Timur Maisak

Complex verbs in Kina Rutul (Nakh-Daghestanian)

1 Introduction: Complex verbs in Nakh-Daghestanian

Nakh-Daghestanian, also known as East Caucasian, is one of the three indigenous language families of the Caucasus together with Abkhaz-Adyghe (West Caucasian) and Kartvelian. Nakh-Daghestanian languages are spoken in the eastern part of the Caucasus, mainly in the Russian republics of Daghestan, Chechnya and Ingushetia, as well as in northern Azerbaijan and eastern Georgia. More than forty languages belong to the family, many of which are minority languages and remain unwritten and understudied; the precise number of languages remains unclear due to the problematicity of the language/ dialect distinction for quite a few idioms spoken in Daghestan. The main branches of the family, going from north-west to south-east, are Nakh, Avar-Andic, Tsezic, Lak, Dargwa, Lezgic and Khinalug; for a recent general overview of the family, see (Ganenkov and Maisak, 2020).

Although, to varying degrees, the contrast between simplex and complex verbs is relevant to all branches of Nakh-Daghestanian, it is especially well-pronounced in the Lezgic languages, which possess large amounts of complex verbs. In many of these languages, verbal compounding is also the only available means of creating new verbal lexemes: there is no other way to make borrowed verb stems inflected. The Lezgic branch, together with a single-language Khinalug branch, is southernmost: it includes nine languages, which are spoken in central and southern Daghestan (Archi, Agul, Tabasaran), or both in Daghestan and northern Azerbaijan (Rutul, Tsakhur, Lezgian), or exclusively on the Azerbaijani side of the current political border (Budugh, Kryz, Udi). Given long-standing contact with Azerbaijani, as well as, in earlier times, with Iranian languages, the Turkic and Iranian influence on the Lezgic languages is quite significant. The spread of complex verbs, which are very typical of both Turkic and Iranian, may be one of the manifestations of this contact influence.

In the Lezgic languages, complex verbs, also known as verbal compounds or light-verb constructions, consist of a lexical component and a postposed light verb. Lexical

Timur Maisak, Institute of Linguistics, Russian Academy of Sciences & HSE University, Moscow, Russia

https://doi.org/10.1515/9783112229330-006

components, sometimes labeled 'nominal parts' or 'coverbs',[1] can be represented by nouns, adjectives, adverbs, ideophones, and also acategorical bound stems which cannot function as autonomous words. Light verbs, which usually host all verbal inflectional marking (tense, aspect, mood, polarity, etc.), typically include such high-frequency verbs with a generalized meaning, as 'be, become', 'do, make', 'give', 'say', 'hit, beat', 'go', 'come' and some others. The most productive means of expanding verbal vocabulary is to borrow verbs from a dominant language (e.g. Russian or Azerbaijani) in combination with a light verb, 'be, become' for intransitives and 'do, make' for transitives. For brief overviews of complex verbs in Nakh-Daghestanian, see Klimov and Alekseev (1980, 182–183), Ganenkov and Maisak (2020, 109–111) and Maisak (2020a, 346–352), as well as papers in the volume on word-formation (Müller et al., 2016). In-depth studies of complex verbs in individual languages remain rare. In her book on the Udi morphosyntax, Harris (2002, ch.4) dedicates a special chapter to the analysis of the Udi complex verbs with respect to the wordhood criteria. In a recent study, Kerimova (2023) scrutinizes the morphosyntatic properties of coverbs in complex verbs, mainly focusing on the Lezgian data. A detailed analysis of complex verbs with the light verb 'say' in Archi is given by Authier (2024), who suggests a semantic classification of verbs belonging to this very numerous class.

An important problem in the description of complex verbs in Nakh-Daghestanian is the identification of the boundaries of this class. As a rule, complex verbs are not a homogeneous group, as coverbs can be phonologically, morphologically and syntactically 'fused' with light verbs to varying degrees. There may be no strict border between what is usually described as complex verbs and free syntactic combinations (or idioms, in case of non-compositional meaning). For example, in the literature it usually remains without any discussion, to what extent verbs like 'to help' which are expressed by a combination "help (noun) + do"[2], are different from ordinary patient–verb combinations like "house + build" or "letter + write".

The present paper aims at a description of complex verbs in Kina Rutul, a dialect of Rutul, a Nakh-Daghestanian language of the Lezgic branch spoken in southern Daghestan, Russia. The data for the study were collected during the author's fieldwork in 2018–2023. Apart from the introduction and the conclusion, there are three main sections. In Section 2, I present a brief linguistic characteristics of Kina Rutul in terms of major peculiarities of its phonological, morphological and syntactic structure. Section 3 gives an overview of structural types of complex verbs, including various types of "adjective + light verb", "noun + light verb", "acategorical coverb + light

1 The term 'coverb' (or 'co-verb') should not be confused with 'converb' which refers to a non-finite verb form that serves to express adverbial subordination.

2 Hereafter, I give the literal meaning of verbal compounds/ complex verbs in double quotes, while their proper meaning is cited in single quotes according to a general convention.

verb", "verb stem + light verb" and "locative adverb + verb" combinations. In Section 4, I look at the morphosyntactic properties of complex verbs, focusing on whether they can be treated as (compound) words.

2 Kina Rutul: The typological profile

Rutul is a relatively small Nakh-Daghestanian language of the Lezgic branch, mainly spoken in a dozen of villages of southern Daghestan (mostly Rutulsky District of the Republic of Daghestan, Russia) and a few settlements in the northern part of Azerbaijan. The total number of speakers is usually estimated as 30,000 or slightly more. Kina Rutul is a single-village variety of Rutul spoken in the village of Kina (Rutulsky District), which is located on the western periphery of the Rutul-speaking area and close to the area where Tsakhur, a sister language of Rutul, is spoken.

According to a dialectal classification of Rutul presented by Ibragimov (1978, 2004), the main dialects of the language are Mukhad, Shinaz, Myukhrek, Ikhrek and Borch-Knnov. Under this approach, Kina Rutul (as well as a few other single-village idioms) is an 'intermediate' or 'mixed' variety, which does not belong to any of these major dialects but rather accumulates linguistic features of Mukhad, Shinaz and Ikhrek dialects. An alternative may be to treat Kina Rutul as a small dialect on its own.

Rutul remains understudied, although there exist a number of grammar sketches of the language published since the early 20th century, including Dirr (1912); Ibragimov (1978, 2004); Alekseyev (1994) and Makhmudova (2001), all of which mostly treat phonological and morphological issues. None of these has the data on the Kina variety, which became a subject of dedicated research only in the last decade. Since 2016, a project aiming at the documentation of Kina Rutul has been running at the HSE University (Moscow, Russia), and the present contribution stems from the author's fieldwork fulfilled within this project. A collective descriptive volume on Kina Rutul is currently being prepared (Daniel et al., to appear), and in the present chapter, I largely follow the transcription and glossing conventions of the volume. In the description of Kina Rutul light verb constructions, I rely both on elicitation and the data from a spoken text corpus collected by the project members: the corpus mainly includes personal narratives and folk tales, and at present comprises ca. 20,000 word tokens. In the examples cited below, those taken from texts are accompanied by the indication 'Text Corpus'; no indication is given for those examples that were elicited.

Like the majority of Nakh-Daghestanian languages, Kina Rutul possesses a rich phonological inventory, with a three-way contrast between voiceless (aspirated), voiced (non-aspirated) and ejective stops (e.g., /k/ ∼ /g/ ∼ /k'/), a number of uvulars

(e.g., /χ/, /q/, /ɢ/), an opposition between short and long vowels (e.g., /a/ ∼ /a:/), and a prosodic pharyngealization feature acting as a secondary articulation with a peak on vowels and/ or uvulars (e.g., *raˤq* 'road', *qʼʷaˤd* 'two', *jɨˤχɨs* 'hit, beat').[3]

In the nominal paradigm, there is an opposition between an unmarked singular and a suffixally marked plural (e.g., *χal* 'house' ∼ *χal-bɨr* 'house-PL'). Each noun is assigned to one of the four genders, also known as '(noun) classes': Gender 1 is masculine, Gender 2 is feminine, the two other genders include non-human referents (most non-human animates belong to Gender 3 and most inanimates belong to Gender 4, but there are exceptions). The category of gender displays itself in agreement, mainly on verb stems, but also on cardinal numerals and the reflexive pronoun. The exponents of both series, which can be prefixes, infixes or suffixes, are single-consonant markers *r, b, d, w, j, l* and also zero. Gender agreement in noun phrases is controlled by the head, whereas on the clause level the controller is the absolutive noun phrase, be it the patient of a transitive verb or the single core argument of an intransitive verb.

Also mirroring a general Nakh-Daghestanian profile, Kina Rutul has a rich case inventory, with the unmarked absolutive case and all the other (oblique) cases being derived from a marked oblique stem. Rutul is a morphologically ergative language, thus agents of transitive verbs are encoded by the ergative case, patients of transitive verbs and core arguments of intransitive verbs are encoded by the absolutive case. A number of experiential, or affective, verbs ('see', 'know', 'want', 'love', among some others) encode their subject in the dative instead of the ergative. The genitive case in Rutul is a nominal instantiation of a more general 'attributive' category, with one and the same attributive suffix being used on various types of nominal modifiers including adjectives, demonstrative, participles, etc.

