

A sketch of Karata - Draft 3

Jérémy Pasquereau

November 2021

1 Introduction

1.1 Area and speakers

Karata (*kɜːʔirχi macɜːʔi*, Russian *karatinskij jazyk*) is an understudied Nakh-Daghestanian language originally spoken in 10 ‘auls’ (i.e. mountain-top villages) in North Western Daghestan. These villages are given in (1) in Karata and Russian¹ (Magomedbekova 1971; 2000). The number of Karata speakers is somewhere between 5,000 and 20,000.²

(1) Karata-speaking villages³

Karata name	Avar name	Russian name
<i>kɜːʔira</i>	kɜːʔaratʔa	Karata
<i>arči</i>	ʃarčo	Archo
<i>ančiq</i> ⁴	ʃančiq	Anchikh
<i>rač^waldi</i>	račabalda	Rachabulda
<i>maš:tada or maš:tajkʔa</i>	maš:tada	Mashtada
<i>čʔabaq:ʔara</i> ⁵	čʔabaq:ʔoro	Chabakoro
<i>raciχ:</i>	raciχ:	Ratsitlʔ
<i>enχ:eli</i> ⁶	enχ:elo	Enkhelo
<i>siukh or siwukh</i> ⁷	siukh	Siukh
<i>tʔukja</i>	tʔokitʔa	Tukita

1. I add Avar as, clearly, the Russian names are based on the Avar names.

2. The first and last census of the Karatas as a distinct people (from the Avars) dates back to 1926 (Kolga 2013). The number was then 5,305. Since then, the Karatas have been counted as Avars. Karata is an Andic language. The Andic branch forms with Avar the Avaro-Andic branch of the Nakh-Daghestanian language family. Karata and Avar are thus fairly closely related within the family, this, in addition to cultural and geographical promiximity may have motivated their classification as one cultural entity (Pasquereau 2009). Magomedova and Khalidova in their 2001 dictionary give the approximate number of 20,000 speakers. In the 2010 census (available at <http://www.gks.ru/>), the Karatas are counted as a subtype of Avars and the number given is 4,787. But this number can only be taken as an indication of the actual number since Karatas are free to declare themselves as Avars or more specifically as Karatas. Given that many consider themselves Avars (and this is in fact what used to be written in their passport), this number must be taken as a mere estimation.

This paper is based on the insights of the grammar published in Magomedbekova 1971 and Magomedova and Khalidova 2001 as well as on my own analyses. The data presented in this chapter come from several sources. I have used examples provided in the dictionary and others which I collected myself (in the field in 2011 and 2012, and working remotely with Karata speakers). I have also made much use of the texts provided in Magomedbekova 1971's grammar: I transcribed and analyzed the first ten texts provided in the grammar and Rashidat Khalidova checked and translated them into Russian.

1.2 Dialects

Karata is described in the literature as having two main dialects: the Tukita dialect spoken in one village, Tukita, and the Karata dialect spoken in nine villages including Karata, the head city of the *akhvakhskiy* administrative district. These nine villages each have their own variant which are mutually intelligible. By contrast, the Tukita dialect and the Karata dialect (as it is spoken in the village of Karata) are not mutually intelligible. The description presented in this paper is based on the variety spoken in the village of Karata.

1.3 Sociolinguistic situation

Karata is spoken by the whole population of (the village of) Karata, including children. The territory inhabited by the Karatas is part of a larger homogeneous cultural Sunni Muslim area which is mostly Avar and traditionally associated with the use of Avar as the regional *lingua franca*. Traditionally, everyone in Karata was fluent in Avar. While Russian progressively replaced Avar as the regional *lingua franca*, the knowledge of Avar was kept up by a few hours of Avar lessons in school but this is no longer the case. Nowadays, with exposure to television and internet, everybody is fluent in Russian.

Since Karata has no written tradition, its use is mainly oral although it is possible to write it using the Avar alphabet, Avar and Karata having similar sound systems. There are a few speakers of Karata who do write and publish in Karata, poetry and stories. In recent years, there seems to have been a significant increase in the awareness of Karata as an endangered language by the Karatas themselves and there is a noticeable increase in the amount of writing published both through traditional and non-traditional means (e.g. social media).

1.4 State of research

The first remarks on the Karata language are found in Erckert 1895 on the basis of the variety spoken in Archo. Dirr 1909 also dedicates a few pages to the Karata variety of Karata. The Tukita variety is first discussed in Bokarev 1938.

The study of Karata really takes off with the work done by Magomedbekova (Magomedbekova 1961; 1967), especially the monograph (Magomedbekova 1971) entirely dedicated

3. Other villages were created in Soviet times, e.g. for winter pastures. These are not listed here.

4. Magomedbekova 1971 gives *ančix:i*.

5. Magomedbekova 1971 gives the name *kaʔa enχ:eli* (Avar *t'as:a enχ:elo*, Russian *Verkhnee Enkhelo*, English *Upper Enkhelo*) but it was not recognized by the two consultants I consulted.

6. Magomedbekova 1971 gives the name *keχ:'i enχ:eli* (Avar *vorχ:'a enχ:elo*, Russian *Nizhnee Enkhelo*, English *Lower Enkhelo*) but it was not recognized by the two consultants I consulted.

7. My consultant tells me *siwukh* is what young Karata speakers say.

to describing the grammar of the language spoken in the village of Karata. To this day, it remains the standard reference on the grammar of the Karata language.

In 2001, a second reference work was published, the Karata-Russian dictionary (Magomedova and Khalidova 2001), and in 2017, a book of Karata tales was published (?). Much work has been carried out in Daghestan, e.g. Khalidova 2016 on lexicology and phraseology, Dalgatov 2015; 2017 on the variety spoken in Siukh, or still Xalidova 1990; Kuraeva 1998; Paxrudinova 2004; Gadžieva 2007; Kuraeva 2010; Umargadžieva 2010; Paxrudinova and Akhmedova 2014; Mallaeva and Xalidova 2016 on various aspect of Karata.

Unfortunately, since then, work on Karata has not progressed much. While Magomedbekova provides a good description of the morphology of the language, very little is known about its syntax (let alone its semantics, for instance the TAM system of Karata has still not been studied).

1.5 Language history

Karata is uncontroversially classified among the Andic sub-branch of the Daghestanian side of the Nakh-Daghestanian language family .

2 Phonology

2.1 Vowels and Consonants

2.1.1 Vowels

Karata has five contrastive vowel qualities (figure 1), each of which can be realized phonetically as short oral, long oral, short nasal, or long nasal.

Figure 1: Contrastive vowel qualities in Karata

	front	back
high	i	u
mid	e	o
low	a	

The extent to which vowels are nasalized varies: from lightly nasalized followed by a distinctly articulated nasal consonant to true nasal vowels with no trace of a consonantal constriction (as are found in French or Portuguese for instance). Vowels are more or less nasalized depending on their position in the word. Clear nasal vowels occur word-finally. In all other positions, if the phonetic nasal vowel is followed by a consonant in the coda of the same syllable, a nasal consonant may still be perceived.

- (2) a. d
 b. d

2.1.2 Consonants

Karata has both central and lateral consonants. Central sounds are articulated in various places whereas lateral sounds are all velar except /l/ which is dental. For the sake of conciseness, I present both central and lateral sounds in the same table (figure 2). In addition, all the phonemes (except bilabial and palatal phonemes) have labialized counterparts which are separate phonemes (see Pasquereau 2018 for arguments). A long consonant (C:) indicates a *fortis* (a.k.a. *strong*) consonant.⁸ The phonetic realization of the *fortis* feature depends on the manner and place of articulation of the consonant that bears it. It can be realized as frication, gemination or length but for some consonants (e.g. affricates) the realization of the *fortis* feature is better characterized as an increase in energy, in the sense that the articulators are more rigid. In keeping with the conventions adopted in this volume, I transcribe *fortis* consonants with the length symbol.⁹

8. In the literature on Nakh-Daghestanian languages, these consonants are called ‘strong consonants’. There is consensus on the fact that the strong character of a consonant is a phonological feature (Charachidzé 1981). Only voiceless consonants can have that feature. How this phonological feature is realized is however still to be investigated systematically.

9. It should be borne in mind though that since no phonetic measurement on Karata consonants has been performed, one should not conclude that all fortis consonants (transcribed as C:) are geminates.

Figure 2: Phonological table of Karata non-labialized consonants.

		Central							Lateral			
		Bilabial	Dental	Alveolar	Post-alv	Palatal	Velar	Uvular	Pharyngeal	Glottal	Dental	Velar
plosives	aspirated	weak	p	t	c	č	k					
	voiceless	strong			c:	č:	k:	q:				χ:
	glottal	weak	p'	t'	c'	č'	k'			ʔ		χ'
		strong			c':	č':	kt'	q':				
	voiced	b	d		dž	g						
fricatives	voiceless	weak		s	š			ɣ	h	h		ɬ
	strong			s:	š:	x:		ɣ:				ɬ:
	voiced			z	ž			ʁ		ʕ		
		nasals	m	n								
non-nasal sonorants			r			j	w					l

2.2 Script and transcription

Every person in the village I have met while on field trips knows how to read and write Russian at least, and most middle-aged and older people know how to read and write Avar. From their knowledge of the Avar writing system, they can read and write Karata. How used they are to reading and writing Karata varies greatly however. Many Karatas have a smartphone with internet access and many households have (at least had in 2012) internet access with a computer providing new opportunities to write in Karata. In addition a few Karatas write poetry and songs in Karata, while others do not have any use for it.

The Karata-Russian dictionary and the collection of texts recently published by Khalidova (2018) use the Avar cyrillic alphabet with a few additions to encode the consonants of Karata that have no equivalent in Avar (e.g. Avar and Karata both have the sounds / λ :/ – transcribed in both languages as $\lambda 1$ – and / λ :’/ – transcribed in both languages as $\kappa\bar{b}$. However, of the two languages, only Karata has the sound / λ :’/ for which Karata has the graph $\kappa\bar{b}1$).

Regarding nasal vowels, the dictionary consistently uses superscripted h (e.g. a^{h} for / \tilde{a} /) but when speakers write, they frequently omit indicating nasality at all. Strong consonants are rendered in at least three different ways: in the dictionary, strong consonants consistently have a macron above the consonant (e.g. $\bar{\kappa}1$ for / k :’/), but when Karatas write their language, one finds that they use different strategies, e.g. the whole symbol is doubled (e.g. $\kappa 1\kappa 1$ for / k :’/) or only part of it is doubled (e.g. $\kappa 11$ for / k :’/).

2.3 Phonotactics

The maximal syllable template is: C G V G C (C). Complex codas are found in loanwords, as in (3h) from Russian.

(3) List of syllable types and examples

	syllable type	underlying form	surface form	meaning
(a)	CV	hane	ha.ne	‘village’
(b)	CGV	dunj \bar{a} l	du. \bar{p} al	‘world’
(c)	GV	j-ah ^w a-e	ja.hwe:	‘she played’
(d)	CVC	bert’in	ber.t’in	‘cheese’
(e)	CVG	w-ož-ud-o-w	wo.ž \bar{u} .dow	‘trustful’
(f)	CVGC	ʃajb	ʃajb	‘guilt’
(g)	GV	waša	wa.ša	‘boy’
(h)	CVCC	port	port	‘harbor’

Sequences of labialized consonants in onset positions are generally not allowed. In addition, a number of assimilation phenomena contribute to obscuring the underlying form of morphemes (for more details see section 2.5, Magomedbekova 1971 and Pasquereau 2018).

2.4 Prosody

Both in the existing grammar (Magomedbekova 1971) and in the dictionary (Magomedova and Khalidova 2001), Karata is described as having lexical stress, i.e. contrastive stress. However, it is not obvious at this point in my research on Karata that this is indeed

the case.¹⁰In any case, the description of the grammar of Karata is not impacted by not indicating stress (if there is indeed such a phenomenon). For this reason, I do not indicate stress in this paper, though it is not a settled issue.

2.5 Morphophonemics

There is a number of phonological changes that occur when stems are affixed to stems. In this section, I give an overview of some of these changes.

