same sentence, Swahili does not show a similar restriction, as evidenced in (34b). Crucially, *faini* has not to be right-dislocated but can co-occur with a coreferential bound object marker within the same clause. If Martin is right, Swahili differs from Gĩkũyũ concerning the possibility to pronominalize the nominal element of an LVC.

The two languages show similarities when it comes to passivization As the examples in (35) show, the nominal element of the LVC *piga mayowe* 'to shout/scream' can become the subject under passivization.

(35) a. *Mama huyo a-li-pig-a mayowe ku-omba msaada.*
1.mother 1.DEM 1-PST-beat-FV 5.shout 15-beg 3.help
'That mother shouted begging for help.'
b. *Mayowe ya-li-pig-wa kwa sauti kubwa.*
5.shout 5-PST-beat-PASS with 9.voice large
'Shouts were given loudly.'

(Martin, 2019, 30)

Like in Gĩkũyũ, some Swahili LVCs simply reject passivization of their nominal element. This is shown for the LVC *piga mnada* 'to auction' (lit. beat auction) in (36). *Mifugo* 'livestock' which is an argument of the LVC can become subject under passivization (36b), the nominal element *mnada* 'auction' cannot (36c). (For additional data, see, Olejarnik, 2011, 158–159.)

(36) a. *Wakusanya kodi wa-li-pig-a mnada mifugo.*
2.collector 9.tax 2-PST-beat-FV 3.auction 4.livestock
'The tax collectors auctioned off the livestock.'

b. *Mifugo i-li-pig-wa mnada ma wakusanya kodi.*
4.livestock 4-PST-beat-PASS 3.auction by 2.collector 9.tax
'The livestock were auctioned off by the tax collectors.'

c. **Mnada u-li-pig-wa mifungo na wakusanya kodi.*
3.auction 3-PST-beat-PASS 4.livestock by 2.collector 9.tax
'The livestock were auctioned off by the tax collectors.'

(Martin, 2019, 30)

That *mayove* can become subject under passivization in (35b) but not *mnada* in (36c), follows from a general syntactic property of Swahili. The language is, in difference to, for instance, Gĩkũyũ, an asymmetric object language (e.g., Fleischhauer, 2023a, 167). Whereas Gĩkũyũ can make any non-subject argument of a ditransitive predicate the direct object (37), Swahili – and similar asymmetric object languages – restrict direct object properties to the beneficiary argument (38), see, for instance, Krifka (1995, 1400).

(37) a. *A-a-he-ir-e mw-anake i-buku.*
1-RMPST-give-PFV-FV 1-boy 5-book
'S/he gave the boy a book.'

b. *A-a-rĩ-he-ir-e mw-anake.*
1-RMPST-5-give-PFV-FV 1-boy
'S/he gave it to the boy.'

c. *A-a-mũ-he-ir-e i-buku.*
1-RMPST-1-give-PFV-FV 5-book
'S/he gave him/her a book.'

(38) a. *M-sichana a-li-m-pat-i-a ki-jana ki-tabu.*
1-girl 1-PST-1-give-CAUS-FV 1-boy 7-book
'The girl gave the boy a book.'

b. **M-sichana a-li-ki-pat-i-a ki-jana.*
1-girl 1-PST-7-give-CAUS-FV 1-boy
'The girl gave it to the boy.'

Mayove in (35b) can become the subject under passivization since it is the only non-subject argument of *piga*. This is different for the example in (36) as *mnada* 'auction' and *mifugo* 'livestock' are both non-subject arguments. In this case, it is *mifugo* which is realized as the direct object and not *mnada*. As shown in the discussion of the Gĩkũyũ LVC *hũra baini* 'to fine', the two non-subject arguments can become the subject under passivization in accordance with the language's status as a symmetric object language. Thus, the asymmetry in (38) does not represent a specific syntactic property of Swahili LVCs but can be explained on the basis of a general syntactic property of the language. In fact, taking this as a property of Swahili complex predicates –

rather than as a general syntactic constrained – obscures the real nature of complex predicates in that language. We therefore reject the descriptive as well as theoretical analysis proposed by Martin (2019).