Sentences (1)–(3) from the text corpus of Kina Rutul illustrate the use of the ergative for marking the agents of transitive verbs (cf. *wɨylira* 'husband', *rɨšera* 'girl', *zad* 'I'), the absolutive for marking the patients (cf. *gʷalaχ* 'job' and *ʁil* 'leg'), and the genitive, or 'attributive', form of a noun in the modifying function (cf. *gedijed* 'boy's'). In (2), we can also see the postpositional phrase *urusašdɨ bejda* 'near the Russians', in which the complement takes the genitive form. As for the gender agreement on verbs, in (1) it is with the Gender 3 absolutive *gʷalaχ* 'work, job', in (2) it is with the elided absolutive plural subject 'we' (hence, the 'human plural' gender marker is used), and in (3) it is with the Gender 3 absolutive *ʁil* 'leg'. Note that in (1)–(2), *gʷalaχ waʔas* 'to work' and *ješemiš hɨkɨs* 'to live' represents two of many types of compound verbs, to be discussed in Section 3.

3 By convention, pharyngealization is marked only on vowels of the first syllable, although it usually 'spreads' to the second syllable as well (e.g., [raˤqˤ] 'road', [qʼʷˤaˤd] 'two', [jɨˤχˤɨˤs] 'hit, beat').

(1) *wɨɣl-ir-a=xa* *gʷalaχ* *w-aʔa-r-i-j*, *za-d=xa* *gʷalaχ*
husband-OBL-ERG=ADD job.ABS 3-do.IPFV-CVB-COP-PST I-ERG=ADD job.ABS
w-aʔa-r-i-j.
3-do.IPFV-CVB-COP-PST
{Many years ago we lived in Volgograd.} 'Both my husband worked and I worked.' (Text Corpus)

(2) *ješemiš* *d-iši-r* *urus-aš-dɨ* *bejda*
live HPL-become.PFV-CVB(AOR) Russian-OBL.PL-ATTR beside
'We lived near the Russians.' (Text Corpus)

(3) *mi* *ged-ije-d* *ʁil-i=xʷa* *mi* [...] *riš-e-ra* *ʁil*
this boy-OBL-ATTR leg-OBL(SUP)=REP this girl-OBL-ERG leg.ABS
s-i<w>χɨ-r=xʷa
DOWN-<3>push.PFV-CVB(AOR)=REP
{Once in a bus ...} 'The girl stepped on the boy's foot.' (Text Corpus)

The case inventory includes a large array of locative cases, which express both localization with respect to the landmark (e.g., 'inside' or 'on the upper surface') and direction of motion ('towards' or 'from') or the lack thereof. Thus, the nominal form *ʁili* 'on a leg' in (3) is the form of the case called 'superessive', as it describes location at or motion towards the upper surface. Localization and direction is also expressed on verbs by means of locative prefixes (or 'preverbs'), some of which are cognate to locative case markers. For example, the verb *s-ɨwχɨr* 'pushed down' (4) contains the prefix *s-* expressing downward motion, *l-iwxir* 'put on' (4) has a prefix *l-* 'on', whereas *kʲ-awar* 'add, mix' (4) includes a prefix *kʲ-* with the meaning 'in contact with the surface, in a substance'.

(4) *hejeg* *l-i<w>xi-r*, *ara* *sa* *q'ʷaˤ-b* *dur* *q'el-dɨ*
cauldron.ABS UP-<3>put.PFV-CVB inside one two-3 spoon.ABS salt-ATTR
kʲ-a<w>a-r=a
CONT-<3>do.IPFV-CVB=be
'After we have put the cauldron, we add a couple (lit. one-two) of spoons of salt.' (Text Corpus)

In (4), the numeral *q'ʷaˤ-b* 'two' agrees in gender with the Gender 3 head 'spoon' it modifies; this is not the case with adjectives which lack gender agreement. Note that adjective stems combine with the attributive suffix only in the noun-modifying function, e.g., *hɨχ-dɨ insan* [good-ATTR person] 'a good person'. They take adverbial suffixes when they modify predicates, as in *hɨχ-ana hac'ara* [good-ADV knows] 'knows well'. There is also a third form of adjectives, labeled 'co-verbal' form, which is found in complex predicates which the adjectives is part of, e.g. *hɨχ-a jiʔi* [good-ADJ is] 'is

well, feels well'. Compound verbs with the co-verbal adjectival form will be treated below in Section 3.

As for the verbal system of Kina Rutul, it is fairly rich and includes quite a few synthetic and periphrastic forms (Maisak, 2020b). The overwhelming majority of verb stems agree in gender and include a morphological slot (prefixal or infixal) for gender agreement. The presence and the type of this slot is a lexical property of a verb stem and has nothing to do with the meaning of the verb or any of its morphosyntactic features. Thus, the verb stem *hiši-*/ *w-iši-*/ *r-iši-* etc. [1.become.PFV/ 2-become.PFV/ 3-become.PFV] has a prefixal gender slot, the Gender 1 being zero-marked, the verb stem *ji<r>q'i-*/ *ji<b>q'i-*/ *ji<d>q'i-* etc. [<1/2>become.PFV/ <3>become.PFV/ <4>become.PFV] has an infixal gender slot, and the verb *gič'e-* 'be afraid' lacks an agreement slot altogether. A group of stative verbs comprising the copula *ji?i*/ *i* and a few existential verbs with locative semantics ('be in', 'be under' etc.), have a reduced paradigm and possess only one stem, from which all forms are derived. To the contrary, 'canonical' verbs possess several inflectional stems, on which the synthetic forms are based: these are the perfective, imperfective and infinitive stems, none of which is generally predictable from the other two. The relation between the stems can also be suppletive, e.g. the perfective stem of a very frequent verb 'become' is *hiši-* (Gender 1), whereas its imperfective stem is *ru?u-* and its infinitive stem is *hiki-*.

While the stative verbs have two finite synthetic tenses (present and past), the synthetic forms of canonical verbs are mostly non-finite forms including converbs, participles, infinitive and verbal noun, but also a few non-indicative forms (imperative, optative). The finite indicative system consists of periphrastic forms, which include two components, a lexical verb and a postposed auxiliary, the latter usually being the short copula *i* or the stative verb *a* 'be (inside)'. In the periphrastic forms, lexical verbs occur in one of the two converbs (perfective or imperfective), one of the three participles (perfective, imperfective and prospective) and the infinitive. Thus, the form *wa?ar-ij* in (1) is a periphrastic past habitual consisting of the imperfective converb *wa?ar-* 'doing' and the past tense copula *ij* 'was', and the form *kʲawar=a* in (4) is a periphrastic present tense consisting of the imperfective converb *kʲawar-* 'adding' and the present tense auxiliary *a* 'is (inside)'. The aorist, which is a high-frequent perfective past tense, is special in that the affirmative copula is always dropped, so that finite forms like *liwxir* in (4) are syncretic with perfective converbs: *liwxir* may mean both 'put on' (past) and 'having put on'.

Traditionally, verbal lexemes in Rutul (and the Lezgic languages in general) are divided into simple, or underived, and two groups of derived verbs, namely prefixal verbs containing one or more locative prefixes, and complex verbs consisting of a coverb and a light verb. The number of simple verbs is very restricted: thus, Alisultanov and Sulejmanova (2019, 480) call the number of simple verbs in the Mukhad dialect 'insignificant', while Makhmudova (2001, 163) suggests that the

underived verb stems account for just 7 or 8 percent of the verbal lexicon. For Kina Rutul, 55 simple verbs have been found (e.g., *ɬikɬis* 'become', *haʔas* 'do', *ɬiwɬis* 'give', *jiqʼas* 'come', *jaχas* 'run', *iles* 'eat', *raʿʁas* 'drink', *jetas* 'kick', *wezas* 'milk', etc.). Prefixal verbs are a more numerous group, while the number of verbal compounds as the only open class of verbal lexemes can be estimated in hundreds.[4] One should take into account, however, that the boundaries of the class of verbal compounds are fuzzy, and in the remainder of the paper I will try to touch upon the issue of relation between verbal compounds in a more general sense and complex verbs in a narrower sense of multi-word verbal lexemes.

The present paper suggests the first comprehensive treatment of complex verbs in Rutul. Starting with Dirr (1912, 97), the existing descriptions of the language always mention the existence of compound (or complex, 'periphrastic' etc.) verbs, but usually only present a few examples (as in Alekseyev, 1994, 226 and Alekseyev, 2016, 3538 for the Luchek dialect). Makhmudova's (2001, 167–168) grammar sketch of the Mukhad dialect of Rutul is an exception is that it contains longer lists of complex verbs, although without much comments on their morphosyntactic behaviour.