A few Karata consonant-final stems, verbs and adjectives, come with a nasal feature which is realized either on a suffixal consonant if it is /b/, /l/, or /r/ or, in the absence of one of these consonants, on a vowel.¹¹ As the examples below show, the verb stem *-oʔ~* ‘go’ comes with a nasal feature which is realized on the suffix-initial vowel, but which in the participle, specialized converb (sp. *cvb*), and masdar forms is realized on the suffixal consonant making it nasal (see (4) illustrating forms of a stem without a nasal feature, and (5) illustrating forms of a stem with a nasal feature).

- (4) No nasalization, e.g. *-oq:-* ‘remove’

	underlying form	surface form
infinitive	b-oq:-aʎa	boq:ʎa
imperfective	b-oq:-ida	boq:ida
participle	b-oq:-o-b	boq:ob
sp. <i>cvb</i>	b-oq:-ala	boq:ala
masdar	b-oq:-e-r	boq:er

- (5) Nasalization, e.g. *-oʔ~* ‘go’

	underlying form	surface form
infinitive	b-oʔ~-aʎa	boʔãʎa
imperfective	b-oʔ~-ida	boʔĩda
participle	b-oʔ~-o-b	boʔom
sp. <i>cvb</i>	b-oʔ~-ala	boʔana
masdar	b-oʔ~-a-r	boʔan

The same is observed with adjectives.

10. There are a few ways in which the seeming lack of stress in Karata (in opposition to what is described in the dictionary and grammar) is yet plausible. First, it is not only the linguist writing these lines who has failed to hear stress, colleagues too have not heard any obvious stress system in the language (although of course their opinion is based on the listening of (just) one text and minimal pairs). Second, verifications of reported minimal pairs (in the grammar and dictionary) and conversations with the second author of the dictionary, who is a native speaker of the language, point towards a situation different from that found in other Daghestanian languages and more specifically in the more closely related Andic languages. For instance, the grammar of Bagvalal (Kibrik 2001) describes a situation in which Bagvalal words are divided into three sets when it comes to their behavior with respect to stress: (i) words with a clearly emphasized syllable (‘as in Russian’ according to speakers); (ii) words with a less clearly emphasized syllable which speakers call ‘weak stress’; and (iii) words with no stress whatsoever. It is of course possible that a similar partition exists in Karata and that I have either missed it or that I just happened to tap into class (iii) when working on accentual minimal pairs according to the Karata dictionary (Magomedova and Khalidova 2001). Unlike in Bagvalal, in Karata prosody seems to be entirely dependent on intonation.

11. For vowel-final stems that come with a nasal feature, e.g. *-aʎ:a~* ‘make similar’, one could alternatively say that the stems ends in a nasal vowel, say *-aʎ:ã-*, which transmits its nasality according to the conditions specified above.

(6) No nasalization, *řačo-* ‘green’

	underlying form	surface form
I	řačo-w	řačow
II	řačo-j	řačoj
III	řačo-b	řačob
IV	řačo-baj	řačobaj
V	řačo-raj	řačoraj

(7) Nasalization, *c’ijo~*- ‘big’

	underlying form	surface form
I	c’ijo~-w	c’ijōw
II	c’ijo~-j	c’ijōj
III	c’ijo~-b	c’ijom
IV	c’ijo~-baj	c’ijomaj
V	c’ijo~-raj	c’ijonaj

Another morphophonological change is consonant labialization arising exclusively as a result of affixation of a morpheme to a stem.¹² In Karata, in vowel-initial verb stems (-V₁C(V₂)...-)¹³ that have a prefixal slot for a gender marker, C must be labialized when both of the following conditions hold:

- (i) /V₁/ is underlyingly¹⁴round and high, i.e. V₁ = /u/
- (ii) [V₁] is realized as an unround vowel as a result of its assimilation to the gender marker prefix added to the stem

In (8a), both conditions (i) and (ii) are met and it looks as though labialization shifts from /u/ to the stem C. For instance the consonant in the underlying stem /uč~/ is realized as labialized once the stem-initial vowel has assimilated (delabialized) to the prefixal feminine gender marker *j-*. However, in (8b), the underlying stem-initial vowel is round but it is not high and, upon assimilation of the vowel to /j-/, the following consonant is not labialized.¹⁵ This labialization change is observed across all forms of the paradigm where /u/ loses its labiality. In these examples (and in the rest of the paper), unless otherwise indicated, verbs are presented in the infinitival form, suffixed with the infinitival morpheme *-ařa* glossed ‘INF’.

(8) a. Labialization of stem consonant

12. For reasons of space, I cannot fully describe the patterns, see Pasquereau 2018 for more detail.

13. The majority of Karata verb stems that can undergo this process follows the pattern VC(V).

14. Non-low stem-initial vowels undergo a number of assimilation processes depending on the gender marker that precedes them. Vowels following the gender I marker *w-* are rounded, those following the gender I marker *j-* are unrounded/palatalized, and those following the gender IV and V markers *b-* and *r-* are lowered to [a] with certain verbs. Because only the neuter singular marker *b-* (gender III) can be followed by any of the five vowel qualities, the form of the verb stem following this marker is taken to be its underlying form. This form, the gender III or singular neuter form, is the citation form in the dictionary (Magomedova and Khalidova 2001).

15. At this point, the reader may very reasonably wonder why an analysis according to which the consonant is labialized underlyingly cannot be entertained. See Pasquereau 2018 for arguments that such an analysis is incorrect.

	Underlying form	Surface form	Meaning
(a)	j-uč̣~aɫa	jič̣ ^w ãɫa	bathe
(b)	j-uč̣'ã-aɫa	jič̣' ^w ã ɫa	open
(c)	j-uχ: ['] ~aɫa	jiχ: ^{'w} ãɫa	share

b. The stem consonant is not labialized

	Underlying form	Surface form	Meaning
(a)	j-oχ:a-aɫa	jeχ:a:ɫa	thrust
(b)	j-oɫ̣~aɫa	jeɫ̣ã ɫa	go
(c)	j-oχ: ['] -a ɫa	jeχ: ['] a ɫa	warm up

3 Morphology: word classes and inflection

3.1 Overview

Save for morpho-phonological processes, the morphology of Karata is agglutinative and the morphemes are easy to segment and recognize. Likewise, parts of speech are overall easy to identify, each presenting a distinctive set of morphological properties.

3.2 Nouns

3.2.1 Noun genders

The gender system of Karata includes three genders (a.k.a. classes in the Caucasological tradition) in the singular and two in the plural. The gender of non-derived nouns is not tied to specific markers in the citation form¹⁶ but their gender becomes apparent when they control agreement with adjectives, demonstrative pronouns, numerals, verbs and derived forms of verbs, and even adverbs. Table (9) shows the five gender/number markers: masculine (M), feminine (F), non-human (N), human plural (HPL), and non-human plural (NPL). As suffixes, the plural gender markers appear with the *-aj* formative.

(9) Gender markers (GM)

sg			pl	
morpheme	gloss		morpheme	gloss
/w/	M	male human	} /b(-aj)/	HPL
/j/	F	female human		
/b/	N	non-human ¹⁷	/r(-aj)/	NPL

The labels ‘feminine’ and ‘non-human’ are only an indication of the kind of nouns one can expect to find in each of these genders. For instance, the word *mak'e* ‘child’ is non-human. Gender affixes appear on the agreement targets as prefixes, suffixes and sometimes as infixes. This is illustrated in (10) with the non-human singular gender marker *b*, which is the gender marker that citation forms are given with (see section 3.6.2).

16. Nouns like *w-aša* ‘boy’ and *j-aše* ‘girl’ are perhaps remnants of an earlier stage of the language where nouns were marked for gender.

- (10) a. Prefix: *b-aλ'is:* ‘middle’
 b. Suffix: *bišdi-b* ‘your’
 c. Infix: *ha-b-al* ‘this’

3.2.2 Oblique stem and number marking

Nominal morphology is added to one of two stems: the nominative or the oblique stem. The absolutive (a.k.a. nominative in the tradition of North-Caucasian studies) is the citation form. The oblique stem is a stem that is, in most cases, distinct from the citation/nominative form of the noun. In addition, each of these stems can be in the singular or plural. Nouns can be inflected for plural number using a wealth of suffixes or vowel alternations (11) that are lexically specified.

(11) Plural suffixes with examples

Suffix (variants)	Example	Example-PL
-bi/-ibi/-abi	šunk'a ‘leg’	šunk'-ibi
-di/-idi/-adi	k'aze ‘shawl’	k'az-idi
-bdi/-ibdi/-abdi	miʁaλ:' ‘crupper’	miʁaλ:'abdi
-badi	wac:i ‘brother’	wac:ibadi
-baj	ɬic'el ‘finger’	ɬic'baj
-j/-aj/-ij	ming ^w a ‘spindle’	ming ^w aj
-i	mak'e ‘child’	mak'i
-li	unsa ‘ox’	unsali

The oblique stem of nouns is formed by adding a suffix or changing the last vowel of the citation form. Oblique stem formatives are distinct depending on the gender of the stem they are suffixed to. Typical oblique-stem formatives are given in (12). Note that while the oblique stem formatives of the human genders (i.e. singular masculine, singular feminine, and plural human) are clear, there is a lot of variation in the oblique formative of the non-human gender (both in the singular and plural).

(12) Common oblique-stem formatives

Masculine	-š:u
Feminine	-ɬ:i
Non-human	-V, -ɬ:i, -lo, -la
Human plural	-lo (-do with pronouns)
Non-human plural	-V (-ɬ:i with nom. adjectives, -da with pronouns, nom. numerals and adjectives.)

An example with *waša* ‘boy, son’ is provided in (13). The citation form/singular nominative form is *waša* and the derived singular oblique stem is *wašaš:u-*. Case suffixes are added to this form in the singular. From *waša* we can also derive the plural nominative form *wašaj* and from that form the plural oblique stem *wašajlo-* to which case suffixes can be added.

(13) Case-stem derivation of *waša* ‘boy’

	singular		plural
nominative	waša	→	waša-j
	↓		↓
oblique	waša-š:u-CASE		wašaj-lo-CASE

3.2.3 Case

3.2.3.1 The inventory of cases

Karata nouns inflect for case. The case inventory includes six grammatical cases and 20 spatial cases. Among the grammatical cases, the nominative (a.k.a. absolutive) is the unmarked case, which is also used as the citation form of nouns. The *-(o)l* morpheme is used to mark both the ergative and the instrumental cases. The two genitives have different distributions as is discussed later on. Note for now that genitive 1 consists of a gender marker (reflecting the gender of the head noun) directly added to the oblique form of the noun. The comitative suffix is *-k’el*.

(14) Case morphemes

‘Grammatical’ cases		Spatial cases
nominative	-∅	CFG + DIR
ergative/instrumental	-(o)l	1 -č’o- LOC -∅
genitive 1	-GM	2 -λ’a- ALL -r
genitive 2	-λ:(aj) ¹⁸	3 -a- ABL -gal
dative	-a	4 -χa-
(comitative	-k’el)	5 -q:-
		6 -i-
		7 -λ:i-
		8 -λ:’i-

3.2.3.2 The ergative/instrumental

The ergative case suffix is *-l* after a vowel, *-ol* after a consonant. It is used to mark the agent in transitive constructions (and very marginally in some intransitive constructions) as well as instruments and temporal adjuncts.

3.2.3.3 The genitive

As shown in (14), Karata has two genitive suffixes: genitive 1 with one of the gender markers, and genitive 2 with the suffix *-λ:(aj)*. The formative *-aj* is added when the possessum is plural. The genitive case suffix is used to mark dependent nominal phrases, like the possessor in a possessive construction. Which genitive suffix can be used depends on the gender/number of the possessor (see section 4.1).

18. The genitive 2 reflects the number of a plural possessum with the formative *aj*.

- (18) a. č'irq'aj hane ɬĩ-λ:i b-ac^w-a:
 Chirkai village water₀-CFG₇[LOC] N-flood-PF
The village of Chirkai flooded. (lit. got plunged in water)
- b. karš:i-λ:i-r c:ʔãji t'ama
 porridge-CFG₇-ALL salt throw.IMP
Put salt in the porridge!
- c. ɬĩji-λ:i-gal b-oq:-e herk'a-m č'ina
 water₀-CFG₇-ABL N-extract-PF big-N LOG
They¹⁹ extracted a big log from (inside) the water.

Spatial cases are also used more abstractly with non-spatial meaning and as part of the argument structure of predicates (Pasquereau 2011). For instance, the locative of CFG₅ is used to express ‘in exchange for’ (19a), and the locative of CFG₇ is used to encode one of the terms in a multiplication (19b).