The restriction on passivization has been related to the fact that Swahili is an asymmetric object language. This, however, does not account for the fact that the nominal element cannot co-occur with a bound argument marker. Swahili generally licenses the co-occurrence of a bound argument marker with a co-referential RP but – as shown above – prohibits it in the case of an LVC's nominal element. However, this restriction can be accounted for in a different general syntactic property of the language. Swahili shows differential object marking, it restricts object marking – in terms of bound argument markers co-occurring with a coreferential RP – to a specific subset of salient nouns (Aissen, 2003). Iemmolo and Klumpp (2014, 272) state that bound object marking is restricted to arguments having animate referents (cf. the examples in 39). Krifka (1995, 1399) states that the bound object marker occurs "regularly with animate NPs, especially if they are definite, and sometimes with inanimate definite NPs". Thus, bound object marking is restricted to salient argument expressions (animate or definite). Salient argument expressions are such which rank high on the animacy or definiteness scale meaning, i.e., having, for instance, an animate or a definite referent (e.g., Aissen, 2003, 437). The nominal element occurring within LVCs usually rank low both on these scales – they have inanimate and indefinite referents – and therefore do not license bound argument marking in Swahili. In Gĩkũyũ, bound object marking is not subject to differential object marking which means that, for instance, animate as well as inanimate object arguments can be pronominalized by a bound object marker. And in fact, we observed bound object marking of the LVC's nominal element in the LVC *hũra thimũ*. Thus, we can again relate the difference between Swahili and Gĩkũyũ concerning the pronominalization of the LVC's nominal element to a more general syntactic difference between the two languages.

(39) a. *Juma a-li-m-pig-a risasi tembo jana usiku.*
Juma 1-PST-1-beat-FV 9.bullet 1.elephant yesterday night
'Juma shot an/the elephant last night.'

b. *Risasi i-li-pig-a m-ti karibu na sisi.*
9.bullet 9-PST-beat-FV 3-tree near with us
'A bullet struck the tree near us.'
(Vitale, 1981, 123–124; glossing slightly adapted)

Thus, there is an independent grammatical reason for why the nominal element of an LVC cannot be pronominalized by a bound argument marker in Swahili: bound object markers are subject to differential object marking and are restricted to object

argument's which have animate/ definite referents. Concerning Swahili, we reject Krifka's strict statement that the LVC's nominal element cannot be direct object as we can explain the grammatical behavior of the LVC's nominal element to more general syntactic properties related to direct object arguments. In addition, some LVCs even allow their nominal element to become the subject under passivization, as shown above. This clearly falsifies Krifka's strict statement. We have no reason to propose that Swahili LVCs differ concerning their grammatical properties from their Gĩkũyũ correspondents. Instead, we have shown that the language-specific differences between Gĩkũyũ and Swahili LVCs follow from more general syntactic properties with respect to which the two languages differ and which are not specifically restricted to complex predicates.

7.2 Direct object properties: A look into Persian

The preceding sections have shown that the nominal element of LVCs show a non-uniform behavior with respect to direct object diagnostics. For Swahili, we argued that this is a consequence of the language's syntactic architecture. We have not come up with a similar explanation for Gĩkũyũ but regard it as idiosyncratic properties of individual LVCs. Persian is an Indo-European language which makes massive use of light verb constructions. Like the two Bantu languages discussed abobe, Persian shows similar non-uniformity when it comes to direct object properties of the LVCs' nominal element.

Within its history, Persian lost most of its monolexical verbs and compensated this loss by use of light verb constructions (e.g., Mohammad and Karimi, 1992, 195, Samvelian, 2018, 256). Light verb constructions in Persian come in different morphosyntactic types but we only focus on the most frequent type consisting of a noun-verb sequence (cf. Dabir-Moghaddam, 1997, 31 for an overview on other morphosyntactic construction types). Unlike the Bantu languages, Persian is a verb-final language and the nominal element of an LVC precedes its light verb. An illustrative example consisting of the noun *šekast* 'defeat' and the light verb *dâdan* 'give' meaning 'to defeat' (lit. defeat give) is shown in (40).

(40) *Ân-hâ došman-râ šekast dâd-and.*
DEM-PL enemy-ACC defeat give.PST-3PL
'They defeated the enemy.'