3 Structural types of light-verb constructions in Kina Rutul

There are quite a few compound verbal lexemes in Kina Rutul which are translational equivalents of simple verbs in English or Russian. Thus, 'to help' is *kumag haʔas* "help do" or *kumag ɬiwɬis* "help give", 'to search' is *aramiš waʔas* "search do", 'to marry (about a man)' is *ɢari raʔas* "wife do", 'to smoke' is *pʼapʼrɬis deʔes* "cigarette pull", and so on. The number of such combinations amounts to hundreds, at the very least. Some of these combinations may turn out to be just periphrastic ways of expressing certain meanings, without displaying any special morphosyntactic

4 The proportion of simple, prefixal and complex verbs is subject to variation even among the languages of the Lezgic group (also, the methodology of counting may differ across the language experts). Thus, in Udi there are about 50–60 simple verbs, and at least 50 prefixal verbs (Maisak, 2008, 98); according to Schulze (2016, 3570), complex verbs account for more than 75% of all verbal lexemes in the language. In Agul, there are at least 120 simple verbs and more that 350 prefixal verbs (Maisak and Ganenkov, 2016, 3585). In Archi, the number of simple verbs is about 170, and the rest of the lexicon consists of complex verbs, as there is no locative prefixation (Chumakina, 2016, 3599). To the contrary, in Tsakhur there seem to be less than ten underived verbs without preverbs (Kibrik and Testelets, 1999, 67).

behaviour, and some of them may turn out to show at least some properties of single words (see further Section 4).

Although various types of verbal compounds exits in many languages, including English (cf. *take a rest* or *look forward*) or Russian (cf. *okazat' pomošč* 'to deliver help' or *soveršit' samoubijstvo* 'to commit suicide'), they are especially prominent in languages like Rutul, where compounds are found among the lexemes expressing very 'basic' concepts. In the present study, I focus on a set of complex verbs that appear in the (unpublished) Kina Rutul dictionary, as well as in elicitation tasks and the available text corpus. Thus, in the Kina Rutul dictionary, collected by the fieldwork team I was part of, almost half (48 per cent) of about 300 verbs included in the dictionary are complex verbs, the rest being simple and prefixal ones.[5]

In this section, I present an overview of compound verbs in Kina Rutul, briefly commenting on their composition and basic properties. The types of verbal compounds surveyed below, ordered by the form of the lexical part, are:

– ADJECTIVE (co-verbal form) + copula
– ADJECTIVE (co-verbal form) + light verb 'become'
– ADJECTIVE (co-verbal form) + light verb 'do'
– NOUN (absolutive case) + light verb 'become'
– NOUN (absolutive case) + light verb 'do'
– NOUN (absolutive case) + other light verbs
– various 'NOUN (oblique case) + verb' combinations
– ACATEGORICAL BOUND COVERB + light verbs 'become', 'do' or 'give'
– VERB STEM + light verb 'do' (= causatives)
– BORROWED VERB STEM + light verbs 'become' and 'do'
– LOCATIVE ADVERB + verb

3.1 Verbal compounds with adjectives

There are three subtypes of verbal compounds which include an adjective as the lexical part. In all three subtypes, adjectives occur in the co-verbal form in *-a*/ *-e* (some adjectives have an unmarked co-verbal form identical to the bare adjective stem).

5 The Kina Rutul dictionary is based on a word list accepted within the international project 'LexCauc – A lexical database for the languages of the Caucasus' (2017–2020), based at the MPI-SHH Jena and led by Diana Forker and Oleg Belyaev. The LexCauc word list includes 1132 entries, of which about 300 represent verbal meanings. The Kina Rutul data were collected by Konstantin Filatov, supervised by Michael Daniel.

3.1.1 Adjective (co-verbal form) + copula

In the first subtype, the light verb is the copula. Verbal compounds of this type describe states. Some of the compound predicates belong to the experiential syntactic class and have a dative subject, e.g. *hɨχa i* 'be good; feel good' (5), while the majority of them belong to the intransitive class, e.g. *hazɨr i* 'be ready; be prepared' (6).

(5) *za-s hɨχ-a jiʔi.*
I-DAT good-ADJ 1.COP
'I'm fine; I feel good.'

(6) *iz-dɨ sin hazɨr i*
I-ATTR all.ABS ready COP
'Everything on my side (lit. my all) is ready.'

3.1.2 Adjective (co-verbal form) + light verb 'become'

In the second subtype, the light verb is *hɨkɨs* 'become'. As the copula has a very reduced paradigm, *hɨkɨs* is used in its place in all the syntactic and tense-aspect contexts that are not available to the copula. Thus, in (7) the aorist form *č'irč'ima wišira* expresses perfective past, in (8) the present tense *ɢuru ruʔura* is used with the present habitual meaning, and in (9) the imperative form *jawaš hiš* is employed.

(7) *nek č'irč'im-a w-iši-r=a*
milk.ABS sour-ADJ 3-become.PFV-CVB=be
'The milk turned sour.'

(8) *wirʁ-ɨ-da walɨg-mar k'ɨb ɢuru ruʔu-r=a*
sun-OBL-APUD clothes-PL.ABS quickly dry.ADJ NPL.become.IPFV-CVB=be
'The clothes get dry quickly in the sun.'

(9) *huʁal jawaš hiš, je-s wirɨʁ w-ɨga-r=a*
rain.ABS slow.ADJ 4.become.IMP we-DAT sun.ABS 3-want-CVB=be
'The rain, stop! we need the sun.' (Text Corpus)

Any 'stative' compound with the copula has a 'dynamic' counterpart with the light verb 'become'. Still, compounds with the same lexical part, like *hazɨr i* 'be ready; be prepared' ~*hazɨr hɨkɨs* 'become ready; become prepared', cannot be treated as belonging to one lexeme, given that the copula and the verb 'become' are two distinct lexical items.

3.1.3 Adjective (co-verbal form) + light verb ‘do’

The third subtype of verbal compounds with adjectives comprises the light verb *haʔas* ‘do, make’. Semantically, such compounds are causative equivalents of compounds with adjectives and the copula or the verb ‘become’. There are numerous examples of compound inchoative/ causative pairs like *jawaš hɨkɨs* ‘come to a stop, become quiet’ *~jawaš haʔas* ‘stop (tr.), make quiet’ or *qʼɨxe hɨkɨs* ‘grow, become older’ *~qʼɨxe haʔas* ‘grow (tr.), raise’, etc.

(10) a. *mij-aː haˤ<t>χiˤ-r ha-bɨr sa-ʔ, mašin jawaš*
here-ELAT <HPL>go.PFV-CVB(AOR) that-PL.ABS down-LAT car.ABS stop.ADJ
ha<w>ɨ-r-diš.
<3>do.PFV-CVB-COP.NEG
‘From here, they went further down, they didn’t stop the car.’ (Text Corpus)

b. *ha-now-a (…) qʼɨx-e d-ɨʔɨ-r ha-d*
that-OBL.H-ERG senior-ADJ HPL-do.PFV-CVB(AOR) that-ATTR
χɨnime-r
child-PL.ABS
‘She raised those children.’ (Text Corpus)

c. *čʼabal-er raˤ<t>χaˤ-r-ɨ-j, čʼabal-eš-dɨ dɨn*
sheep-PL.ABS <4>shear.IPFV-CVB-COP-PST sheep-OBL.PL-ATTR wool.ABS
ji<d>ʁɨ-r tamɨz hiʔɨ-r
4.wash.PFV-CVB clean 4.do.PFV-CVB(AOR)
‘They sheared the sheep, washed the wool and made it clean.’ (Text Corpus)

3.2 Verbal compounds with nouns in the absolutive case

3.2.1 Noun (absolutive) + light verb ‘become’

The combinations of a noun in the absolutive case and the verb ‘become’ are not numerous among verbal compounds, as normally such combinations simply represent subject–verb constructions, as in (11a)–(11b), and not complex verb lexemes.

(11) a. *nu haː-sa-d žizni w-iši-r*
PART[R] that-MNR-ATTR life[R].ABS 3-become.PFV-CVB(AOR)
‘Well, life was like that.’ (Text Corpus)

b. *wɨɣɨl-dɨ kulfat w-iši-r-diš=x^{w}a, riši-jmar*
male-ATTR child.ABS 3-become.PFV-CVB-COP.NEG=REP sister-PL.ABS
d-iši-r […]
HPL-become.PFV-CVB(AOR)
'No boys were born, sisters were born …' (Text Corpus)

However, one can subsume under verbal compounds such combinations as *sur hikɨs* and *q'at' hikɨs* 'to break, to be divided', which include the light verb 'become' and the nouns *sur* 'side, half' and *q'at'* 'part, piece'. Both verbs describe the process of division or separation, and both are intransitive verbs taking the absolutive subject, cf. *k'azit* 'paper; letter' in (12a) and *berad ǯibɨr* 'spade handle' in (13a). Note that it is the absolutive subject that controls gender agreement: while in (12a) and (13a) this is not obvious, this is clear from (12b) and (13b) where the nouns *badu* 'trousers' and *rub* 'needle' both belong to gender 3. The nouns *sur* and *q'at'*, as independent nouns, belong to gender 4, hence in (12b) and (13b) with gender 3 agreement on the verb, it is definitely not them that are agreement controllers.

(12) a. *k'azit sur hiši-r=a*
paper.ABS side 4.become.PFV-CVB=be
'The paper tore apart.'
b. *badu sur w-iši-r=a*
trousers.ABS side 3-become.PFV-CVB=be
'The trousers tore apart.'