- (19) a. x^wani-q: ič^wa χisa:ɬa
 horse₀-CFG₅[LOC] mare exchange.INF
Exchange a mare for a horse.
- b. boʔoda-λ:i ĩnštuda λi'abajbar k'ijac'ada bik'uda
 four-CFG₇[LOC] five multiply:COND twenty N:be:IPF
If you multiply five by four, you get twenty.

3.2.3.7 Other cases (or case-like forms)

Case-like suffixes in this section attach to non-oblique forms of nouns. These are the functive-transformative case (Creissels 2014) *-ɬe*, the directional suffix *-χ:il* which indicates spatial motion *towards* when combined with a locative form or *from* when combined with an ablative form. It does not appear on allative forms. More research is needed to pin down its meaning and morpho-syntax. In particular, it does not seem to be part of the same paradigm as ‘spatial cases’ since it seems to be able to be added to any case marked form. It is cognate with Chamalal *χ:in* (?) and probably cognate with the Avar directional suffix *χ:un* (Magomedbekova 1971, 175).

- (20) a. ɬanda-χ:il
 near-DIR
 ‘to the closest side’ (Magomedbekova 1971: 175)
- b. ɬanda-χ:il-a-gal
 near-DIR-CFG₃-ABL
 ‘from the closest side’

19. This example does not contain an overt subject. It is indeed very common in Karata to leave out arguments if it is recoverable from context or if the argument in question is meant to be very general.

- (21) hadoʔa-l č'am-e-b-oχa, boʁon ẽχ:e-λ:i-χ:il (ẽχ:e-λ:i-r) t'am-e
 head-ADD chew-PF-N-PF.CVB pig stream-CFG7-DIR stream-CFG7-ALL throw-PF.CVB
 idja.
 COP

He [the fox] chewed his head [the pig's], and threw the pig into the river. (Magomedbekova 1971, text 10, line 12).

- (22) mesdo=λ'ibar sinl hark'e-ɬe j-eɬ-a idja.
 Mesedo=? bear-ERG wife-FUNC F-lead-PF.CVB COP

The bear married Mesedo. (Khalidova 2017, text Xanšuj Mesdo, line 8).

3.2.4 Definiteness

There are no known markers of definiteness.

3.2.5 The formation of nouns

3.2.5.1 Nominal compounds

By far the most common way to form nominal compounds is by juxtaposing two nouns. The two nouns can be independently-meaningful as in (23).

- (23) a. *bazar-zebu* ‘day of the market/Sunday’ from *bazar* ‘market’ and *zebu* ‘day’
 b. *ʕamal-ʕadlu* ‘behavior’ from *ʕamal* ‘character’ and *ʕadlu* ‘discipline’
 c. *kunt'a-hark:e* ‘family’ from *kunt'a* ‘husband’ and *hark:e* ‘wife’

Other noun-noun compounds are composed of one independently-meaningful noun and another noun which is not found outside of the compound (24).

- (24) a. *χabalo-haλ:u* ‘wool comb’, from *haλ:u* ‘comb’
 b. *alaj-bulaj* ‘pretention’, from *alaj* ‘pretention’²⁰
 c. *adab-χatir* ‘respect, politeness’, from *adab* ‘respect, politeness’

I am aware of two nouns that result from the compounding of a noun and an adjective, they are the proper names in (25).

- (25) a. *k^waj-herk'am* ‘Big Dipper (constellation)’ from *kwaj* ‘scoop, ladle’ and *herk'am* ‘big’
 b. *k^waj-mik'ob* ‘Little Dipper (constellation)’ from *kwaj* ‘scoop, ladle’ and *mik'o-b* ‘small’

20. This could be the result of a reduplication process.

3.2.5.2 Nominal derivation

Two suffixes are used fairly productively to form nouns from various parts of speech. The suffix *-ɫer* forms abstract nouns from nouns (26a), adjectives (26b), adverbs (26c).

- (26) a. *č oq'ʔa* ‘friend’ + *-ɫer* = *č oq'ʔaɫer* ‘friendship’
 b. *džigaro-b* ‘energetic’ + *-ɫer* = *džigarɫer* ‘energy’
 c. *sveruq'ʔ* ‘around’ + *-ɫer* = *sveruq'ʔɫer* ‘the surroundings’

The suffix *-q:an* forms nouns of profession from other nouns (27).

- (27) a. *dane* ‘drum’ + *q:an* = *daneq:an* ‘drummer’
 b. *lajla* ‘religious service’ + *q:an* = *lajlaq:an* ‘religious singer?’
 c. *paxxu* ‘copper’ + *q:an* = *paxxuq:an* ‘welder?’

The most important source of derived nouns is verbs however. So-called *masdars*, i.e. deverbal nouns, are derived by adding the suffix *-r* to the past (a.k.a perfective in the Caucasological tradition) form of the verb. The derived noun retains the valency frame of the source verb. Participles can be nominalized via the addition of one of the oblique formatives.

3.3 Adjectives

3.3.1 Adjectival inflection

One can distinguish three groups of adjectives: adjectives with both a prefixal and suffixal gender marker (*b-eč'et'iro-b* ‘black’), adjectives with a suffixal gender marker (*ɣačo-b* ‘green’), and finally adjectives with no gender marker (*q:ajiλ:* ‘blue’). As far as I am aware, there are no adjectives that have a prefixal gender marker but no suffixal gender marker.

- (28) a. *b-ač'ak'o-b* ‘short’ b. *ɣačo-b* ‘green’ c. *q:ajiλ:* ‘blue’

All adjectives have the same inflection. In combination with a noun they modify, they express gender agreement with their head provided they have a gender marker. When the noun referring to the notion they modify is not expressed, adjectives occupy the last position in the noun phrase and are inflected for gender and case. In the nominative, they take one of the (suffixal) gender markers whereas in the other cases, they take an oblique stem formative, followed by the case marker. The oblique marker for feminine and non-human adjectives is *-ɫ:i*, it is added *after* the gender marker of the non-human plural gender.

- (29) Case inflection of the nominalized adjective *hac'a-GM* ‘white’

	M	F	N	HPL	NPL
NOM	<i>hac'a-w</i>	<i>hac'a-j</i>	<i>hac'a-b</i>	<i>hac'a-baj</i>	<i>hac'a-raj</i> ²¹
ERG	<i>hac'a-š:u-l</i>	<i>hac'a-ɫ:i-l</i>	<i>hac'a-ɫ:i-l</i>	<i>hac'a-lo-l</i>	<i>hac'a-raj-ɫ:i-l</i>
GEN	<i>hac'a-š:u-b</i>	<i>hac'a-ɫ:i-λ:</i>	<i>hac'a-ɫ:i-λ:</i>	<i>hac'a-lo-b</i>	<i>hac'a-raj-ɫ:i-λ:</i>
DAT	<i>hac'a-š:u-wa</i>	<i>hac'a-ɫ:i-ja</i>	<i>hac'a-ɫ:i-ja</i>	<i>hac'a-lo-wa</i>	<i>hac'a-raj-ɫ:i-ja</i>
COM	<i>hac'a-š:u-k'el</i>	<i>hac'a-ɫ:i-k'el</i>	<i>hac'a-ɫ:i-k'el</i>	<i>hac'a-lo-k'el</i>	<i>hac'a-raj-ɫ:i-k'el</i>

3.3.2 The formation of adjectives

3.3.2.1 Adjectival compounds

Adjectives can be created by compounding two adjectives together (30).

- (30) a. *anzeʔo* ‘snow-like’ + *hac’a-b* ‘white’ = *anzeʔo-hac’ab* ‘white like snow’
b. *c’ajʔo* ‘fire-like’ + *hero-b* ‘red’ = *c’ajʔo-herob* ‘red like fire’
c. *minaro-b* ‘other’ + *c’ijo-m* ‘new’ = *minarob-c’ijom* ‘new, fresh’

3.3.2.2 Adjectival derivation

Adjectives in Karata can be derived from nouns by means of the suffix *-ʔo* as in *waša-ʔo-b* ‘like a son’ from *waša* ‘son’, the privative suffix *-q:’u-b* as in *unš:i-q:’u-b* ‘landless’ from *unš:i* ‘land’, or the similitive suffix *-godo-b* as in *ila-godo-b* ‘like a mother’ from *ila* ‘mother’. The propriative suffix *-χ:u-b* derives adjectives from both nouns (a) and other adjectives (b).

- (31) a. *šiw-χ:u-b*
milk-ADJZ-N
milky (of a cow)
- b. *hero-χ:u-b*
red-ADJZ-N
reddish

Karata makes productive use of the genitive form of nouns with the suffix *-λ:* in the function fulfilled in other languages by relational adjectives, for example *q:aj-λ:* ‘blue’ from *q:aj* ‘dark clay’. Adjectives equivalent to English ‘such’ can be derived from demonstratives by adding the suffix *-šdo-b* to demonstrative stems, e.g. *hošdo-b*.

Participial morphology very productively derives adjectives from verbs, for instance *g^wanB-o-b* ‘light’ from *g^wanBaʔa* ‘become light(er)’. Finally Karata productively uses the suffix *-(i)s:* to derive adjectives from adverbs and nouns. Interestingly, this suffix can be added to case-marked nouns, converting them into noun modifiers as in (32).

- (32) *han-λ:i-s:* hawa
village-CFG₆[LOC]-ADJZ climate
climate like in a village = mountain climate

3.4 Pronouns

3.4.1 Personal pronouns

Karata has personal pronouns for first and second person, with an inclusive/exclusive distinction for the 1 person plural pronouns. The function fulfilled in other languages by specialized third person pronouns is fulfilled in Karata by demonstratives. Singular first and second person pronouns have a distinct, suppletive oblique stem. There is however no distinction of the direct/oblique stem in plural pronouns. Note that, exceptionally, the ergative suffix of the first/second singular pronoun is added to the nominative stem.

21. Magomedbekova 1971 (p. 89) notes that plural non-human forms can be formed with the oblique formative *-da*, thus NOM *hac’a-raj*, ERG *hac’a-da-l*, GEN *hac’a-da-λ:*, DAT *hac’a-da-a*. She does not provide the comitative form.

(33)	First and second person personal pronouns					
	1 SG	2 SG	1 PL INCL	1 PL EXCL	2 PL	
Nominative	den	men	iʎi	iš:i	biš:di	
Ergative	den-a	men-a	iʎi-l	iš:i-l	biš:di-l	
Genitive	di-GM	du-GM	iʎi-GM	iš:i-GM	biš:di-GM	
Dative	dij-a	duw-a	iʎij-a	iš:ij-a	biš:dij-a	
Comitative	di-k'el	du-k'el	iʎi-k'el	iš:i-k'el	biš:di-k'el	

The suffix *-da* can be added to personal and demonstrative pronouns to emphasize the identity of the referent of the pronoun as in *den-da*.

3.4.2 Demonstrative pronouns

There is a range of demonstratives in Karata. They can situate their referent according to two parameters: its height and distance relative to the speaker. They are all based on the roots *ha* (proximal), *ho* (medial), and *hu* (distal), alone or enlarged by one of the following three formatives expressing vertical deixis: *-ʎ* (higher), *d* (same level), or *-g* (lower than the deictic center). The suffixal gender marker reflects the gender of the lexical item associated with the referent of the demonstrative.²²

(34)	Demonstratives pronouns in the nominative ²³			
		proximal	medial	distal
	no height specification	ha-b	ho-b	hu-b
	higher	ha-ʎ-ib	hu-ʎ-ub	wu-ʎ-ub
	same level	ha-d-ib	hu-d-ub	wu-d-ub
	lower	ha-g-ib	hu-g-ub	wu-g-ub

Demonstrative pronouns inflect for gender and case. For all cases other than the nominative, case suffixes are added to their oblique stem, formed via *-šru* (M), *-ʎri* (F, N), *-do* (HPL), or *-da* (NPL).