(Fleischhauer, 2021, 44)

A further difference between Persian, on the one hand, and the Bantu languages, on the other, is that Persian is a dependent-marking language. Rather than bound argument markers, it uses a combination of case marking and adpositional marking

to indicate the grammatical relations of the argument expressions. Persian has an accusative case marker *-râ* which attaches to direct object arguments. Indirect objects, which do not play a role in the current section, are marked by the preposition *be* 'to'. Similar to Swahili, Persian displays differential object marking. Accusative case marking is restricted to salient direct object arguments either having indefinite specific or definite reference (cf. Bossong, 1985; Lazard, 1992; Ghomeshi, 1997).

In case of the LVC *šekast dâdan*, accusative case marking is not found on the LVC's nominal element but on the noun *došman* 'enemy' which is the one being defeated. However, other LVCs show case marking on the LVCs nominal element. This is, for instance, the case for *latme zadan* 'to damage' (lit. damage hit) in (41).

(41) *Tegarg-e diruz în latme-râ be bâq-e man zad.*
hail-EZ yesterday DEM damage-ACC to garden-EZ 1SG hit.PST
'The yesterday's hail caused this damage to my garden.'
(Karimi-Doostan, 2011, 71; glossing slightly adapted)

The examples in (42) show that the nominal element of the LVC *jade kešidan* 'built a road' (lit. road pull) can receive accusative case marking (42a) and can become subject under passivization (42b). The noun exhibits *jâde* both relevant direct object properties. This is different from the considered Bantu languages, where nouns exhibit only a subset of direct object properties.

(42) a. *In âqâya in jade-râ be forudgâh kešid-and.*
DEM men DEM road-ACC to airport pull.PST-3PL
'These men build the road to the airport.'
b. *Jâde-hâ-i be forudgâh kešide šod-and.*
road-PL-INDEF to airport pull.PTCP become.PST-3PL
'(Some) Roads to the airport were build.'
(Fleischhauer, 2021, 48)

As the discussion reveals, a non-uniform distribution of direct object properties among the LVCs' nominal elements is attested in genetically unrelated languages. Concerning Persian LVCs, there exists no consensus in the literature whether there is any systematicity underlying the distribution of direct object properties.[17] It might be the case that it is an LVCs idiosyncratic property whether its nominal element bears direct object properties. But it could also be the case that whether an LVCs nominal element bears – some or all – direct object properties or not reflects a difference in the degree of the LVC's lexicalization. The nominal elements in stronger lexicalized

17 For a more detailed discussion of the grammatical properties of Persian light verb constructions, see Fleischhauer and Neisani (2020) and Fleischhauer (2020, 2021).

LVCs lacking – some or all – direct object properties, whereas does in – not yet – lexicalized LVCs retain them. We cannot provide even a speculative answer on this question but conceive of it as an interesting probably cross-linguistic project for future work.

8 Conclusion

The paper started with the observation that Gĩkũyũ possesses at least three different predicative constructions consisting of a semantically reduced verb and a nominal element. One of these constructions – represented by the sequence *hũra thimũ* 'to call' – resembles what in other languages, e.g., Swahili, is called a light verb construction. Departing from this observation we asked whether light verb constructions of the type exemplified by *hũra thimũ* can be distinguished on the basis of grammatical properties (i) from regular predicate-argument constructions headed, for instance, by *hũra* as a lexical full verb meaning 'to beat' and (ii) from semi-auxiliary constructions such as prospective '*enda* + INF' and inceptive '*rika* + INF'.

The analysis revealed that *hũra*-LVCs do not show any systematic grammatical differences compared to heavy uses of the same verb. However, we observed that pronominalization as well as passivization of the LVC's nominal element is subject to – as it seems – idiosyncratic restrictions. We also observed the same variance in other languages, indicating that LVCs are generally compatible with its nominal element bearing a grammatical relation – notably direct object – to the light verb.

LVCs differ from semi-auxiliary constructions with respect to their grammatical properties. The analyses in sections 4 and 5 revealed that semi-auxiliary constructions show more grammatical restrictions than LVCs. The restrictions concern voice marking, especially passivization. The severe restrictions on passivization – in line with the fact that the nominal element cannot be pronominalized by a bound argument marker – indicate that the nominal element entirely lost its direct object properties. We consider this as a reflex of the auxiliarization process which the respective verbs undergo. Such grammatical restrictions can be analyzed as 'decategorization', i.e., the loss of morphosyntactic properties of a lexical unit. Decategorization represents a central parameter in grammaticalization theory (e.g., Narrog and Heine, 2021, 72–78), and the results suggest that semi-auxiliaries show signs of grammaticalization – but are not yet fully grammaticalized – while the light verb does not exhibit this.