(13) a. *ber-a-d ǯibɨr q'at' hiši-r*
spade-OBL-ATTR tip.ABS piece 4.become.PFV-CVB(AOR)
'The spade handle broke off.'
b. *rub q'at' w-iši-r*
needle.ABS piece 3-become.PFV-CVB(AOR)
'The needle broke.'

3.2.2 Noun (absolutive) + light verb 'do'

As already mentioned above, combinations of nouns in the absolutive case with the verb 'do' are very numerous in Kina Rutul and often correspond to single-word verbs in languages like Russian or English: thus, 'to work' is *g^{w}alaχ waʔas* ("work do"), 'to moo' is *maʿh waʔas* ("moo do"),[6] 'to dance' is *muk' waʔas* ("dance do"), 'to

6 Note that I subsume ideophones as parts of verbal compounds under nouns, as they do not display a behaviour noticeably different from nouns in their inflection or syntactic distribution.

think' is *fikɨr waʔas* ("thought do"), and so on. As a rule, such combinations describe physical activity, social activity or sound production (rarer, also mental activity). As examples below show, compounds of this type take an agent in the ergative.

(14) a. *zir-ɨ-ra guʒli-na maʿh w-aʔa-r=a*
cow-OBL-ERG strong-ADV IDEOPH 3-do.IPFV-CVB=be
'The cow is mooing loudly.'

b. *iz-dɨ χinχ-ɨ-ra hiχ-ana xed haʔa-r=a*
I-ATTR child-OBL-ERG good-ADV water.ABS 4.do.IPFV-CVB=be
'My son swims well.'

c. *iz-dɨ riši-ra bat'r-ana muk' w-aʔa-r=a*
I-ATTR sister-ERG beautiful-ADV dance.ABS 3-do.IPFV-CVB=be
'My sister dances gracefully.'

d. *wa-d hiji-d fikɨr w-aʔa-r=a?*
you-ERG what.OBL-ATTR thought.ABS 3-do.IPFV-CVB=be
'What are you thinking about?'

Probably, the most interesting in this class are those compounds which assign another absolutive argument besides the one that occurs in the 'noun + light verb' combination. Thus, in the following examples *sur* 'side, half', *q'at'* 'part, piece' and *peškeš* 'gift' are nouns in the absolute case, but there are also other absolutive noun phrases ('trousers', 'bread', 'knife') which are patients of transitive verbs 'tear', 'break' and 'give as a gift', respectively. It is these 'external' absolutive NPs that control gender agreement on the light verb (for example, *badu* 'trousers', a gender 3 noun, controls agreement on *wɨʔɨra* 'did', whereas *sur* 'side' cannot be the controller here, as it belongs to gender 4). See Section 4 for the discussion of whether such nouns can really be treated as absolutive NPs.

(15) a. *za-d badu sur w-ɨʔɨ-r=a*
I-ERG trousers.ABS side 3-do.PFV-CVB=be
'I tore the trousers.'

b. *za-d xɨw q'at' hiʔɨ-r*
I-ERG bread.ABS piece 4.do.PFV-CVB(AOR)
'I broke the bread (in two).'

c. *za-d ha-nowu-s kant' peškeš hiʔɨ-r*
I-ERG that-OBL.H-DAT knife.ABS gift 4.do.PFV-CVB(AOR)
'I gave him the knife as a present.'

3.2.3 Noun (absolutive) + other light verbs

There is a number of more or less idiomatic combinations with a noun in the absolutive and a transitive verb other that 'do', e.g. 'give' or 'pull'. In such combinations, the noun seems to simply fill the position of the patient argument. Note that for some nouns, there exist semantically equivalent combinations with 'do' or a different verb, e.g. *kef deʔes*/ *kef haʔas* 'to enjoy' ("pleasure pull"/ "pleasure do"), *kumag hɨwɨs*/ *kumag waʔas* 'to help' ("help give"/ "help do").

(16) a. *ha-now dawat-mɨ-kʲ kef d-eʔe-r=a*
that-OBL.H(ERG) wedding-OBL.PL-CONT pleasure.ABS 4-pull.IPFV-CVB=be
'He is chilling at the wedding.'

b. *iz-dɨ šu-ra bala d-eʔe-r=a p'ap'rɨs*
I-ATTR brother-ERG much 4-pull.IPFV-CVB=be cigarette.ABS
'My brother smokes much.'

c. *je-d wa-s kumag wɨlc'a-r=a*
we-ERG you-DAT help.ABS 3.give.IPFV-CVB=be
'We are helping you!' (Text Corpus)

d. *χɨnχ-ɨ-ra heh raˤʁa-r=a*
child-OBL-ERG IDEOPH 4.drink.IPFV-CVB=be
'The boy is yawning (lit. drinks a yawn).'

e. *tɨl-ije-ra za-s miz jɨˤ<w>χɨ-r*
dog-OBL-ERG I-DAT tongue.ABS <3>hit.PFV-CVB(AOR)
'The dog licked me (lit. hit tongue at me).'

3.3 Verbal compounds with nouns in oblique cases

A number of compounds include nouns in oblique cases, which usually refer to locations or instruments. As a rule, such combinations are parts of standard syntactic constructions and neither semantically, not morphosyntactically seem to represent complex verb lexemes. For example, the dative noun *wɨɣlis* as part of *wɨɣlis rurus* 'to marry (about a woman)' is the beneficiary or goal NP 'to the husband' accompanying the motion verb (17a). The ergative noun *šinara* as part of *šinara hɨkɨs* 'to sweat' is the instrumental noun phrase 'with sweat' combined with the verb 'become' (17b). In the combination *neqʲa: kiʁipxus* 'to wake up', literally "to fall out from a dream" (17c), *neqʼa:* is a locative case (inelative) expressing the source of motion.

(17) a. *wɨ-dɨ riši wɨɣl-ɨ-s r-ɨxɨ-r=a-m*?
you-ATTR sister.ABS husband-OBL-DAT 2-go_to.PFV-CVB=be-Q
'Have your sister married?'

b. *siɣin-e-ra hejwamar šin-a-ra hiši-r=a*
heat-OBL-ERG horse.PL.ABS sweat-OBL-ERG NPL.become.PFV-CVB=be
'Horses sweated because of the heat.'

c. *wi ʁijʁa mis neqʲ-a: ki-ʁ-i<r>xu-r*?
you.ABS today when sleep-IN.ELAT PV-PV-<1>wake_up.PFV-CVB(AOR)
'When did you wake up today?'

However, in some compounds oblique nouns seem to be 'incorporated' elements, forming a closer bond with the verb than ordinary arguments. Thus, in *xije ewč'us* 'to wash oneself' ("to get into water") and *xiji liwxis* 'to drown' ("to appear upon water"), the location of drowning is expressed separately from the noun phrase *xije*/*xiji* 'in/ on water', which hence looks as a syntactically 'incorporated' lexical part.

(18) a. *zi lec'-ur-a xij-e eč'u-r*
I.ABS river-OBL-IN water.OBL-IN 1.bathe.PFV-CVB(AOR)
'I bathed in the river.'

b. *armi-j-a: majit q-i<d>q'i-r*,
army[R]-OBL-IN.ELAT corpse.ABS RE-<HPL>come.PFV-CVB(AOR)
mori-j-a xiji l-i<r>xu-r
sea[R]-OBL-IN water.OBL(SUP) PV-<2>appear.PFV-CVB(AOR)
'They brought the body from the Army, (he) drowned in the sea.' (Text Corpus)

3.4 Verbal compounds with acategorical bound coverbs

An important type of verbal compounds is based on coverbs that do not belong to any major lexical class. The lexical part is morphosyntactically bound in that it does not occur elsewhere besides a compound it is part of, and it is acategorical in that it is neither noun, nor adjective, adverb or verb. Unlike nouns used as coverbs, acategorical coverbs do not control gender agreement. When combined with 'become', they do not fill the subject slot and when combined with 'do', they do not fill the patient slot. Some of the acategorical coverbs (e.g. *χaˁr* 'to learn; learning', *dagul* 'to steal; stealing') can combine with both light verbs 'become' and 'do', deriving inchoative/causative pairs, see (19).

(19) a. *nin-e χinɨχ kitab q'ale haʔa-s χaˁr hiʔi-r*
mother-ERG child.ABS book.ABS read 4.do-INF learn 1.do.PFV-CVB(AOR)
'The mother taught the child to read books.'

b. *rasul juž žu-s χaʿr hiši-r haʿrf-bir*
Rasul.ABS 1.self.ABS 1.self-DAT learn 1.become.PFV-CVB(AOR) letter-PL.ABS
q'ale haʔas
read NPL.do-INF
'Rasul learned to read letters by himself.'

c. *rasul-a iz-dɨ t'exʲ dagul ha<w>ɨ-r*
Rasul-ERG I-ATTR sheep.ABS disappear <3>do.PFV-CVB(AOR)
'Rasul stole my sheep.'

d. *χɨniχ iškab-a dagul hiši-r*
child.ABS wardrobe-IN disappear 1.become.PFV-CVB(AOR)
'The little boy hid himself in the wardrobe.'