(35)	Case inflection of the demonstrative pronoun <i>wudu-GM</i> ‘this one’				
NOM	wudu-w	wudu-j / wudu-b	wudu-baj	wudu-raj	
ERG	wudu-š:u-l	wudu-ʎ:i-l	wudu-do-l	wudi-da-l	
GEN	wudu-š:u-b	wudu-ʎ:i-ʎ:	wudu-do-b	wudi-da-ʎ:	
DAT	wudu-š:u-wa	wudu-ʎ:i-ja	wudu-do-wa	wudi-da-ja	
COM	wudu-š:u-k'el	wudu-ʎ:i-k'el	wudu-do-k'el	wudi-da-k'el	

3.4.3 Possessive pronouns

Possessive pronouns are formed by taking the genitive 1 form of the relevant pronoun or demonstrative; the suffixal gender marker cross-referencing the gender of its head noun – *di-b* ‘my’, *du-b* ‘your’, *iʎ:i-b* ‘our (incl)’, *iš:i* ‘our (excl)’, *biš:di* ‘your (pl)’. The way the

22. For demonstrative determiners, see section 4.1.

23. Different authors give slightly different forms for the medial and distal demonstratives. Magomedbekova 1971; 2000 give the forms indicated in (34) but in Magomedova and Khalidova 2001 two forms are found: forms with a suffix *-ib* are given (e.g. huʎib, hudib, hugib, wuʎib, wudib, wugib) in the grammatical sketch at the end of the dictionary whereas forms with the suffix *-ub* are listed in the dictionary.

ergative; genitive and dative cases are formed depends on the gender of the (unexpressed) head noun: if it is singular (masculine, feminine or neuter), the ergative; genitive; and dative exponents are added to the oblique form of these pronouns, which is formed by adding *š:u* if the head noun is masculine, *ɫ:i* if it is feminine or neuter. If the head noun is plural, the oblique formatives – *lo* for humans and *ɫ:i* for non-humans – are added to the gender markers. In the genitive, the gender marker reflects the gender of the head noun.

(36) Possessive pronouns *di-GM* ‘mine’

NOM	di-w	di-j / di-b	di-baj	di-raj
ERG	di-š:u-l	di-ɫ:i-l	di-baj-lo-l	di-raj-ɫ:i-l
GEN	di-š:u-GM	di-ɫ:i-GM	di-baj-lo-GM	di-raj-ɫ:i-GM
DAT	di-š:u-wa	di-ɫ:i-ja	di-baj-lw-a	di-raj-ɫ:i-j-a

An example of a possessive pronoun is given in (37): the first person singular pronoun in B’s response to A has the oblique marker *-ɫ:i-* because it refers to B’s car – the word for *car* is neuter – and is in the ergative case, as required by the predicate *awara gahaɫa* ‘have an accident’.

(37) A: Iš:i-b mašina b-iɪ-e idja
 1PL-N car stop COP
Our car stopped.
 B: Di-ɫ:i-l awara ge:
 1SG-OBL-ERG accident did
Mine had an accident.

(38) A: jac:o-λ: mašina b-iɪ-e idja
 sister-GEN car stop COP
My sister’s car stopped.
 B: Di-ɫ:i-b mašino-l awara ge:
 1SG-OBL-GEN car-ERG accident did
Mine’s car had an accident.

3.4.4 Intensive and reflexive pronouns

First and second person reflexive pronouns are formed by suffixing the morpheme *-da* to pronouns and demonstratives after the gender or case suffix.

(39) a. den w-ok’-ã du-χa-r
 1SG M-look-PF 2SG₀-CFG₄-ALL
I looked at you.
 b. den w-ok’-ã di-χa-r-da
 1SG M-look-PF 1SG₀-CFG₄-ALL-INT
I looked at myself.

For the third person reflexive pronoun, the intensive suffix *-da* is added to the pronoun *že-b*, with the suppletive oblique stem *in-*, which is inflected like other demonstrative pronouns.

(40) Logophoric pronoun

NOM	že-w	že-j / že-b	že-baj	že-raj
ERG	in-š:u-l	in-ł:i-l	in-do-l	in-da-l
GEN	in-š:u-b	in-ł:i-ł:	in-do-b	in-da-ł:
DAT	in-š:u-a	in-ł:i:j-a	in-dow-a	in-da-a (=inda:)

When not suffixed with *-da*, *žeb* is used as a long-distance reflexive, both in logophoric and non-logophoric contexts. In strictly local configurations (i.e. when pronoun and its antecedent are clause-mates), the suffix *-da* is required on any bound pronoun.

3.4.5 Reciprocal pronouns

The reciprocal pronoun is formed by combining two forms of *ce-b* ‘one’: one occurrence of *ce-b* bears the case of the controller of the reciprocal, the other occurrence of *ce-b* bears the case of the argument it fills.

(41)

bišdi	hał:u-j-da-l	hinc:u-č’o	bihi	ce-š:u-l	ce-š:u-č’o
2PL	seven-COLL-INT-ADD	door-CFG ₁ [LOC]	N-stand-IMP	one-OBL.M-ERG	one-OBL.M-CFG ₁ [LOC]
hadoʔa-bdi	un-e-mχ ^w a				
head-PL	touch-PF-N-[PF]CVB				

You seven stand by the door, hanging your heads against each other. (X1yЛОВШВАГАЛ
гоб мугъа, line 165.1)

3.4.6 Interrogative pronouns

Karata has the interrogative pronouns in (42) (see section 3.5.7 for wh-quantity words and section 3.7.1 for wh-adverbs). They must be used with one of the question particles (see section 4.3).

(42) Karata interrogative pronouns

<i>hede</i>	‘what’
<i>heme</i>	‘who’
<i>ło</i>	‘who’
<i>hinšto-b</i>	‘which’

The pronoun *hinštob* ‘which’, when used as a determiner, reflects the gender of the noun it takes as complement.

3.4.7 Indefinite pronouns

The numeral *ce-b* ‘one’ is used as an indefinite determiner and pronoun, with the suffixal gender marker reflecting the gender associated with its referent. It is also the basis for derived indefinite determiner/pronouns. The particle *-al* is added to *ce-b* to mean ‘one (of two)’, it can also be reduplicated to mean ‘some’.

- (43) gordi-λ: ce-b-al kwaɣa b-eɣ:ela-b idja
 dress-GEN one-N-? sleeve N-long-N COP
One sleeve of the dress is longer.

Interrogative pronouns (without the question particle) are also the basis for a number of indefinite pronouns. Indefinites can be formed by suffixing *bik'u(barel)* (e.g. *hede-bik'u(barel)* ‘someone’). The formative *bik'u* seems cognate with the past tense form of the verb *bik^waɬa* ‘be’.²⁴

- (44) men-a hede-bik'u q:’am-as:
 2SG-ERG what-INDEF eat-FUT
You will eat something.

In clauses whose main verb is negated, suffixing *ɣagi* (e.g. *hede-ɣagi* ‘anyone’) to wh-indefinites forms negative polarity pronouns.

- (45) den ɬo-k’el-ɣagi ɣurmi ge-da hač’e
 1SG who₀-COM-INDEF life do-IPF COP.NEG
I do not live with anyone.

3.5 Numerals

3.5.1 Cardinal numerals

The numeral system is decimal. Numerals from 1 to 10 are non-derived simplex stems and they all include the morpheme *-(V)da* except for *ce-b* ‘one’. Numerals from 11 to 19 are formed by adding the numerals from 1 to 9 to the word *hac’a-l* ‘ten’. The stem *hac’a-l* is the word for ‘ten’ minus the particle *da* and to which the particle *-l*²⁵ has been added.

Table 1: Counting from 1 to 19 in Karata

Basic numerals		Numerals from 11 to 19	
1	ceb	11	hac’a-l ceb
2	k’eda	12	hac’a-l k’eda
3	ɬabda	13	hac’a-l ɬabda
4	boʔoda	14	hac’a-l boʔoda
5	ĩš:tuda	15	hac’a-l ĩš:tuda
6	ĩλ:ida	16	hac’a-l ĩλ:ida
7	haλ’:ruda	17	hac’a-l haλ’:ruda
8	biλ’:ida	18	hac’a-l biλ’:ida
9	hač ^w ada	19	hac’a-l hač ^w ada
10	hac’ada		

24. This is reminiscent of how French forms indefinites from a wh-word, e.g. *quoi* ‘what’ and *que ce soit* as in *quoi que ce soit* ‘anything’, or how Spanish does the same from a wh-word, e.g. *quien* and *fuera* as in *quien fuera*. Thanks to Patricia Cabredo-Hofherr for pointing this parallel out to me.

25. This particle is presumably related to the additive particle.

Multiples of tens are formed by adding the formative *-ac'ada* – which is the word *hac'ada* ‘ten’ without the initial /h/ – to (*V*)*da*-less stems from 1 - 9. In case of hiatus, a glide is inserted, thus /k'e-ac'ada/ is [k'ejac'ada] ‘twenty’. Numbers are added to multiples of tens up to 99 by compounding the numerals from 1 - 9 to the *-ada*-less forms of multiples of ten. For instance, the word for ‘forty’ is composed of *boʔ-* – from *boʔoda* – and *-ac'* – from *hac'ada*. The word for ‘forty one’ is, as in English, the word for ‘forty’ minus the *-ada* ending – i.e. *boʔac'* – with the word for ‘one’ – *ce-b*.

Table 2: Counting in Karata

	Multiples of ten	Multiples of ten + 1-9
-	-	11 hac'al ceb
20	k'ej-ac'-ada	22 k'ejac' k'eda
30	ʃab-ac'-ada	33 ʃabac' ʃabda
40	boʔ-ac'-ada	44 boʔac' boʔoda
50	ĩš:t ^w -ac'-ada	55 ĩš:t ^w ac' ĩš:tuda
60	ĩʃ:ac'-ada	66 ĩʃ:ac' ĩʃ:ida
70	haʃ': ^w -ac'-ada	77 haʃ': ^w ac' haʃ':ruda
80	biʃ':ac'-ada	88 biʃ':ac' biʃ':ida
90	hač' ^w -ac'ada	99 hač' ^w ac' hač' ^w ada

The term for 100 is *bešanda* and multiples of 100 are formed by compounding (*V*)*da*-less forms of the basic numerals and *bešanda*. Numerals are further added to *da*-less *n*-hundred forms after they have been suffixed with the word *boʃ:e*²⁶. The word for ‘thousand’ is *ʃazar*. Multiples of 1000 are formed by combining the (*V*)*da*-less form of the basic numeral and the word for *thousand* without the initial pharyngeal. Further numbers are added to the thousand (multiples), the word *boʃ:e* is used after the *n*-thousand term.

Table 3: Counting from 100 in Karata

100	bešanda	1000	ʃazar
150	bešanboʃ:e ĩnš:t ^w wac'ada	2032	k'ejazar boʃ:e ʃabac' k'eda
200	k'ebešanda	5847	ĩnšt ^w azarboʃ:e biʃ':ibešanboʃ:e
267	k'ebešanboʃ:e ĩʃ:ac' haʃ':ruda		boʔac' haʃ':ruda
300	ʃabešanda		

Cardinal numerals can be nominalized and thus bear case inflection after suffixation of an oblique formative (46).

(46) Case inflection of the cardinal number *ce-GM* ‘one’

NOM	ce-w	ce-j / ce-b	ce-baj	ce-raj
ERG	ce-š:u-l	ce-ʃ:i-l	ce-ba-lo-l	ce-raj-da-l
GEN	ce-š:u-b	ce-ʃ:i-ʃ:	ce-ba-lo-b	ce-raj-da-ʃ:
DAT	ce-š:u-wa	ce-ʃ:i-ja	ce-ba-lo-wa	ce-raj-da-a
COM	ce-š:u-k'el	ce-ʃ:i-k'el	ce-ba-lo-k'el	ce-raj-da-k'el

26. The word *boʃ:e* might be related to the past tense form of the verb *boʃ:aʃa* ‘happen’.

3.5.2 Ordinal numerals

Ordinal numbers are formed by suffixing the morpheme $-\lambda:o-b^{27}$ to full forms of numbers (i.e. forms that include the suffix $-da$): $k'eda+\lambda:ob$ ‘second’ or $bešanda+\lambda:ob$ ‘hundredth’. The ordinal form of ceb ‘one’ is either the regular $ceb\lambda:ob$ or $heč'es:igis:eb$, which is composed of $heč'e$ ‘most’, $srigi$ ‘in front of’, the suffix $s:$ ‘ATTR’, and a gender marker.

- (47) wugu-w c':al-da ida hač'wada- $\lambda:o-b$ kalas:ij-a.
 DEM.LOW-M study-IPF COP nine-ORD-N class-CFG₃[LOC]
He is in the 9th grade.