As the overall result, the case study shows that *hũra*-LVCs are grammatically more similar to regular predicate-argument constructions than to semi-auxiliary constructions. In fact, there are not any relevant grammatical differences between LVCs and regular predicate-argument constructions. The same is true of the related Bantu

language Swahili, as we attributed any restrictions on passivization and pronominalization to the existence of more general syntactic properties (i.e., asymmetry in the realization of the two non-subject arguments and differential object marking).

The results on the comparison of *hũra*-LVCs to corresponding full verb uses of *hũra* as well as to the semi-auxiliary constructions discussed in the paper should, in a next step, validated on a data set contain a larger number of *hũra*-LVCs. In addition, the criteria should be tested against a different data set containing other types of light verb constructions as well as semi-auxiliary constructions within the same but also other language.

Acknowledgments

The work on this study was partially funded by a grant from the German Research Foundation (DFG) and is part of the DFG-funded project 'Funktionsverbgefüge: Familien & Komposition' (HE 8721/1-3).

Bibliography

Aissen, Judith. 2003. Differential Object Marking: Iconicity vs. Economy. *Natural Language & Linguistic Theory* 21. 435–483. https://doi.org/10.1023/A:1024109008573.

Allerton, David. 2003. *Streched verb constructions in English*. London: Routledge.

Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9. 5–16.

Barlow, A. Ruffell. 1960. *Studies in Kikuyu grammar and idiom*. Edinburgh: William Blackwood.

Benson, T. Godfrey. 1964. *Kikuyu-English dictionary*. London: Clarendon Press.

Bergvall, Victoria L. 1987. *Focus in Kikuyu and Universal Grammar*. Doctoral dissertation: Harvard University.

Bossong, Georg. 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

Bresnan, Joan and Lioba Moshi. 1990. Object Asymmetries in Comparative Bantu Syntax. *Linguistic Inquiry* 21(2). 147–185.

Butt, Miriam and Aditi Lahiri. 2013. Diachronic pertinacity of light verbs. *Lingua* 135. 7–29. https://doi.org/10.1016/j.lingua.2012.11.006.

Cable, Seth. 2013. Beyond the past, present, and future: towards the semantics of 'graded tense' in Gĩkũyũ. *Natural Language Semantics* 21. 219–276. https://doi.org/10.1007/s11050-012-9092-3.

Clements, George N. 1984. Binding domains in Kikuyu. *Studies in the Linguistic Sciences* 14. 37–56.

Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.

Croft, William. 2022. *Morphosyntax*. Cambridge: Cambridge University Press.

Dabir-Moghaddam, Mohammed. 1997. Compound verbs in Persian. *Studies in the Linguistic Sciences* 27(2). 25–59.

Dik, Simon. 1997. *The Theory of Functional Grammar. Part 1: The structure of the clause*. Berlin: De Gruyter Mouton.

Fleischhauer, Jens. 2020. Predicative multi-word expressions in Persian. In *Proceedings of the 34th Pacific Asia Conference on Language, Information and Computation*, 552–561. Hanoi: Association for Computational Linguistics. https://aclanthology.org/2020.paclic-1.63.

Fleischhauer, Jens. 2021. Simplex and Complex Predicates in Persian: An RRG Analysis. In Robert Van Valin (ed.), *Challenges at the Syntax-Semantics-Pragmatics Interface*, 31–62. Newcastle upon Tyne: Cambridge Scholars Publishing.

Fleischhauer, Jens. 2023a. Argument doubling and right-dislocation – An RRG analysis of head-marking in isiZulu. In Jens Fleischhauer and Claudius Patrick Kihara (eds.), *African languages from a Role and Reference Grammar perspective*, 163–191. Berlin: De Gruyter. https://doi.org/10.1515/9783110795295-007.

Fleischhauer, Jens. 2023b. German light verb constructions with non-eventive nouns. In Jean-Pierre Colson (ed.), *Phraseology, constructions and translation. Corpus-based, computational and cultural aspects*, 387–398. Louvain: Presses Universitaires de Louvain.