Others acategorical coverbs are restricted to combinations with just one light verb: e.g., *q'ale*/ *q'ɨle* 'to read; reading' only combines with 'do' (20a). An acategorical component *masa* 'to sell; sale' combines with the light verb 'give' (20b).

(20) a. *q'urʔan q'ɨle w-aʔa-r-i-j man, jasin*
Quran.ABS read 3-do.IPFV-CVB-COP-PST PART Yasin.ABS
w-aʔa-r-i-j wɨyle-š-e
3-do.IPFV-CVB-COP-PST husband-OBL.PL-ERG
'The men used to read the Quran, they read Surah Yasin.' (Text Corpus)

b. *za-d č'abɨl-er masa hɨwɨ-r, artɨχ-na*
I-ERG sheep-PL.ABS sell NPL.give.PFV-CVB(AOR) excessive-ADV
s-atɨ-r-diš
PV-NPL.leave.PFV-CVB-COP.NEG
'I sold the sheep, did not keep any extra ones.' (Text Corpus)

3.5 Causatives with the light verb 'do' as a type of verbal compounds

There are no morphological (affixal) causative derivations in Kina Rutul. Causatives are light verb constructions with a special verb form and the light verb *haʔas* 'do'. Only intransitive and experiential verbs can derive this type of causatives, but not transitive verbs with an ergative subject. Still, the group of compound 'do'-causatives is very numerous.

The lexical part of 'do'-causatives is a special 'periphrasis form' of the verb. With most verbs, it is formally identical with the 'full form' of the imperative, i.e. the form with a vocalic suffix. Thus, the imperative 'die!' is *jiq'-e* [1.die-IMP], although a truncated form *jiq'* is also available. In a compound 'do'-causative which means 'to kill', namely *jiq'-e haʔas* [die-IMP do.INF], only the form *jiq'-e*, not *jiq'* can be employed

(21a). The verb *lat'us* 'be finished, come to an end' (21b) is also intransitive, and its causative is a means to express intentional completion ('finish, bring to an end'). The verb *ʁag^{w}as* 'see' (21c) belongs to a class of bivalent experiential verbs with a dative subject; in the possibility of deriving causatives experiential verbs pattern with intransitives.

(21) a. *mi si-ri-ra mi edemi jiq'-e ha?a-r=a*
this bear-OBL-ERG this man.ABS 1.die-PER 1.do.IPFV-CVB=be
'And the bear kills ("makes die") the man.' (Text Corpus)

b. *armi-j-e služba l-a<p>t'-a ha<w>i-r*
army[R]-OBL-IN service[R].ABS PV-<3>end-PER <3>do.PFV-CVB
q-i<r>q'i-r
RE-<1>come.PFV-CVB(AOR)
'He finished ("made end") his service and came back.' (Text Corpus)

c. *xela-la qu-? za-d ha-biši-s ʁ-ag^{w}-a*
then-ELAT back-LAT I-ERG that-OBL.HPL-DAT PV-4.see-PER
q-i?i-r-diš
RE-4.do.PFV-CVB-COP.NEG
'After that, I didn't show ("make see") this to them anymore.' (Text Corpus)

3.6 Verbal compounds with borrowed verb stems

Lexical parts of verbal compounds can be represented by a borrowed Azerbaijani or Russian verb. Both types of borrowed verbs are bound components, which do not occur as independent words outside complex verbs.

This type of verbal compounds is especially important, because it represents the only way of creating new verbal lexemes. Compounds based on Azerbaijani verbs belong to an older layer. Azerbaijani used to be an important contact language in southern Daghestan for centuries, and also functioned as a language of school education in early Soviet times, until Russian was finally introduced in this role in 1950s. Nowadays, the command of Azerbaijani among the Rutul people is low.[7] To the contrary, the Rutul–Russian bilingualism is almost complete, as elsewhere in Daghestan: only very small children or very old people (especially women, who do not often leave their own villages) have little to no command of Russian. Thus, the role of Russian in the extension of the lexicon, be it the borrowing of nouns or

7 See Chechuro et al. (2021) for a detailed account of Azerbaijani bilingualism in southern Daghestan.

verbs (as part of verbal compounds, in the latter case) has become higher in the last decades.

The borrowed Azerbaijani verbs occur in the form in *-miš*, which is originally the perfect participle in Azerbaijani. As a borrowed item, the form in *-miš* is acategorical and simply represents the lexical verb without any inflections. The choice of this particular form as the lexical part of complex verbs is typical both in Nakh-Daghestanian languages and elsewhere, see Bağrıaçık et al. (2015, 110) on verbs borrowed from Oghuz Turkic. The light verbs is either 'become', for intransitives, or 'do', for transitives.

Some of the lexical parts are attested with just one light verb: thus, *aramiš* is only found in the transitive compound *aramiš haʔas* 'to look for, search' (22a), while *ješemiš* only occurs within the intransitive compound *ješemiš hiкis* 'to live' (22b).

(22) a. *gad-ije-ra ǯu-s lišanči aramiš r-aʔa-r=a.*
boy-OBL-ERG 1.REFL-DAT bride.ABS search 2-do.IPFV-CVB=be
'The boy is looking for a bride.'

b. *je mij ješemiš du-ruʔu-r=a.*
we.ABS here live HPL-become.IPFV-CVB=be
'We live here.'

To the contrary, the lexical part *jaʁmiš* is 'labile' in the sense that it can co-occur with both intransitive and transitive light verbs, cf. *jaʁmiš hiкis* 'to gather together (intr.)' vs. *jaʁmiš haʔas* 'to gather, collect (tr.)'. There are other coverbs with similar behaviour, e.g., *bezmiš* which occurs both in a transitive compound *bezmiš haʔas* 'to annoy, make bored' (23c) and an intransitive one *bezmiš hiкis* 'to get bored'; with the latter, the source/ stimulus is encoded by a locative (superelative) case form (23d).

(23) a. *za-d bala uq' jaʁmiš hiʔi-r=a.*
I-ERG much hay.ABS gather 4.do.PFV-CVB=be
'I have collected much hay.'

b. *je gim-a jaʁmiš d-iši-r=a-j.*
we.ABS godekan-IN gather HPL-become.PFV-CVB=be-PST
'We gathered at the godekan (a gathering place for men in a village).'

c. *ha-now zi bezmiš hiʔi-r=a.*
that-OBL.H(ERG) I.ABS annoy 1.do.PFV-CVB=be
'He annoyed me.'

d. *zi ha-now-la bezmiš hiši-r=a.*
I.ABS that-OBL.H(SUP)-ELAT annoy 1.become.PFV-CVB=be
'I got sick and tired of him.'

The borrowed Russian verbs occur in the form of the Russian infinitive (in *-at*/ *-at^{j}*),[8] which is another pattern typical for Nakh-Daghestanian languages and many other languages of Russia (Forker and Grenoble, 2021, 261–263). Like the Azerbaijani *miš*-form, the borrowed Russian infinitives are acategorical and do not bear any other inflections. The light verbs are 'become' for intransitives and 'do' for transitives. The following examples are taken from the spoken text corpus.

(24) a. *ti xu-d-di mu^{ʕ}ɢ-u-la za-d perwij mesto*
that five-4-ATTR village-OBL(SUP)-ELAT I-ERG first[R] place[R].ABS
zanimat^{j} ha<w>i-r, sowχoz-a hixi-r
take[R] <3>do.PFV-CVB(AOR) sovkhoz[R]-IN 1.go_to.PFV-CVB
'Among the five villages, I took the first place (in a competition), when I went to sovkhoz (a collective farm in Soviet times).' (Text Corpus)

b. *xele ma^{ʕ}ʔalim-a-ra ha-bir nakazat^{j} d-aʔa-r-i-j*
after teacher-OBL-ERG that-PL.ABS punish[R] HPL-do.IPFV-CVB-COP-PST
'Then the teacher punished them.' (Text Corpus)

As both (24a) and (24b) show, the lexical parts of complex verbs *zanimat^{j} haʔas* and *nakazat^{j} haʔas* are not assigned any thematic role: both verbs have their own patient NPs, namely *perwij mesto* 'the first place' in (24a) and *habir* 'they' in (24b).

Unlike the compounds with Azerbaijani forms in *-miš*, which are relatively numerous and common in the text corpus, the compounds with Russian infinitives are relatively rare. Interestingly, only compounds with Azerbaijani forms in *-miš* can be found in the dictionary of Standard Rutul based on the Mukhad dialect (Alisultanov and Sulejmanova, 2019), whereas those with Russian infinitives are absent in this source. This may point at the perception of Russian verbs in Rutul compounds as instances of code-mixing and/ or imperfect command of one's native language. Indeed, at least some of such verbs can be regarded as occasional combinations used whenever a Rutul speaker finds appropriate to combine a Russian lexical part with a light verb instead of using a Rutul lexeme. Such compound verbs can be often seen in those contexts where the description of a situation includes a whole range of Russian borrowings, like in (24a): here, *perwij mesto* 'the first place' and *sowχoz* 'sovkhoz, collective farm' are also Russian loans, and the situation described is a workplace competition among the collective farms in late Soviet or post-Soviet period.