Table 4: Ordinal numerals

1st	ce-b- $\lambda:o-b$
2nd	k'e-da- $\lambda:o-b$
3rd	ʃab-da- $\lambda:o-b$
4th	boʔoda- $\lambda:o-b$
5th	ĩnštu-da- $\lambda:o-b$ / ĩnšd ^w a- $\lambda:o-b$
6th	ĩn $\lambda:i$ -da- $\lambda:o-b$
100th	bešan-da- $\lambda:o-b$
1000th	azar-da- $\lambda:o-b$

Cardinal numerals can be nominalized and thus bear case inflection after suffixation of an oblique formative (48).

- (48) Case inflection of the ordinal number $ʃabda\lambda:o-GM$ ‘the third’
- | | | | | |
|-----|----------------------------|---|---------------------------|-------------------------------|
| NOM | ʃabda $\lambda:o-w$ | ʃabda $\lambda:o-j$ / ʃabda $\lambda:o-b$ | ʃabda $\lambda:o-baj$ | ʃabda $\lambda:o-raj$ |
| ERG | ʃabda $\lambda:o-š:u-l$ | ʃabda $\lambda:o-ʃ:i-l$ | ʃabda $\lambda:o-lo-l$ | ʃabda $\lambda:o-da-l$ |
| GEN | ʃabda $\lambda:o-š:u-b$ | ʃabda $\lambda:o-ʃ:i-\lambda:$ | ʃabda $\lambda:o-lo-b$ | ʃabda $\lambda:o-da-\lambda:$ |
| DAT | ʃabda $\lambda:o-š:u-wa$ | ʃabda $\lambda:o-ʃ:i-ja$ | ʃabda $\lambda:o-lo-wa$ | ʃabda $\lambda:o-da-a$ |
| COM | ʃabda $\lambda:o-š:u-k'el$ | ʃabda $\lambda:o-ʃ:i-k'el$ | ʃabda $\lambda:o-lo-k'el$ | ʃabda $\lambda:o-da-k'el$ |

3.5.3 Distributive numerals

Distributive numerals can be formed in two ways: by reduplication only or by (partial) reduplication and addition of the word *bik:e*.

- (49) a. $cebceb$ ‘one each’
 b. $cebce bik:e$ ‘one each’

3.5.4 Collective numerals

Collective numerals are formed by suffixation of $-(i)j$

27. The suffix $-\lambda:o-b$ might be related to the participle form of the verb $bo\lambda:a\lambda a$ ‘happen’.

(50) *ʔab-ij* ‘three, трое in Russian’, boʔoj, in[̃]s dij, inλ:ij, haλ:’uj, biλ:’ij, ʔ, hac’aj

The numeral *two* has an irregular collective form.

(51) k’enaj ‘двое’

3.5.5 Multiplicative numerals

The suffix *-c’e* added to the short form of a numeral *x* produces forms that mean ‘x times’: *bešan-c’e* ‘a hundred times’.

(52) *ʔabc’e* ‘three times’

The numeral for two is irregular in that it is derived from another stem *k’an-c’e* ‘twice’.

3.5.6 Year numerals

The morpheme *-ilja-b* is suffixed to the (*V*)*da*-less form of a numeral *x* to form an adjective that means ‘being x years old’ about animals.

(53) *inλ:-ilja-b* ‘being six years old’

3.5.7 Other quantity words

Other quantity words include those in (54) and (55).

(54) Quantifiers

- a. *mik’i* ‘little, few’
- b. *λ:’wani* ‘much, many’
- c. *ce-b-k’eda* ‘a few’

(55) Quantity wh-words

- a. *čāc’e* ‘how many times’
- b. *čami* ‘how much/many’

The suffix *-šdoʔo* is added to demonstrative stems to express ‘as much’ (e.g. *ho-šdoʔo(-b)*).

3.6 Verbs

3.6.1 Types of verb stems

Karata has both simple and complex verbs (compound and derived verbs). Bare verb stems cannot be used on their own: they must combine with affixes. All suffixes begin with a vowel. Some stems end in a consonant while others end in a vowel. This is, as far as I know, lexically specified.

There are two types of verbal stems in Karata: consonant-initial and vowel-initial stems (figure 3). C-initial stems never take a gender prefix while V-initial stems may: there is a lexical distinction between vowel-initial stems that have a morphological slot for a gender prefix and those that do not.

Figure 3: Verb stems in Karata (underlying forms)

Vowel-initial stems		Consonant-initial stems
VC		CVC
–VC	VC	
–ah- ‘buy’	aba- ‘sprinkle’	barka- ‘congratulate’
–iɪ- ‘stop’	ij- ‘attach’	bih˜- ‘heat up’
–eʔ- ‘ripen’	er- ‘lean on’	beχʔ- ‘plough’
–utʔ- ‘hollow out’	urɪ- ‘miss’	bur- ‘rise’
–oqʔ- ‘remove’	ob- ‘shake’	boršː- ‘peel’

Simple verb stems in Karata are overwhelmingly monosyllabic (VC(V) or CVC), occasionally bisyllabic, and rarely longer. Vowel-final stems end in an oral or nasalized vowel (see footnote ??). The TAM markers are suffixed to the verb stems.

3.6.2 Agreement and cross-reference

Gender/number agreement is limited to some verb stems beginning with a vowel. As shown below, some TAM suffixes include an optional or obligatory slot for a gender-number marker.

For instance, in (56) the gender marker on the verb and on the adverb reflects the gender of the nominative argument (in a. *waša* is gender I and in b. *roša* is gender III).

- (56) a. *waša* w-oχ:elał-er-iq: w-oʔ-ã
 boy M-lengthen-MSD-CFG₅[LOC] M-go-PF
The boy has grown.
- b. *roša* b-eχ:elałeri-q: b-oʔ-ã
 tree N-lengthen-MSD-CFG₅[LOC] N-go-PF
The tree has grown.

3.6.3 TAM, evidentiality, and polarity

TAM and negative polarity suffixes are added to the (light) verb stem. These are synthetic verb forms. In addition to these synthetic forms, Karata has analytic verb forms where

the lexical verb (in a converbial, participial, or infinitival form) is combined with another verb inflected for TAM categories. The copula *idja* is optionally added to these forms to signal indirect evidentiality.

3.6.3.1 Independent synthetic verb forms

The synthetic verb forms that can be the nucleus of independent clauses are characterized by the paradigm of suffixes listed in (57). Karata has suffixes for both past, present, and future temporal reference. In the tradition of Daghestanian studies, these are known respectively as perfective, imperfective, and future. They occupy the same slot and cannot combine.

(57) List of synthetic-form forming suffixes

	Positive	Negative
perfective	-a/i/e	-a/i/e-č'e
imperfective	-(i)da	-(i)da-č'e
future	-as:	-ibič'e
imperative	-a (tr) / -i (intr)	-ibis:e
optative	-o-b	-ibis:o-b

As far as suffixes are concerned, all verbs have the same suffixes except for the imperative suffix which makes a formal distinction between intransitive and transitive constructions (58).

- (58) a. ʃāda-χ:il-a-gal b-ah-a
 close-PROX-CFG₃-ABL N-take-IMP.TR
Remove it from the closest spot!
- b. miʃila-λ'i-gal χidi j-eʔ-ĩ
 sun₀-CFG₈-ABL away F-go-IMP.INTR
Go away from the sun! (lit. from under the sun)

Karata uses a suffix -č'e on verbs to negate them. Sometimes the negative form of a inflectional TAM suffix is suppletive (e.g. the negation of the copula *idja* is *hač'e*).

3.6.3.2 Independent analytic verb forms

In addition to the synthetic tenses listed in section 3.6.3.1, Karata also has a host of analytic verb forms formed by combining a lexical verb with the copula *idja* in the present or *bik^waʃa* 'to be' in the past or future tenses, or both to express aspectual and evidential distinctions. Analytic verb forms follow the pattern in (59).

- (59) lexical verb - auxiliary - evidential

The following generalizations hold. Aspectual distinctions are expressed by the form the lexical verb takes whereas temporal reference is set by the auxiliary. In addition, in the past (temporal reference), an evidentiality distinction is expressed: the addition of the copula *idja* conveys that the event was not witnessed. Future temporal reference

seems to be conveyed most often via the use of the verb *b-isãła* ‘turn out, find oneself’ but other verbs are also possible, e.g. *b-aq:ãła* ‘end’.²⁸

Table 5: List of analytic verb forms

	progressive	habitual
present	q: ^w ar-da idja	q: ^w ar-d-o-b idja
past	q: ^w ar-da bik ^w a	q: ^w ar-d-o-b bik ^w a
past unwit.	q: ^w ar-da bik ^w a(-b-χ ^w a) idja	q: ^w ar-d-o-b bik ^w a(-b-χ ^w a) idja
future	q: ^w ar-da bisās:	q: ^w ar-d-o-b bisās:
	perfect	prospective
present	q: ^w ar-e(-b-χ ^w a) idja	q: ^w ar-ała idja
past	q: ^w ar-e(-b-χ ^w a) bik ^w a	q: ^w ar-ała bik ^w a
past unwit.	q: ^w ar-e(-b-χ ^w a) bik ^w a(-b-χ ^w a) idja	q: ^w ar-ała bik ^w a idja
future	q: ^w ar-e(-b-χ ^w a) bisās:	q: ^w ar-ała bisās:

3.6.4 Negation

Negation is usually expressed by means of the suffix *-č’e* but it can also be expressed by a negative auxiliary (60). In addition, there are a number of suppletive negative forms, e.g. the future in an affirmative sentence is marked with the suffix *-as:* whereas, in a negative sentence, it is marked with the suffix *-bič’e*.

- (60) a. ho-b zini iš:i-b idja
DEM-N cow 1PL[GEN]-N COP
This cow is ours.
- b. ho-b zini iš:i-b hač’e
DEM-N cow 1PL[GEN]-N COP.NEG
This cow is not ours.

3.6.5 Valency-changing derivation

The only valence-changing derivation in Karata is the causative, productively formed via suffixation of the morpheme *-a-* directly to the verb stem or via the verb *itała* ‘let’ (e.g. *ħawała itała* ‘make burn’).

The morpheme *-a-* is often realized fused to the morpheme in the next verbal slot, i.e. tense, mood, participial or converbial morpheme, thus yielding diverse morpho-phonological phenomena (lengthening (61a), devocalisation (61b)).

- (61) Causative of *-iB-* ‘stop (intr)’
- a. Infinitive: b + iB+ a + ała = biB:ała
- b. Imperfective: b + iB+ a + idja = biB:a:jdja

28. These remarks reflect the knowledge of the system as per Magomedbekova (1971)’s observations and my own fieldwork data. In particular, I had the opportunity to elicit TMA-specific data using Dahl 1985’s TMA questionnaire as a result of which I obtained data that was not reported in Magomedbekova 1971 nor in Magomedova and Khalidova 2001. These remarks summarize my current knowledge of the system but further work is needed to fully understand it.

3.6.6 Dependent verb forms

Non-finite forms in Karata include participles, converbs, and infinitive (with the suffix *-aʎa*). There are three participles: participial forms are formed by suffixing *-o-b*, where *b* is a gender marker, to the perfective or imperfective stem of the verb thus giving a perfective versus imperfective participle. The prospective participle is formed by adding *ʎ:o-b* to the infinitival form or to the stem.

Table 6: Participles

Type of participle	Positive	Negative	Example
perfect	o-b		q:am-o-b
present	o-b		q:an-d-o-b
prospective	ʎ:o-b		q:am-ʎ:o-b / q:amaʎa-ʎ:o-b ²⁹

ADD SOMETHING ABOUT CONDITIONAL FORMS

Table 7: Conditional forms

Type of ????	Positive	Negative	Example
Conditional	[past stem]-bar	[past stem]-č'war	
Wishing present	[present ptcp]-xoror		
Conditional past	[past ptcp]-xoror	[past ptcp]-č'e-xoror	

Converbs are non-finite forms, similar to gerunds in English, which express a variety of relations between the event denoted by the verb and the main event. There are two main kinds of converbs: general and specialized converbs. General converbs only constrain the temporal orientation of the event denoted by the verb they mark relative to the main clause event.

(62) List of positive and negative general converb forms of *q:'amaʎa* 'eat'

	Positive	Negative
pf cvb	q:'am-e(-b-χ ^w a)	q:'am-e-ʁe(-b-χ ^w a)
impf cvb	q:'an-da-da	q:'an-da-ʁe

For instance in (63), the event of the putting on masculine clothes by the girls precedes the ploughing by the girls.