Fleischhauer, Jens. 2023c. Prospective aspect and current relevance – A case study of the German prospective '*stehen vor* NP'-construction. *Journal of Germanic Linguistics* 35(4). 371–408. https://doi.org/10.1017/S1470542722000198.

Fleischhauer, Jens and Thomas Gamerschlag. 2019. Deriving the meaning of light verb constructions – a frame account of German *stehen* 'stand'. In Constanze Juchem-Grundmann, Michael Pleyer and Monika Pleyer (eds.), *Yearbook of the German Cognitive Linguistics Association, vol. 7*, 137–156. Berlin: De Gruyter Mouton. https://doi.org/10.1515/gcla-2019-0009.

Fleischhauer, Jens and Mozhgan Neisani. 2020. Adverbial and attributive modification of Persian separable light verb constructions. *Journal of Linguistics* 56(1). 45–85. https://doi.org/10.1017/S0022226718000646.

Ghomeshi, Jila. 1997. Topics in Persian VPs. *Lingua* 102(2–3). 133–167. https://doi.org/10.1016/S0024-3841(97)00005-3.

Gibson, Hannah and Lutz Marten. 2016. Variation and Grammaticalisation in Bantu Complex Verbal Constructions: The Dynamics of Information Growth in Swahili, Rangi and SiSwati. In Léa nash and Pollet Samvelian (eds.), *Approaches to Complex Predicates*, 70–109. Leiden: Brill.

Good, Jeff. 2005. Reconstructing morpheme order in Bantu. The case of causativization and applicativization. *Diachronica* 22(1). 3–57. https://doi.org/10.1075/dia.22.1.02goo?locatt=mode:legacy.

Good, Jeff. 2007. Slouching towards deponency: A family of mismatches in the Bantu verb stem. In Matthew Baerman, Greville G. Corbett, Dunstan Brown and Andrew Hippisley (eds.), *Deponency and morphological mismatches*, 203–230. Oxford: Oxford University Press.

Grimm, Scott and Beth Levin. 2017. Artifact Nouns: Reference and Countability. In Andrew Lamont and Katerina Tetzloff (eds.), *Proceedings of the 47th Annual Meeting of the North East Linguistic Society (NELS 47)*, 45–64. Amherst, MA: University of Massachusetts.

Guthrie, Malcolm. 1967. *The classification of Bantu languages*. London: Dawson.

Heine, Bernd. 1994. On the genesis of aspect in African languages: The proximative. *Proceedings of the 20th Annual Meeting of the Berkeley Linguistics Society* 35–46.

Hopper, Paul and Elizabeth Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165525.

Hyman, Larry and Alessandro Duranti. 1982. On the object relation in Bantu. In Paul Hopper and Sandra Thompson (eds.), *Studies in Transitivity*, 217–239. Leiden: Brill.

Iemmolo, Giorgio and Gerson Klumpp. 2014. Introduction to the special issue 'Differential Object Marking: theoretical and empirical issues'. *Linguistics* 52(2). 271–279. https://doi.org/10.1515/ling-2013-0062.

Karimi-Doostan, Gholamhossein. 2011. Separability of light verb constructions in Persian. *Studia Linguistica* 65(1). 70–95. https://doi.org/10.1111/j.1467-9582.2011.01178.x.

Kihara, Claudius Patrick. 2016. *Aspects of Gĩkũyũ (Kikuyu) Complex Sentences: A Role and Reference Grammar Analysis*. Doctoral dissertation: Heinrich-Heine University Düsseldorf.

Krifka, Manfred. 1995. Swahili. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld and Theo Vennemann (eds.), *Syntax: An international handbook of contemporary research*, 1397–1418. Berlin: De Gruyter.

Kuteva, Tania. 2001. *Auxiliation: An enquiry into the nature of grammaticalization*. Oxford: Oxford University Press.

Kuteva, Tania, Bas Aarts, Gergana Popova and Anvita Abbi. 2019. The grammar of 'non-realization'. *Studies in Language* 43(4). 850–895. https://doi.org/10.1075/sl.18044.kut.

Lambrecht, Knud. 2001. Dislocation. In Martin Haspelmath, Ekkehard König, Wolfgang Raible and Wulf Oesterreicher (eds.), *Language typology and language universals: An international handbook. Vol. 2*, 1050–1078. Berlin: De Gruyter.