8 In Russian, the infinitive suffix *-t'* (-ть) is a palatalized /t^{j}/; in Rutul, however, borrowed Russian infinitives do not always keep the palatalization on the final consonant.

3.7 Verbal compounds with locative adverbs

Compounds that represent combinations of a preposed locative adverb and a verb stand apart from the rest of the group for two reasons. First, only in this type do we see the use of locative adverbs in verbal compounds. The adverbs used in compounds are mostly those with essive (locational) semantics, e.g. *u:* 'on top of', *a:* 'below', *ara* 'inside', or those with lative (directional) semantics, e.g. *la?* 'up', *sa?* 'down', *a?* 'inside', *ʁa?* 'outside', *xu?* 'in front'.

Second, the verbs used with locative adverbs are not the semantically general light verbs like 'do' or 'become' which we see in the other types of compounds. In the "adverb + verb" combinations, the verbs are mainly verbs with locative prefixes which already express a locational meaning by themselves. Thus, in all three compounds *la? luzas* ("up + stand") 'to get up; stand up', *a: lukas* ("below + lie") 'to lie down' and *xu? la?as* ("in front + throw") 'to vomit' the verb starts with *l-*, which is a locative prefix with the meaning 'up, on'. In a sense, in "adverb + verb" compounds, adverbs function as an additional layer of spatial marking specifying the spatial configuration of the event. In some cases, this yields semantically transparent combinations like *la? luzas* 'to get up; stand up' or *sa? sirxus* ("down + get down") 'to fall down', in which the adverb reinforces and specifies the semantics of the verb. Some combinations, like *xu? la?as* ("in front + throw on") 'to vomit' or *ara jiˤχas* ("inside + hit, kick") 'to collide, to intermix', are idiomatic. In any case, it appears that quite often the most natural way to describe a situation is by using an "adverb + verb" combination, and not the verb alone.[9]

The following examples illustrate the use of "adverb + verb" compounds *a: lukas* 'to lie down', *la? luzas* 'to get up; stand up' and *ara jiˤχas* 'to collide, to intermix'.

(25) a. *ha<w>i-r=xʷa* *kašir*, *mij-a:*
<3>do.PFV-CVB(AOR)=REP porridge.ABS here-ELAT
ha<w>i-r=xʷa, *mi-bir* *a:* *l-ü<t>kü-r*
<3>do.PFV-CVB(AOR)=REP this-PL.ABS below PV-<HPL>lie.PFV-CVB
s-e<t>χi-r=xʷa
PV-<HPL>sleep.PFV-CVB(AOR)=REP
'They made the porridge, and after they made it, they lied down to sleep.'
(Text Corpus)

9 Nasledskova and Netkačev (2020, 812) mention that in the combination *sa? ?irxus* 'to fall down' (the verb *?irxus* 'to get into, to appear' includes the locative prefix *?-* 'inside'), the use of the adverb *sa?* 'down' is optional in some cases, but is obligatory for the description of those situations when the starting point of falling is mentioned explicitly and it is higher than the final point.

b. *haj-a:* *dibir* *la-ʔ* *l-uzu-r* *ruxʷa-r=a=xʷa*
DEM-ELAT mullah.ABS up-LAT PV-1.stand.PFV-CVB 4.say.IPFV-CVB=be=REP
'Then the mullah stood up and said …' (Text Corpus)

c. *qʼʷaˤ-d čabɨl-eš-dɨ* *süri* *ara* *jiˤ<l>χi-r=a-j*
two-4 sheep-OBL.PL-ATTR flock.ABS inside <APL>hit.PFV-CVB=be-PST
'The two flocks of sheep mixed up.'

Note that combinations of one and the same adverb with two different verbs are not uncommon. Apart from the pair *saʔ sirxus* ("down + get down") ~*saʔ ʔirxus* ("down + get into") 'to fall down', cf. the pair with a difference in meaning in (26a)–(26b). Whereas *u: liʔir* ("on top + put on") 'to put on' is used to describe putting on of clothes (26a), its counterpart with a different verb bearing the same locative prefix, namely *u: luʁur* ("on top + lift on") 'to place onto' describes lifting something on an upper landmark, as in (26b).

(26) a. *za-d* *ʁijʁa* *bɨtʼrɨ-d* *uχun* *u:* *l-a<w>a-s-ɨ*
I-ERG today beautiful-ATTR dress.ABS on_top PV-<3>put_on-INF-COP
'I shall put on a beautiful dress today.'

b. *hejwan-a* *χinɨχ* *u:* *l-aʁʷ*
horse-OBL(SUP) child.ABS on_top PV-1.lift.IMP
'Mount the child on the horse!'

4 Properties of light-verb constructions in Kina Rutul

As I mentioned in the Introduction, it may be not so easy to draw a boundary between ordinary syntactic combinations (especially, 'patient NP + transitive verb') and complex verbs in the sense of compound (bi-componental) verbal lexemes, whose behaviour is different from ordinary syntactic combinations and which should be lexically listed. Thus, among the numerous combinations like "work do", "big become", "die do", "live become" or "down fall" there may be both just periphrastic ways of expressing certain verbal meanings, without special morphosyntactic behaviour, or compound lexical items, displaying at least some properties of single words.

In order to scrutinize this distinction, quite a few parameters of wordhood may potentially be taken into account, although not all of them may be applicable to every language. For example, Harris (2002, 76–87) in her study of complex verbs in Udi, argues that complex verbs in this language are single verbs with an 'incorporated' lexical part, by addressing the following criteria: complex verbs have the stress of a

single verb; complex verbs are written as single words in texts and dictionaries; all of the deverbal derivational processes (e.g., derivation of various non-finite forms and causatives) take complex verbs as input; the negative marker occurs before the whole complex verb; complex verbs do not incorporate phrases with modifiers; conjoining parts of complex verbs is impossible; gapping parts of complex verbs is impossible; parts of complex verbs are anaphoric islands (as they cannot be replaced with proforms). Although for Kina Rutul, a dedicated in-depth study involving the application of these (and, potentially, also other) criteria to all the groups of verbal compounds remains a task for the future, in the present section I will partially approach the problem of wordhood of Kina Rutul compounds, focusing on the syntactic and morphological fusion of their parts. In particular, I will look at whether the order of the components is fixed (by default, the lexical part is followed by the light verb) and whether the insertion of other lexical material between the two parts is allowed, whether the 'nominal' lexical component may be assigned a thematic role (e.g. the patient) or be modified, and whether lexical parts of compounds or light verbs can become morphologically bound and 'incorporated', turning former bi-componental structures into simple verb stems.

4.1 Word order

Describing complex verbs in Archi, also a Lezgic language, Chumakina (2016, 3600) states that "[s]yntactically, all types of complex verbs demonstrate the characteristics of a single word: the order of the parts is fixed (the lexical part is followed by the light verb) and the insertion of other lexical material between these parts is not, as a rule, allowed". Kerimova (2023) make a similar conclusion about the complex verbs in Lezgian, although showing at the same time that in Agul, a closely related language, it is possible both to change the order of components and to insert other material between them. Kina Rutul displays the same behaviour of complex verbs as described for Agul (see also Maisak and Ganenkov, 2016, 3580–3583): although, for all type of complex verbs, the default order is 'lexical part + light verb', the reverse order is judged acceptable by native speakers. Likewise, the insertion of other material between the parts did not make the respective examples ungrammatical:

(27) a. *za-d q'ile w-i?i-r=a-j maskow-di*
I-ERG read 3-do.PERF-CVB=be-PST Moscow-OBL(SUP)
'I studied in Moscow.'

b. *za-d maskow-di w-i?i-r=a-j q'ile*
I-I-ERG Moscow-OBL(SUP) 3-do.PERF-CVB=be-PST read

c. *za-d q'ile maskow-di w-iʔi-r=a-j*
I-ERG read Moscow-OBL(SUP) 3-do.PERF-CVB=be-PST

4.2 Thematic roles of coverbs

In the 'noun (absolutive) + light verb' combinations like *gʷalaχ waʔas* ("work do") 'to work' or *kumag hɬwɬs* ("help give") 'to help', the light verbs are transitive verbs that also function as independent lexical items. In their turn, the coverbs like *gʷalaχ* 'work, job' or *kumag* 'help' also function as independent nouns. As parts of the combinations mentioned above, these nouns seem to occupy the position of patient noun phrases (in the absolutive case), and they also control gender agreement on the verb. In sentences like (1), (14a)–(14d) or (16a)–(16e), there are no other nouns that might be alternative candidates for the patient role.

The situation is different for compound verbs like "side do", "piece do", "present do" mentioned in Section 3.2: although their coverbs *sur* 'side, half', *q'at'* 'part, piece' and *peškeš* 'gift' are also 'normal' nouns, it is other absolutive noun phrases that occupy the patient slot in (15a)–(15c). The nominal parts of the compounds *sur haʔas* 'to break, to tear (tr.)' and the like thus seem to represent syntactically 'incorporated' components, i.e. indispensable parts of complex verbs rather than free absolutive NPs. But let us have a look at the intransitive combinations with the same coverbs, e.g., *sur hɬkɬs* and *q'at' hɬkɬs* 'to break, to tear (intr.)', with an absolutive subject distinct from the nouns *sur* or *q'at'* (12a)–(13a). One may wonder, whether the lexical parts of such verbs actually fill the Theme slot of the intransitive verb 'become': "[Subject$_{ABS}$] becomes [Theme$_{ABS}$]", a construction that can be illustrated with (28). Then, in (12a) it is the trousers that become a 'half', and in (13a), it is the spade handle that becomes a 'piece'. Under such approach, the construction would be morphosyntactically canonical, like (28), and does not represent a complex verb as such.