(63) jaš-i kunt'alo-b raʎ'ar-el b-aʎ'-a-b-aχ^wa beʎ':i-aʎa b-aʎ-ã idja
 girl-PL masculine-N clothes-ADD N-wear-PF-N-CVB plough-INF H+-go-PF COP

The girls went ploughing after they had put on masculine clothes (Text 3 Magomedbekova 1971)

29. Magomedbekova 1971 and Magomedova and Khalidova 2001 report that the suffix *-ʎ:o-b* must be added to the infinitival form of the verb. In my texts, I have also found *-ʎ:o-b* added to the stem of the verb.

Specialized converbs express more specific notions like causation with $-\lambda'oxa$ (see section 4.4.5).

- (64) List of positive and negative specialized converb forms of $q:'ama\acute{t}a$ 'eat'
- | | Positive | Negative |
|-------------|----------|----------|
| because ... | -gil | -č'e-gil |
| while ... | -laš | |

3.6.7 The formation of verbs

3.6.7.1 Verbal derivation (other than valency-changing)

The suffix $-t-$ can derive intransitive verbs from adjectives.

- (65) $q:aj\lambda:$ 'blue' + t + $a\acute{t}a$ = $q:aj\lambda:t'e\acute{t}a\acute{t}a$ 'become blue'

The suffix $-\chi:(^w)-$ can derive verbs from both adjectives and nouns.

- (66) a. $re\acute{t}ba-b$ 'barren' + $\chi:$ + $a\acute{t}a$ = $re\acute{t}ba\chi:^w a\acute{t}a$ 'become barren'
 b. $re\acute{t}e$ 'laughter' + $\chi:$ + $a\acute{t}a$ = $re\acute{t}e\chi:a\acute{t}a$ 'to laugh'

3.6.7.2 Verbal compounds

Some verbal compounds are composed of a noun and a verb. Some of the most common verbs are in (67).

- (67) a. $gaha\acute{t}a$ 'do' (e.g. $awara\ gaha\acute{t}a$ 'damage, lit. do damage')
 b. $t'ama\acute{t}a$ 'throw' (e.g. $rak^w a\ t'ama\acute{t}a$ 'be bored, lit. throw (the) heart')
 c. $baha\acute{t}a$ 'take' (e.g. $hox:el\ baha\acute{t}a$ 'breathe, lit. take air')

Other compounds from a noun and verb are such that the verb only exists in that compound, e.g. $hedela\ pila\acute{t}a$ 'have pimples'. Other verbs are formed from the morphological fusion of a noun and a verb as (68)³⁰.

- (68) a. eli 'mouth' + $b-a?a\acute{t}a$ 'to take away' = $elba?a\acute{t}a$ 'to warn someone'
 b. $kilo$ 'manure' + $b-oq:a\acute{t}a$ 'remove' = $kiloq:a\acute{t}a$ 'fertilise'

Just like new nouns can be obtained from compounding two nouns, new verbs can be obtained from compounding two independently-meaningful verbs (e.g. $s:ore-bik:a\acute{t}a$ 'surround' from $s:ora\acute{t}a$ 'spin' and $bik:a\acute{t}a$ 'hold') or two verbs only one of which has a meaning (e.g. $bit'e-bi\chi:a\acute{t}a$ 'get settled, solved' from $bit'a:t$ 'straighten up').

3.6.7.3 Verbal reduplication

Reduplication is very productive in Karata to derive so-called iterative forms of verbs (e.g. $\acute{i}k^w a\acute{t}a$ 'eat' and $\acute{i}k^w ak^w a\acute{t}a$ 'eat many times'). This area has, to my knowledge, not been studied at all for any of the Daghestanian languages that has it, although see Plungian 1997 for Chamalal.

30. Thanks to Denis Creissels (p.c.) for pointing out these examples to me.

3.7 Adverbs

3.7.1 Deictic, interrogative and indefinite adverbs

Non-derived adverbs are of different kinds. Spatial adverbs can be inflected for directional case while other adverbs (time, manner) do not.

(69) Deictic spatial adverbs

	close	medial	distal
higher	ha-ł-i	hu-ł-i	wu-ł-i
	ha-ł-i-r	hu-ł-i-r	wu-ł-i-r
	ha-ł-i-gal	hu-ł-i-gal	wu-ł-i-gal
same level	ha-d-i-	hu-d-i-	wu-d-i-
	ha-d-i-r	hu-d-i-r	wu-d-i-r
	ha-d-i-gal	hu-d-i-gal	wu-d-i-gal
lower	ha-g-i	hu-g-i	wu-g-i
	ha-g-i-r	hu-g-i-r	wu-g-i-r
	ha-g-i-gal	hu-g-i-gal	wu-g-i-gal

Deictic manner adverbs (equivalent to English ‘thus’) are formed from demonstratives by adding the suffix *-šda* (e.g. *ho-šda*). Other adverbs can be derived from adverbs (70a), adjectives (70b), and nouns (70c) using the configuration marker 5 in the locative *-q:*.

- (70) a. *χidi* ‘behind’ + *q:* = *χidiq:* ‘behind’
b. *b-it’o-b* ‘straight’ (adj) + *q:* = *b-it’oq:* ‘straight’ (adv)
c. *q’ino* ‘summer’ + *q:* = *q’inoq:* ‘in the summer’

Finally the manner, temporal, spatial, and causal interrogative adverbs are listed in (71).³¹

(71) Karata interrogative adverbs

<i>hinšda</i> ‘how’	<i>hinge</i> ‘where/to where’
<i>hinda</i> ‘when’	<i>hindir</i> ‘to where’
<i>hense</i> ‘why’	<i>hingal</i> ‘from where’
<i>heła</i> ‘why, what for’	

3.7.2 Case-inflected non-deictic spatial adverbs

In addition, some spatial adverbs have the same directional case inflection as spatial forms of nouns. Their identification as adverbs follows from the fact that they cannot be analyzed as including a stem combinable with non-spatial cases. The ordinary nouns used to express these concepts (e.g. ‘room’ *misa*, ‘hands’ *reλ’a*, ‘face’ *bała*) are not related to the spatial adverbs.

31. Quantity interrogative words are given in section 3.5.7 and pronominal interrogative words are given in section 3.4.6.

		LOC	ALL	ABL
(72)	‘in the room/house’	eši	eši-r	eši-gal
	‘in the hands’	k ^w adi	k ^w adi-r	k ^w adi-gal
	‘in the face’	hadi	hadi-r	hadi-gal

3.8 Postpositions

Karata has many postpositions (e.g. *caqʔa* ‘like’). Some postpositions of Karata are words that can also be used as adverbs, that is, they can be used on their own or they can be used with an NP complement. They are listed in (73).

(73)	Adverbs/Postposition		
	postposition/adverb	meaning	CFG of NP
	hini	in (hollow space)	CFG ₁ , CFG ₃ , CFG ₆ , CFG ₇
	kaʔa	on	CFG ₁ , CFG ₃ , CFG ₆
	keχ:i	in (substance)	CFG ₁ , CFG ₃ , CFG ₇
	baχ’i	between	CFG ₁ , CFG ₃ , CFG ₇
	keχ’i	under	CFG ₁ , CFG ₃ , CFG ₈
	s:igi	in front of	CFG ₁ , CFG ₃
	χigi	behind	CFG ₁ , CFG ₃
	χigi-s:igi	around	CFG ₁ , CFG ₃

When used with a complement NP, the postposition obligatorily inflects for directional case. Whether the NP inflects for a directional case is optional, but if it does, it must be in the same directional case as the postposition. In (74a), the word *kaʔa* is used as an adverb but in (74b) it is used as a postposition assigning dative case (or locative 3 case) to its complement NP (see Pasquereau 2011 for more details).

- (74) a. ho-w **kaʔa** w-uχ:-u
 DEM-M on M-end_up-PF
He wound up up there.
- b. **ustulj-a** **kaʔa** kuruška b-iłj-a
 table₀-CFG₃[LOC] ON cup N-put-IMP
Put the cup on the table.

3.9 Minor classes

Karata has interjections (e.g. *wuj* to express surprise, *hew*, *hej*, *jo* to attract someone’s attention) and ideophones (e.g. *q:^warq:^wari* ‘rustling’, *k^wark^wark^war* ‘the sound made when drinking’).

4 Syntax

4.1 Noun phrase and Postpositional phrase

Every noun phrase is marked for case. In general the head of a phrase is aligned to the right with its dependent phrase on the left with marking expressing the dependence

relation such as the attributizer *-s:* in (75).

- (75) hã-λi-s: hawa
 village-CFG₇[LOC]-ADJZ climate
village (mountain) climate

The main way to mark a nominal dependent is the genitive. As shown in (14), Karata forms the genitive in two ways: genitive 1 with one of the gender markers, and genitive 2 with the suffix *-λ:(aj)*. Genitive 1 is used to form the genitive form of personal pronouns (including demonstratives) and M and HPL nouns. In (76a), the word *maduhal* ‘neighbor’ is in the genitive 1 case as the gender marker *-j* is added directly to the oblique form built with the suffix *-š:u*, the gender marker *-j* reflects the gender of the possessum *jaše* ‘daughter’. The pronoun *iš:i* ‘we (exclusive)’ is also in the genitive form, the gender marker *-w* reflects the gender of the possessum *maduhalš:u-* ‘neighbor’ whose referent is masculine as indicated by the choice of the gender I oblique formative *-š:u*. Similarly, in (76b), the pronoun *ho* is in the genitive 1 case because the gender marker *-raj* is added directly to the oblique form built with the suffix *-ł:i*, the gender marker *-raj* reflects the gender of the possessum *mak’i* ‘children’.

- (76) a. iš:i-w maduhal-š:u-j jaše
 1PL.EXCL[GEN]-M neighbor-OBL-F girl
Our neighbor’s daughter.
- b. ho-ł:i-raj mak’-i c’aq’:a kuca: rχ^wa idja
 DEM-F₀[GEN]-NPL child-PL well breed-PF.CVB COP
Her children are well-bred. (Magomedova and Khalidova 2001)

The genitive 2 suffix *-λ:(aj)* is used to mark other possessors. The formative *aj* is added when the possessum is plural.

- (77) a. jaše-ł:i-λ: mak’e
 girl-OBL.F-GEN child
Daughter’s child.
- b. jaše-ł:i-λ:aj mak’-i
 girl-OBL.F-GEN.PL child-PL
Daughter’s children.

Variation in the order of the head and its dependent(s) and its possible effect on the information structure on the sentence need more research. See section 3.8 on postpositional phrases in Karata.

4.2 Clause structure

4.2.1 Constituent order

Karata clause structure is characterized by extreme flexibility of constituent order, which plays no role in the expression of argument structure. One notable exception is wh-questions where the wh-word must occur before the verb (Pasquereau and Khalidova

2017). In all clauses, the verb tends to occur in the clause-final position, but this is just a tendency. It is very common in Karata texts and speech to leave out arguments whose reference is recoverable from context, or indefinite or arbitrary. For instance in (78) the unique argument of *sleep* is not mentioned but its reference is recoverable from the context; the example in (79) illustrates this even more clearly as the logophoric pronoun in the embedded clause is necessarily interpreted with respect to the unmentioned individual whose speech is reported. Proverbs are typical cases where arguments are left out, e.g. (80) where the sentence is not about anyone in particular.

(78) Context: The merchant told himself:
 cemik'i maχ' u-l ga:s:
 a_little sleep-ADD do.FUT
I will sleep a little. (*X1yловшвагал гоб мугъа*, line 165.1)

(79) in-š:u-l it-ała hač'e-χ':e ke-χ':e.
 LOG-M.OBL-ERG let-INF COP.NEG-COMPL speak-PF
He said that he was not going to let us. (?)