Lazard, Gilbert. 1992. *A Grammar of Contemporary Persian*. Costa Mesa, CA: Mazda Publishers.

Levin, Beth, Lelia Glass and Dan Jurafsky. 2019. Systematicity in the semantics of noun compounds: The role of artifacts vs. natural kinds. *Linguistics* 57 (3). 429–471.

Longacre, R. E. 1976. *An Anatomy of Speech Notions*. Lisse: The Peter de Ridder Press.

Maho, Jouni F. 2003. A classification of the Bantu languages: an update on Guthrie's referential system. In Derek Nurse and Gérard Philippson (eds.), *The Bantu Languages*, 639–651. London: Routledge.

Martin, Joshua. 2019. Analyzing Complex Predicates in Swahili. In *Proceedings of the Florida Linguistics Yearly Meeting (FLYM) 5*, vol. 6, 25–36.

Mohammad, Jan and Simin Karimi. 1992. Light verbs are taking over: Complex verbs in Persian. In J. A. Nevis and V. Samiian (eds.), *Proceedings of the Western Conference on Linguistics. Vol. 5*, 195–212. Fresno: California State University.

Mugane, John. 2003. Hybrid Constructions in Gĩkũyũ: Agentive Nominalizations and Infinitive-gerund constructions. In Miriam Butt and Tracy Holloway King (eds.), *Nominals: Inside and Out*, 235–266. Stanford: CSLI Publications.

Narrog, Heiko and Bernd Heine. 2021. *Grammaticalization*. Oxford: Oxford University Press.

Nichols, Johanna. 1986. Head-Marking and Dependent-Marking Grammar. *Language* 62(1). 59–119. https://doi.org/10.2307/415601.

Nichols, Lynn. 2008. Lexical semantic constraints on noun roots and noun borrowability. *Studies in Language* 32 (3). 683–700.

Nicolle, Steve. 2023. *A Grammar of Digo*. Dallas: SLI Institute.

Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description: Vol. 2, Complex Constructions*, 52–150. Cambridge: Cambridge University Press.

Olejarnik, Magdalena. 2009. *Complex Predicates in Swahili: An LFG Approach*. Doctoral dissertation: Universytet Warszawski.

Olejarnik, Magdalena. 2011. An LFG Approach to Complex Predicates in Swahili. In Piotr Banski, Beata Lukaszewicz, Monika Opalinska and Joanna Zaleska (eds.), *Generative Investigations: Syntax, Morphology, and Phonology*, 139–165. Newcastle upon Tyne: Cambridge Scholars Publishing.

Pustejovski, James. 1991. The Generative Lexicon. *Computational linguistics* 17. 409–441.

Pustejovski, James. 1995. *The Generative Lexicon*. Boston: The MIT Press.

Saeddi, Zari. 2016. *Complex predicates in modern Persian*. Sheffield: Equinox.

Samvelian, Pollet. 2018. Specific features of Persian syntax. In Anousha Sedighi and Pouneh Shabani-Jadidi (eds.), *The Oxford Handbook of Persian Linguistics*, 226–269. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198736745.013.9.

Schwarz, Florian. 2003. Focus Marking in Kikuyu. *ZAS Papers in Linguistics* 30. 41–118.

Van Valin, Robert. 2008. RPs and the nature of lexical and syntactic categories in Role and Reference Grammar. In Robert Van Valin (ed.), *Investigations of the Syntax-Semantics-Pragmatics Interface*, 161–178. Amsterdam: John Benjamins.

Van Valin, Robert. 2023. Principles of Role and Reference Grammar. In Delia Bentley, Wataru Nakamura Ricardo Mairal Usón and Robert Van Valin (eds.), *The Cambridge Handbook of Role and Reference Grammar*, 17–180. Cambridge: Cambridge University Press.

Vitale, Anthony J. 1981. *Swahili syntax*. Dordrecht: Foris.

Wichmann, Søren and Jan Wohlgemuth. 2008. Loan verbs in a typological perspective. In Thomas Stolz, Dik Bakker and Rosa Salas Palomo (eds.), *Aspects of Language Contact*, 89–122. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110206043.89.

Wittke, Jonas. 2015. Verbs. In Robert Englebretson (ed.), *A Basic Sketch Grammar of Gĩkũyũ*, 52–81. Houston: Rice Working Papers in Linguistics.