(28) *riši=xa šu=xa duχtur-ar d-iši-r=a*
sister.ABS=ADD brother.ABS=ADD doctor-PL.ABS HPL-become.PERF-CVB=be
'Both the sister and the brother became doctors.'

The same logic can be applied to combinations with the verb 'do': constructions like *sur haʔas* and *q'at' haʔas* (15a)–(15c) can be seen as merely causative counterparts of "[Subject$_{ABS}$] becomes [Theme$_{ABS}$]" clauses. That is, 'I broke the bread (in two)' can be possibly represented as an instance of "[Agent$_{ERG}$] makes [Patient$_{ABS}$] a [Theme$_{ABS}$]", namely "I (Agent) made the bread (Patient) a piece (Theme)". This, again, gives us an opportunity to treat 'noun + verb' compounds as regular syntactic combinations and not as complex verbs in the narrow sense of bi-componental lexemes.

4.3 Modification of coverbs

We might expect that for those light verb constructions whose lexical parts are already 'incorporated' these lexical parts cannot be modified or bear any inflections. Although this it true for acategorical bound components, this is not so for 'noun + light verb' combinations. We find examples, where nouns in such compounds can be inflected for number (i.e. bear the plural suffix), as in (29a) and (29b), or be modified. Thus, in (29b) the noun 'work, job', as part of the "work + do" combination, is preceded by an adjective meaning 'hard', which modifies the noun. In such instances, the lexical component can again only be analyzed as the absolutive patient noun phrase and not as an incorporated part of a complex lexical item.[10]

(29) a. *muˁgʷ-a-d q'ix-dɨ q'aˁs-dɨ insan-aši-s za-d*
village-OBL-ATTR big-ATTR old-ATTR person-OBL.PL-DAT I-ERG
hɨxɨ-r kumag-bɨr haʔa-r-i-j
1.go_to.PFV-CVB help-PL NPL.do.IPFV-CVB-COP-PST
'I used to go and help the old people of the village, the elders.' (Text Corpus)

b. *gal-a d-ɨxɨ-r gʷalaχ-bɨr haʔa-r=a* […]
Gal-IN HPL-go.PFV-CVB job-PL.ABS NPL.do.IPFV-CVB=be
'They go to Gal (a valley in Azerbaijan) and do various kinds of work …' (Text Corpus)

c. *haj-a: qu-ʔ za-da: gʷalaχ* […] *juˁq'-dɨ gʷalaχ*
there-IN.ELAT back-LAT I-APUD.ELAT job.ABS hard-ATTR job.ABS
qa-w-aʔa-s ruʔu-r-diš
RE-3-do-INF 4.become.IPFV-CVB-COP.NEG
'After that I cannot do any hard work.' (Text Corpus)

4.4 Morphologically bound coverbs?

Discussing complex verbs in Agul, another Lezgic language, Maisak and Ganenkov (2016, 3582) claim that the class of complex verbs in Agul is not uniform in that "some of them are close to free syntactic combinations of verbs and object noun phrases, while others are lexicalized to a considerable degree and approach simplex verb stems". In particular, as regards the latter group, lexical parts of some compound

10 In (29c), the light verb 'do' bears a repetitive prefix *q-* 'again' (here, in the context of negation, it adds the meaning 'not anymore'). However, this prefix is so regular that we can probably speak here about a 'repetitive form' of the same verb 'do' and not of a different light verb. See also Section 4.4 on repetitive prefixation.

verbs can become so tightly fused, that it is not the light verb, but the whole 'coverb + light verb' complex that can serve as an input for derivational processes. For example, prefixes in Agul, including the productive repetitive prefix (meaning 'again' or 'backwards'), only attach to verb stems and, in case of compound verbs, to light verbs, as in (30a). In just a few compound verbs, however, it is possible to prefix the repetitive marker to the lexical part, as in (30b).[11] In both examples, the complex verbs *gunt' xas* 'to gather (intr.)' and *gunt' aq'as* 'to gather (tr.)' include the lexical part *gunt'*, which is a bound element, probably related to the noun *k'unt'* 'heap'. The light verbs are 'become' and 'do', which correspond to independent verbs of the language. The behaviour of *gunt'* in (30b) shows, that, as part of a complex verb, it became not only syntactically, but also morphologically fused with the light verb: former bi-componental combinations *gunt' aq'-* "gather do" and *gunt' x-* "gather become" have been reanalyzed (even if optionally) as single stems *gunt'aq'-*, *gunt'x-*, to which a verbal prefix can be added.

(30) a. *saje-wur-i-f-as gunt' q-aq'a-je ħupː-ar e.*
other-PL-OBL-APUD-ELAT gather RE-do.IPFV-PTCP:PRS sheep-PL.ABS COP
'These were the sheep which they used to collect from the others.'

b. *aχpːaj bagajmi qa-gunt'-xa-s-e wari.*
then in_the_morning RE-gather-become.IPFV-INF-COP all.ABS
'Then in the morning, everyone is getting together again.'

For Kina Rutul, we are not aware of similar cases of complex verbs becoming simplex verb stems. (The only exception might be the verb 'to sneeze', which will be treated in the next section.) All complex verbs known to us still represent two components, without morphological fusion. Even bound lexical parts like the Azerbaijani and Russian roots do not become reanalyzed as simply parts of verb stems. An obvious exception, of course, are periphrasis forms of Rutul verbs as parts of causative compounds (Section 3.5), which are naturally verbal and retain the inflectional potential (like gender agreement) they originally possessed as verb stems.

4.5 Morphologically bound or 'zero' light verbs?

The light verbs occurring in verbal compounds can potentially include morphologically bound elements that are not used elsewhere as autonomous verbal predicates. This situation is observed in Udi, where, apart from the verbs 'do', 'become' and

11 Examples (30a) and (30b) are taken from the unpublished spoken corpus of Huppuq' Agul, collected by Dmitry Ganenkov, Timur Maisak and Solmaz Merdanova in the 2000s.

'say', which have counterparts among independent lexical items, a handful of other light verbs can be identified. The light verbs of this latter group are not found as free verbs in the modern language, and only tentative etymologies can be suggested for them. For example, the light verb *-ec-* found in intransitive/ decausative compounds, belongs to the same inflectional type as motion verbs, and might be related to a motion verb 'go, come' historically, and the light verb *-d-* found in transitive/ causative compounds, is very probably related to the root 'give', which in modern Udi is only attested in a prefixed stem *tad-* 'give'; see Maisak (2008, 99–102) and Schulze (2016, 3569) for discussion.

In Kina Rutul, the verb 'to sneeze' may at first glance look as an instance of a complex verb with a bound light verb *urχas* combined with an ideophonic lexical part *čiʔ*. The putative light verb *urχas*, which possess an infixal gender agreement slot (31), does not occur elsewhere in the language. And for the Lezgic languages in general, it is not uncommon to express the meaning 'to sneeze' by means of a compound predicate with a phonetically similar ideophonic part associated with the sound of sneezing, see e.g. *ʕamči aq'as* "sneezing do" in Agul, *inči ap'uz* "sneezing do" in Tabasaran or *aʿnšaʿ bos* "sneezing say" in Archi. However, as it turns out, nothing can intervene between the (putative) components of the Kina Rutul verb, nor their linear order can be shifted. Given that both components are restricted to just this combination, one has to conclude that *čiʔurχas* in fact represents a simplex verb and not a compound.[12]

(31) *χɨnɨχ čiʔurχa-r=a/ rɨš čiʔu<r>urχa-r=a*
child.ABS 1.sneeze.IPFV-CVB=be girl.ABS <2>sneeze.IPFV-CVB=be
'The boy is sneezing.'/ 'The girl is sneezing.'

In some languages possessing complex verbs, it is even possible that the light verb 'disappears' as a result of phonological reduction, so that former compounds cannot be formally identified as bipartite at all. Thus, in Lezgian complex verbs including the light verb *awun* 'do' (e.g., *k'ʷalaχ awun* 'to work', lit. "work do") can occur in their full or reduced forms. In their full forms, the verb is present, cf. the perfective past tense *k'ʷalaχ awuna* 'worked'. In the reduced form, the root 'do' is not visible, cf. simply *k'ʷalaχ-na*, with the tense-aspect suffix *-na* added directly to the coverb (Haspelmath, 1993, 178). Similarly, the Udi light verb *b-* 'do' undergoes devoicing and can even drop if occurs in the middle of a consonant cluster, cf. *äš-e-b-sa* [work-3SG-do-PRS] 's/he

12 Interestingly, in Mukhad Rutul, which is a dialect relatively close to Kina Rutul, the verb 'to sneeze' is much shorter and does not even superficially look like a compound, cf. *čirχas* (Gender 1), *čirirχas* (Gender 2), *čibχas* (Gender 3), etc. (Alisultanov and Sulejmanova, 2019, 429).

works', but *äš-pː-sa* [work-do-PRS] or even simply *äš-sa* in case the personal marker does not occupy the position between the lexical part and the light verb.