(80) du-b-da metra b-eš:da hek'o-b reχ':u b-as-ĩdja.
 2SG-M.OBL[GEN]-N-INT meter N-leave-PRS.CVB foreign-N measure_unit-N N-measure-IPF
The pot calling the kettle black (lit. ignoring your own meter, you measure someone else's X, where X is the distance between the middle and little fingers of one hand)

4.2.2 Case alignment

4.2.2.1 Basic transitive and intransitive alignment

Karata has ergative alignment both in case marking and gender agreement. Karata has no person agreement. In transitive constructions, the agent is in the ergative case and is not indexed on the verb, whereas the patient is in the nominative case and controls verb agreement in gender. In (81a), the absolutive argument of the verb *b-ekr-ała* 'give' is the patient *k'eda waša* 'two boys' the gender of which, singular masculine, is reflected in the verb prefix. In intransitive constructions, the sole argument is in the nominative case and controls verb agreement in gender. In (81b), the verb *b-ekr^w-ała* 'happen, end up' shows the agreement prefix for 'singular masculine' which in this example is the gender of the unique argument *den* 'I'.

- (81) a. den rox:o-χ:i-r w-okr:u
 1SG forest.OBL-CFG7-ALL M-happen-PF
I wound up in the forest.
- b. kunt'^w-a den-a k'eda waša w-okr:e
 husband.OBL-DAT 1SG-ERG two boy M-give-PF
I gave my husband two sons.

4.2.2.2 Valency alternations

These basic codings/alignments can be altered by various valency alternations. As mentioned above, the only valency-changing morphosyntactically-marked derivation is the causative as marked by the suffix *-a* or the use of auxiliaries like *t'amata* 'throw'. In addition to these lexically marked alternations, Karata has a number of constructions which differ from the basic intransitive and transitive ones while not bearing any morphosyntactic marker on the verb.

In most cases, the effect of the causative is more often than not (there are exceptions) to create an ergative slot which is to be filled by a new participant assuming the role of an agent causing the action expressed by the verb. Causatives can be formed from intransitive (including experiential) and transitive constructions. The causative of an intransitive construction results in the addition of a causer in the ergative case as in (61).

- (82) a. c'elt'a b-iʔ-u
 plate N-break-PF
The plate broke.
- b. den-a urʔe:da b-iʔ^w-a: ci'āt'ur
 1SG-ERG on_purpose N-break-CAUS.PF plate
I broke the plate on purpose.

The causative of so-called experiential or mental-activity verbs, i.e. intransitive verbs whose experiencer is in the dative case, results in the addition of a causer in the ergative case as well (83).

- (83) a. dij-a χ:ex:a b-ečeč-ida hedela
 1SG.OBL-DAT fast N-forget-IPF thing
I quickly forget things.
- b. w-užu-w mak^wa b-ečeč-a:s: den-a duw-a
 M-grow_up-PF.PTCP-M place N-forget-CAUS.FUT 1SG-ERG 2SG.OBL-DAT
I'll make you forget where you were born.

Causativized transitives are attested but very rare. For more information on the causative, see Pasquereau 2011.

Although Karata has ergative alignment, it is possible in certain TAM configurations for the agent of a transitive construction to be coded in the nominative case. For instance, in analytical (progressive) forms, an argument normally marked in the ergative case can be marked in the nominative.

- (84) a. ustar-š:u-l q'in ge-da idja
 artisan-M₀-ERG wall do-IPF COP
The artisan is building the wall.
- b. ustar q'in ge-da idja
 artisan wall do-IPF COP
The artisan is building the wall.

4.2.2.3 Reflexive, reciprocal, anaphora

As shown in section 3.4.4, first and second person reflexive pronouns are formed by suffixing the particle *-da* to the corresponding personal pronouns. The third person local reflexive pronoun is formed by suffixing *-da* to *že-b* in the appropriate case form.

- (85) hugu-w w-ok'-ã in-š:u-χa-r-da
 DEM-M M-look-PF LOG-M₀-CFG₄-ALL-INT
He looked at himself.

The pronoun *že-b* can also be used as a long-distance reflexive or as a logophoric pronoun. In these uses, it does not need the suffix *-da*.

- (86) b-ax:ʷa-lada χan-š:u-l_i [in-š:u-b_i anɬer b-ič'a:č'e-b-χ^wa]-χ'e,
 HPL-come-SPCVB khan-OBL.M-ERG LOG-OBL.M[GEN]-N word N-destroy-NEG-N-CVB-COMPL
 keχ'ɾ-ē idja: [že-w_i x:ʷanij-a-gal ruš:t'an-č'e]
 say-PF COP LOG-M horse.OBL-CFG₃-ABL dismount-NEG
When we arrived, in order not to break his word, the khan said: I did not dismount my horse.

In reciprocal constructions, each element of the reciprocal pronoun (reduplicated *ce-b* 'one') bears the case of the arguments in the reciprocal relation.

- (87) a. mak'-i-lo-l kumak ge: jaš-i-lo-wa.
 child-PL-HPL-ERG help do.PF girl-PL-HPL-DAT
The children helped the women.
- b. mak'-i-lo-l kumak ge: **ce-š:u-l ce-š:ʷ-a.**
 child-PL-HPL-ERG help do.PF one-M₀-ERG one-M₀-DAT
The children helped each other.

4.2.3 Gender agreement & cross-reference

As was shown in section 3.2, nominative arguments are indexed on verbs. Karata also displays less usual agreement patterns. Some adverbs have a slot for gender agreement with the absolutive (88).

- (88) a. han-χ:i-gal mirka<ba.j>da b-aʔanɬa
 village-CFG₇-ABL forever<HPL> HPL-go-INF
Leave the village forever.
- b. armi baq:~e mirka<w>da w-oxa.
 army end-CVB forever<M> M-come-PF
After the army service ended, he came back for good.

4.4.2 Relative clauses

Relative clauses are formed with the participial form of a verb modifying the head of the relative clause. There are three participles in Karata (see section 3.6.6). Depending on the ordering of the time of the events in the relative and main clauses, the past, present, or future participle is used. For instance, in (92), the event of going to bring the ox follows the event of meeting a man, and so the prospective participle is used.

- (92) [unsal b-eł-aχ^wa w-oʔ-ã-λ:o-w] χan-š:u-w waša-š:uw-a dande
 ox N-lead-CVB M-go-PF-FUT.PTCP-M khan-M₀[GEN]-M boy-M₀-DAT together
 w-ok:r-u idja [q:ut'a bičar:ła w-oʔ-ã-λ:o-w] ce-w hek'wa.
 M-happen-PF COP rooster sell-INF M-go-PF-FUT.PTCP-M one-M man

The khan's son who was to lead the ox met a man who was to go sell a rooster.

All sorts of noun phrases can be relativized. In (93), the relative clause modifies the word *zebu* 'day'. The suffixal gender marker on the relative clause verb reflects the gender of the modified NP *zebu* 'day' while the prefixal gender marker agrees with the absolutive argument inside the clause, i.e. *men* 'you'.

- (93) [men w-ox:a-b zebu]-ł:i-l den eši w-uk'-a-č'e.
 2SG M-come-PTCP.N day-N₀-ERG 1SG home[LOC] M-be.PF-NEG

The day you came, I was not at home.

4.4.3 Complement clauses

Karata has one complementizer, the suffix *λ're* – cognate with *keλ':ała* 'say' – which is at the right edge of a sentence (94). This suffix is used to mark reported speech, but it is also used for more general complementation.

- (94) χan-š:u-w waša-š:u-l keλ':z-ẽ idja : [ha-b q':oldeni-q:
 khan-M₀[GEN]-M boy-M₀[GEN]-ERG say-PF COP DEM-N hauberk₀-CFG₅[LOC]
 χisa-bič'-o-le x:wane]-λ're
 exchange-FUT.NEG-PTCP-Q horse-QUOT

*The khan's son said: (you) will you not exchange this hauberk for (your) horse?
 (text 7 Magomedbekova 1971)*

Another complementation strategy used in Karata is to use the masdar form of the main verb of the complement clause. Whether this strategy is possible depends on the selecting verb (95).

- (95) [zini be-r-ja-gal-da] rak'wa-č'o reč':-e den
 cow buy-MSD-CFG₃-ABL-INT heart-CFG₁[LOC] pity-PF 1SG

I regret buying (a/the) cow.

Infinitival clauses are also possible (96).

- (96) [mak'i-lo-χa-r bek'ãła] řadã-j q'or-a idja
 child.PL-H₀⁺-CFG₄-ALL care.INF person-F want-PF[CVB] COP
A woman is needed to look after the children.

Another strategy used where other languages would use a complementizer is to use a relatively general noun modified by a relative clause. For instance in (97), the object of the verb *basãła* ‘say’ is the word *hedela* ‘thing’ modified by a relative clause.

- (97) č'wač'warda:riq: bas-ã [b-oλ:-o-b] hedela
 without_hesitation tell-PF N-happen-PF.PTCP-PF thing
He said what happened without hesitation.

4.4.4 Clause chaining

Karata has clause chains that combine a sequence of non-finite clauses with a final clause. The non-finite clauses are headed by general converbs (section 3.6.6) and the final clause must have a fully inflected independent verb form (section 3.6.3). Clause-chaining constitutes the usual way to encode series of events.

- (98) [ho-b χ:abarij-a-l bež-u<j>eχ^wa] ho-j řadan ho-š:uw-a
 DEM-N news₀-DAT-ADD believe-PF.CVB<F> DEM-F person.N DEM-M₀-DAT after
 χirχir j-ex:^w-a idja.
 F-come-PF COP
Having believed this story, this woman ran after him. (Tetx 4, Magomedbekova)

4.4.5 Adverbial clauses

Karata has a range of so-called specialized converbs (section 3.6.6). Specialized converbs express different relations between the main clause and the subordinate clauses. For instance the specialized converb *-ajgil* can be roughly translated as ‘when’ (99).

- (99) [ha-j in-š:uw-a kařa-r j-ex:^w-ajgil] ha-š:u-l c'aq':a-b adab-q:atir
 DEM-F REFL-M₀-CFG₃[LOC] on-ALL F-come-SPCVB DEM-M₀-ERG very-N politeness
 gah-ała q':as: geh-e.
 do-INF intention do-PF
When she came to him, he tried to be very polite.

The specialized converb *-aλ:a* expresses simultaneity (100).

- (100) ce-b zamana-łi-l hařã [w-ox:-aλ:a] aχran.
 one-N time-NM₀-ERG see-PF M-come-SPCVB keeper
At this moment, we saw the keeper coming.

4.5 Negation

In Karata, clausal negation is expressed via verb inflection (see section 3). In clauses whose nucleus is a negative verb form, the determiners, pronouns, and adverbs expressing free choice in positive clauses function as negative determiners, pronouns, or adverbs.

4.6 Comparative constructions

Adjectives do not have special comparative morphology. Karata forms adjectival comparatives by adding a standard of comparison, marked by the ablative case, to a sentence that is otherwise the same as the corresponding non-comparative construction.

- (101) jac:^w-a-gal herk'ã-w idja waci
sister₀-CFG₃-ABL old-M COP brother
(My) brother is older than (my) sister.

A measure phrase, in the ergative/instrumental case, can optionally be added.

- (102) hač'^wada rešĩ-ł:i-l herk'ã-w idja ho-w dij-a-gal
nine year-N₀-ERG old-M COP DEM-M 1SG₀-CFG₃-ABL
He is nine years older than me.

Comparative constructions expressing equality are formed by suffixing the word *-godo-b* to the standard of comparison.

- (103) waci-l jac:i-godo-j x:o:ho-w idja.
brother-ADD sister-EQUAL-F good-M COP
Brother is as good as sister.

Superlatives are formed with the word *heč'e*. The standard is then either in the genitive case or in the LOC₇ case.

- (104) den hor-do-w heč'e herk'ã-w wa hor-do-ł:i heč'e resmaħal
1SG DEM-H₀⁺[GEN]-M very old-M and DEM-H₀⁺-CFG₇[LOC] very wealth
idja-w, bečeda-w idja.
COP[PTCP]-M wealthy-M COP
I am the oldest of them and I am the wealthiest among them. (text 7, Magomed-bekova 1971)

4.7 Co-ordination

Borrowed particles like *wa* 'and', *am:ɔ* 'but' can coordinate phrases of different sizes, NPs (105a) and full clauses (105b).

- (105) a. ho-b-da zamana-ʎ:i-l b-ik^wa ida ʁalmaʁal t'ukida-k'el **wa**
 DEM-N-INT time-NM₀-ERG N-be-PF COP argument people_of_Tukita-COM and
 halbi-lo-k'el **wa** axi^wada-k'el.
 people_of_Khunzakh-H₀⁺-COM and Akhvakh-COM
*At that same time there was an argument with the people of Tukita, the
 people of Khunzakh and the Akhvaks. (text 2, Magomedbekova 1971)*
- b. ho-b zamana-ʎ:i-l urʁe: ida hordo-baj ʁaq':lu ida-baj
 DEM-N time-NM₀-ERG think[PF] COP DEM-H₀⁺-HPL intelligence COP[PTCP]-HPL
 ʁadan-di **wa** s:oreq: i-dow-a-lo-l beš:ti-da-ʁe b-ak^wa-rik'el
 person-PL and around REFL-H₀⁺-DAT-H₀⁺-ERG leave-IPF-SPCVB HPL-be[PF]-SPCVB
 saro hor-do-l dande geh-e ida inš:uda hane.
 together DEM-H₀⁺-ERG together do[PF]-PF.CVB COP five village
*At that time, intelligent people thought and because those around them did
 not leave them in peace, they united 5 villages. (text 2, Magomedbekova
 1971)*

The clitic *-(e)l* is used to mean *and* on each coordinands (106a) and it is used to translate 'also' (106b).

- (106) a. mesedo-l musa-l ida iš:i-č'o
 Mesedo-ADD Musa-ADD COP 1PL-CFG₁[LOC]
Mesedo and Musa are at our place.
- b. dij-a-r-el bas-ã
 1SG₀-CFG₃-ALL-ADD TELL-IMP
Tell me too!

Finally another conjunction is *jaʁibar* or *ja* 'or'.

- (107) jaʁibar miʁi-ʎ:i-r w-oʔ-ĩ, jaʁibar rox:o-ʎ:i-r w-oʔ-ĩ
 or field₀-CFG₇-ALL M-go-IMP or forest₀-CFG₇-ALL M-go-IMP
Either go to the field or to the forest.

4.8 Non-verbal predication

Almost every clause in Karata must contain an element akin to a verb as diagnosed by the possibility to inflect for tense, aspect, and mood. There is one kind of clause where such element is arguably absent, clauses where the predicative part of the clause is headed by the copula *ida*, usually pronounced with a palatalized consonant *idja*, which cannot be inflected (except for participial and converbial morphology) and can thus only be used in the present tense. In the past and future tenses, the verb *bik^waʁa* 'be' is used.

- (108) a. di-č'o χigi ida wudu-w
 1SG₀-CFG₁[LOC] behind[LOC] COP DEM-M
He is behind me.
- b. c:ibero-q: iši hã-χ:i b-ak'^w-a
 winter₀-CFG₅[LOC] 1pl village₀-CFG₇[LOC] HPL-be-PF
We were in the village in the winter.

The copula does not express gender agreement but it expresses polarity (*idja* is the positive copula, *hač'e* is the negative copula, see section 3.6.3.1).

4.9 Information structure

In Karata, intonation is essential for the expression of information structure. At present, not much is known about information structure in Karata.

5 Lexicon

The lexicon of Karata shows no salient characteristic in comparison with the other Andic languages. In addition to lexemes inherited from Proto-Andic, or whose origin cannot be established, it includes a sizeable proportion of borrowings from Avar, and also from Arabic, Persian, and Turkic languages. Russian is now the main source of borrowings. As mentioned earlier the only dictionary of Karata is Magomedova and Khalidova 2001.

6 Sample text

Text 1 from Magomedbekova 1971, told by Magomed Radžabov in 1954, checked and translated to Russian by R. Khalidova, analyzed and translated to English by J. Pasquereau.

azar-boχ:e hač'^wa-bešan-boχ:e inš:t^wac'adaχ:ob rešenɬ:il hac'al inš:dudaχ:ob
 thousand-PCL nine-hundred-PCL five-ten-ORD-N year-N-ERG ten five-ORD-N
 eli bik:e riχ:i, reɬada saɬat ɬabda q:'ot'i ge: baʔanɬa
 mouth N-hold-PF moment-CFG[LOC] at_night hour three agreement do.PF H-go-INF
 harč':aq':arar x:indaɬerowa χirχir.
 Harchaqara-CFG-ALL fruit-OBL-CFG[LOC] after

In the year 1950, on the fifteenth day of Ramadan, at three o'clock at night, we agreed to go to Harchaqara for fruit.

iš:i ɬabi bak'^wa horge ɬax:al-ɬeboχa.
 1PL.EXCL three H-be[PF] there ??-ess
All of us three went there.

hob reɬaɬ:il iš:i bak'^wa sawetija dežurniɬeboχa.
 DEM-N night-N-ERG we H-be council-CFG[LOC] watchman-ESS
That night we were in the townhall as watchmen.

sawetijagalteboxa iš:i baʔã počijar.
 council-CFG-MED 1PL.EXCL H-go[PF] post-CFG-ALL
We made our way through the townhall to the post office.

horge cebk'eda istakan tamox:ul q'ere, χiri baλ:e harč':aq':arar.
 there one-two glass tobacco cut-PF after H-go-PF harchaqara
There we cut up the equivalent of two glasses of tobacco, then we went to Harchaqara.

hugir baʔã ri-λ:i k^wanog^wal bik'^wač'e.
 there.ALL H-go[PF] moment-CFG[LOC] light-even N-be.PF-NEG
There was not even light when we went there.

rošibar χ:areboxa mic':ab ʕeče gehida bak'^wa.
 tree-PL-CFG[ALL] climb-PF-N-CVB sweet-N apple pick-IPF H-be.PF
We climbed into the tree and picked apples.

k^wano haʔindaβe riλ:i ʕeče gahała zaħmałe.
 light see-IPF-NEG.CVB moment-CFG[LOC] apple pick-INF become_difficult-PF
It became difficult to pick apples at that moment there not being any light.

Acknowledgments:

Abbreviations:

References

- Bokarev, A. A. (1938). *Materialy po izučeniu ando-tsezskix jazykov. Narečie aula Tokita.* [Materials for the study of Ando-Tsezic languages. Dialect of Tukita.].
- Charachidzé, G. (1981). *Grammaire de la langue avar.* Paris: Editions Jean-Favard.
- Creissels, D. (2014). Functive-transformative marking in Akhvakh and other Caucasian languages. *Jazyk, Konstanty, Peremennye*, 430–449.
- Dahl, Ö. (1985). *Tense and aspect systems.* Basil Blackwell.
- Dalgatov, I. K. (2015). Siwuxskij govor v sisteme dialektov avaro-ando-tsezskoj gruppy dagestanskix jazykov. *Istoričeskaja i sotsial'no-obrazovatel'naja mysl'* 7(6-1).
- Dalgatov, I. K. (2017). *Siwuxskij govor karatinskogo jazyka.* Ph. D. thesis, Dagestanskij gosudarstvennyj pedagogičeskij universitet.
- Dirr, A. (1909). *Materialy dlja izuchenija jazykov i narechij ando-didojskoj gruppy* [Materials for the study of the languages and dialects of the Ando-Didoic group]. Tbilissi.

- Erckert, R. V. (1895). *Die Sprachen des Kaukasischen Stammes*, Volume 1. A. Hölder.
- Gadžieva, M. Z. (2007). Fonetičeskie i morfoložičeskie osobennosti tukitinskogo dialektu karatinskogo jazyka. Master's thesis, Dagestanskij Gosudarstvennyj Pedagogičeskij Universitet.
- Khalidova, R. (2017). *Karatinskie skazki [Karata tales]*. Maxačkala: Institut jazyka, literatury i iskusstva imena G. Cadasy Dagestanskij naučnij centr Rossijskoj akademii nauk i Federal'noe gosudarstvennoe bjužetnoe obrazovatel'noe učreždenie vyššego obrazovanija.
- Khalidova, R. S. (2016). Žhivotnovodčeskaja leksika karatinskogo jazyka. *Izvestija Dagestanskogo Gosudarstvennogo Pedagogičeskogo Universiteta. Obščestvennye i gumanitarnye nauki*. 10(3).
- Khalidova, R. S. (2018). Osobennosti funkcionirovanija glagol'nyx form v skazotčnom diskurse karatinskogo jazyka. *Izvestija Dagestanskogo gosudarstvennogo pedagogičeskogo universiteta. Obščestvennye i gumanitarnye nauki*. 12(2).
- Kibrik, A. E. e. (2001). *Bagvalinskij jazyk: grammatika, teksty, slovani* (*The Bagvalal language: grammar, texts, dictionaries*). Moscow: Nasl'ed'e.
- Kolga, M. (2013). *The red book of the peoples of the Russian Empire*. eBookIt.com.
- Kuraeva, M. N. (1998). *Somatičeskaja frazeologija karatinskogo jazyka*. Ph. D. thesis.
- Kuraeva, P. G. (2010). *Leksičeskie i morfoložičeskie sredstva vyraženiija kategorii lokalizacii v avaro-andijskix jazykax*. Ph. D. thesis, Institut jazyka, literatury i iskusstva imeni G. Cadasy Dagestanskogo naučnogo centra Rossijskoj akademii nauk.
- Magomedbekova, P. (1971). *Karatinskij jazyk: grammatičeskij očerok, teksty, slovar'* (*Karata: grammar sketch, texts, dictionary*). Mecniereba: Tbilisi.
- Magomedbekova, Z. M. (1961). *O fonetičeskix osobennostjax tokitinskogo dialektu karatinskogo jazyka [On phonetic peculiarities of the dialect of Tukita of the Karata language]*. Tbilisi.
- Magomedbekova, Z. M. (1967). *Jazyki narodov SSSR [The languages of the peoples of the USSR]*, Chapter Karatinskij jazyk [The Karata language].
- Magomedbekova, Z. M. (2000). Karatinskij jazyk [The Karata language]. In I. K. Abdullaev, U. A. Mejlanova, A. G. Magomedov, M. C. M. Musaev, P. A. Saidova, B. B. Talibov, M. S. Xalilov, and B. G. Khanmagomedov (Eds.), *Jazyki Dagestana [The languages of Daghestan]*, Jazyki narodov Rosii [The languages of the peoples of Russia], pp. 408 – 421. Russian Academy of Sciences.
- Magomedova, P. T. and R. S. Khalidova (2001). *Karatinsko-ruskij slovar'*. (*Karata-Russian dictionary*). Maxačkala: Dagestanskij Naučnyj Centr Rossiskoj Akademii Nauk.
- Mallaeva, Z. M. and R. v. Xalidova (2016). *Sistemy glagola v avarskom i karatinskom jazykax*. LAP LAMBERT Academic Publishing.

- Pasquereau, J. (2009). *Eléments de sociolinguistique du karata*. Unpublished manuscript.
- Pasquereau, J. (2010). *Spatial cases in Karata: spatial and non-spatial uses*. Master's thesis, Université Lyon II.
- Pasquereau, J. (2011). *Valency in Karata: a preliminary study*. Master's thesis, Université Lyon II.
- Pasquereau, J. (2018). Phonological degrees of labiality. *Language* 4(94), e216–e265.
- Pasquereau, J. and R. Khalidova (2017). A preliminary investigation of the formation of constituent questions in Karata (Nakh-Daghestanian). *Proceedings of the Linguistic Society of America* 2, 50–1.
- Paxrudinova, R. O. (2004). *Slovoobrazovanie v karatinskom jazyke*. Master's thesis.
- Paxrudinova, R. O. and Z. M. Akhmedova (2014). Suffiksals'noe obrazovanie imen suščestvitel'nyx v karatinskom i axvaxskom jazykax. *Izvestija Dagestanskogo Gosudarstvennogo Pedagogičeskogo Universiteta. Obščestvennye i gumanitarnye nauki* (1 (26)).
- Plungian, V. A. (1997). Plurality of situations in Chamalal. In V. S. Xrakovskij (Ed.), *Typology of iterative constructions*, pp. 1. München: LINCUM Europa.
- Talmy, L. (1972). *Semantic Structure in English and Atsugewi*. Ph. D. thesis, University of California, Berkeley.
- Umargadžieva, A. G. (2010). *Fitonimičeskaja i zoonimičeskaja leksika v andijskix jazykax: na materiale karatinskogo, axvaxskogo i tindinskogo jazykov*. Master's thesis, Dagestanskij Gosudarstvennyj Pedagogičeskij Universitet.
- Xalidova, R. v. (1990). *Lingvodidaktičeskie osnovy obučenija russkomu literaturnomu proiznošeniju učaščixsja-nositelej bespis'mennyx jazykov (karatinskogo i ahvahskogo)*. Ph. D. thesis.