In Kina Rutul, no similar instances of light verb reduction or loss have been found: both the coverb and the light verb are always 'full' and prosodically autonomous words. In the available dictionaries of the Mukhad dialect (Alisultanov and Sulejmanova, 2019) and the Ikhrek dialect (Dzhamalov and Semedov, 2006), both employing the Cyrillic script, parts of verbal compounds are written separately, as two independent words.

5 Conclusion

As I have tried to demonstrate in the present contribution, Kina Rutul, just like its closest relatives in the Nakh-Daghestanian family, possesses a rich number of verbal compounds, or light verb constructions, which represent combinations of light verbs, mostly 'do' and 'become', with adjectives, nouns (in the absolutive case, but also some other cases), locative adverbs, verb stems (in the causative construction only) and coverbs with the acategorical status. Among the light verbs, there are no verbs whose use is restricted to compounds, all the verbs can be also used as independent lexical items.

It is highly probable that the majority of verbal compounds, namely all combinations including nouns, adjectives and adverbs, can be analyzed simply as 'noun + verb', 'adjective + verb' and 'adverb + verb' combinations, respectively. They represent a periphrastic way of expressing certain verbal meanings, which in some languages can be expressed by single words. Thus, such compounds are only 'complex verbs' from the semantic point of view. It cannot be excluded, however, that a more fine-tuned analysis of the morphosyntactic behaviour of such combinations can reveal certain restrictions that would allow us to qualify verbal compounds of the respective types as 'complex verbs' in a more narrow sense, namely bi-componental single words. Likewise, for the causative 'verb stem + do' combinations, additional evidence is needed to see whether such combinations represent monoclausal or biclausal structures, i.e. whether they are indeed single words or syntactic constructions.

Still, there are genuine complex verbs in Kina Rutul, whose coverbs cannot be assigned a thematic role and are syntactically 'incorporated' elements of verb lexemes, only responsible for the expression of the lexical meaning. Such complex verbs include coverbs that are neither visibly nominal nor verbal, and thus should be best treated as acategorical. Diachronically, these converbs can represents former nouns or verbs which only survived in compounds (cf. the converbs in *χaʕr hɨkɨs*

'to learn', *q'ɨle haʔas* 'to read; study' or *masa hɨwɨs* 'to sell'), but also can be direct borrowings, especially of the Azerbaijani and Russian verbs (cf. *ješemiš hɨkɨs* 'to live' or *nakazatʲ haʔas* 'to punish'). It is the latter group of complex verbs with borrowed coverbs that can still be supplemented by new members. Whether other types of verbal compounds in Kina Rutul, or at least some representatives of these types, will be proved to represent complex verbs as well, compounds with acategorical coverbs will probably still remain at the core of the complex verb class.

Bibliography

Alekseyev, Mikhail E. 1994. Rutul. In Rieks Smeets (ed.), *The Indigenous languages of the Caucasus. Vol. 3. North East Caucasian languages. Part 2*, 213–258. Delmar: Caravam.

Alekseyev, Mikhail E. 2016. Rutul. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen and Franz Rainer (eds.), *Word-Formation: An international handbook of the languages of Europe. Vol. 5*, 3536–3545. Berlin: De Gruyter. https://doi.org/10.1515/9783110424942-027.

Alisultanov, A. S. and T. A. Sulejmanova. 2019. *Rutul´sko-russkij slovar´ [Rutul–Russian dictionary]*. Makhachkala: Alef.

Authier, Gilles. 2024. Ideophonic verb compounds in Archi. *Linguistic Typology at the Crossroads* 4(1). 325–380. https://doi.org/10.6092/issn.2785-0943/16518.

Bağrıaçık, Metin, Angela Ralli and Dimitra Melissaropoulou. 2015. Borrowing verbs from Oghuz Turkic: two linguistic areas. In Francesco Gardani, Peter Arkadiev and Nino Amiridze (eds.), *Borrowed morphology*, 109–136. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614513209.109.

Chechuro, Ilia, Michael Daniel and Samira Verhees. 2021. Small-scale multilingualism through the prism of lexical borrowing. *International Journal of Bilingualism* 25(4). 1019–1039. https://doi.org/10.1177/13670069211023141.

Chumakina, Marina. 2016. Archi. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen and Franz Rainer (eds.), *Word-Formation: An international handbook of the languages of Europe. Vol. 5*, 3595–3604. Berlin: De Gruyter. https://doi.org/10.1515/9783110424942-031.

Daniel, Michael, Nina Dobrushina and Timur Maisak. to appear. *Studies in Kina Rutul*.

Dirr, Adolf M. 1912. Rutul'skij jazyk [The Rutul language]. *Sbornik materialov dlja opisanija mestnostej i plemën Kavkaza* 42(3). 1–204.

Dzhamalov, Kamal È. and Semed A. Semedov. 2006. *Rutul´sko-russkij slovar´ (ixrekskij dialekt) [Rutul-Russian dictionary (Ikhrek dialect)]*. Moscow: ÈKON-Inform.

Forker, Diana and Lenore A. Grenoble. 2021. Some structural similarities in the outcomes of language contact with Russian. In Diana Forker and Lenore A. Grenoble (eds.), *Language contact in the territory of the former Soviet Union*, 259–287. Amsterdam: John Benjamins. https://doi.org/10.1075/impact.50.09for.

Ganenkov, Dmitry and Timur Maisak. 2020. Nakh-Dagestanian languages. In Maria Polinsky (ed.), *The Oxford handbook of the languages of the Caucasus*, 87–145. Oxford: Oxford University Press.

Harris, Alice C. 2002. *Endoclitics and the origins of Udi morphosyntax*. Oxford: Oxford University Press.

Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin: De Gruyter Mouton.

Ibragimov, Garun Kh. 1978. *Rutul´skij jazyk [The Rutul language]*. Moscow: Nauka.

Ibragimov, Garun Kh. 2004. *Rutul´skij jazyk. Sinxronija i diaxronija [The Rutul language: synchrony and diachrony]*. Makhachkala: DNC RAN.

Kerimova, Zarina R. 2023. *Aspekty grammatičeskoj struktury složnogo glagola lezginskogo i agul´skogo jazykov [Aspects of the grammatical structure of complex verbs in Lezgian and Agul]*. Master Thesis: RGGU, Moscow.

Kibrik, Aleksandr E. and Yakov G. Testelets. 1999. *Èlementy caxurskogo jazyka v tipologičeskom osveščenii [Elements of Tsakhur in a typological perspective]*. Moscow: Nasledie.

Klimov, Georgij A. and Mikhail E. Alekseev. 1980. *Tipologija kavkazskix jazykov [Typology of Caucasian languages]*. Moscow: Nauka.

Maisak, Timur. 2008. Glagol´naja paradigma udinskogo jazyka [The verbal paradigm of Udi]. In M. E. Alekseyev, T. A. Maisak, D. S. Ganenkov and Yu. A. Lander (eds.), *Udinskij sbornik. Grammatika. Leksika. Istorija jazyka*, 96–161. Moscow: Academia.

Maisak, Timur. 2020a. Grammaticalization in Lezgic (East Caucasian). In Walter Bisang and Andrej Malchukov (eds.), *Grammaticalization Scenarios: Cross-linguistic Variation and Universal Tendencies. Vol. 1: Grammaticalization Scenarios from Europe and Asia*, 309–360. Berlin: de Gruyter. https://doi.org/10.1515/9783110563146-006.

Maisak, Timur. 2020b. The verb paradigm in Kina Rutul. HSE Working papers. National Research University Higher School of Economics, Moscow. https://wp.hse.ru/data/2020/10/27/1359252160/95LNG2020.pdf, last access 27/9/2024.

Maisak, Timur and Dmitry Ganenkov. 2016. Aghul. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen and Franz Rainer (eds.), *Word-Formation: An International Handbook of the Languages of Europe, Vol. 5*, 3579–3594. Berlin: de Gruyter. https://doi.org/10.1515/9783110424942-030.

Makhmudova, Svetlana M. 2001. *Morfologija rutul´skogo jazyka [Morphology of Rutul]*. Moscow: Sovetskij pisatel'.

Müller, Peter O., Ingeborg Ohnheiser, Susan Olsen and Franz Rainer. 2016. *Word-Formation: An International Handbook of the Languages of Europe, Vol. 5*. Berlin: de Gruyter. https://doi.org/10.1515/9783110424942.

Nasledskova, Polina L. and Ivan V. Netkačev. 2020. K opisaniju semantičeskogo polja padenija v rutul'skom jazyke [Towards a description of the semantic field of falling in Rutul]. *Acta Linguistica Petropolitana* 16(1). 801–820.

Schulze, Wolfgang. 2016. Udi. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen and Franz Rainer (eds.), *Word-Formation: An International Handbook of the Languages of Europe, Vol. 5*, 3564–3578. Berlin: de Gruyter. https://doi.org/10.1515/9783110424942-029.