



Universiteit
Leiden
The Netherlands

A grammar of Kumzari : a mixed Perso-Arabian language of Oman

Wal Anonby, C.A. van der

Citation

Wal Anonby, C. A. van der. (2015, April 22). *A grammar of Kumzari : a mixed Perso-Arabian language of Oman*. Retrieved from <https://hdl.handle.net/1887/32793>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/32793>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/32793> holds various files of this Leiden University dissertation.

Author: Wal Anonby, Christina van der

Title: A grammar of Kumzari : a mixed Perso-Arabian language of Oman

Issue Date: 2015-04-22

2 Phonology

Kumzari has eight vowels and twenty-eight consonants. The tables of vowels (Table 2) and consonants (Table 3) below give the phonemic inventory of Kumzari. Table 10 at the end of this chapter shows IPA¹⁶ correlates of the forms: the conventions for vowels and consonants written in phonemic script used in this book, and in the Kumzari writing system¹⁷.

2.1 Vowels

Table 2. Vowel phonemes in Kumzari

	front	central	back
high	ī		ū
	i		u
mid	ē		ō
low		a	
		ā	

Kumzari has eight vowel phonemes: five long vowels /ī/ /ū/ /ē/ /ō/ /ā/ and three short vowels /i/ /u/ /a/. Phonetically, the three short vowels have a mid-centralised quality (and in the case of /a/ [ɐ], toward close) as compared to their long counterparts. In fact, the short *i* and short *u* alternate with both high and mid long vowels in some environments. There are no vowel sequences in Kumzari. Two vowels are always separated by at minimum a glide or glottal stop, even when cliticised. Examples given below phonemically begin with a vowel, but phonetically all vowel-initial words in fact begin with a glottal stop, e.g. [ʔi:ʃu:] ‘living’, [ʔu:ʃu:] ‘kindling’, [ʔe:lɪŋg] ‘bracelet’, etc.

The phoneme /ī/ is a long close front unrounded vowel [i:], as in the word *īšū* ‘living’.

The phoneme /ū/ is a long close back rounded vowel [u:], found in the word *ūrū* ‘kindling’.

The phoneme /ē/ is a long mid front unrounded vowel [e:], represented in the word *ēling* ‘bracelet’. Its allophone is [e] as in *sāhenē* ‘powdered sardines’.

The phoneme /ō/ is a long mid back rounded vowel [o:], as in the word *ōl* ‘mountain peak’.

The phoneme /ā/ is a long open central unrounded vowel [a:], shown by the word *ād* ‘silent’.

The phoneme /i/ is a short near-close centralised front unrounded vowel [ɪ], that is in the word *illit* ‘dirty’. It has an allophone [ʊ] as in [gɪsm]/[gʊsm] ‘I have taken’.

The phoneme /u/ is a short near-close centralised back rounded vowel [ʊ], represented by the word *uff* ‘blowing’.

The phoneme /a/ is a short near-open central unrounded vowel [ɐ], found in the word *atta* ‘wet’. It is nasalised in one word: *ā’ā* ‘no’ and in morphemes shortened from a nasal consonant, e.g. *pāšumbur* ‘Thursday’ from *panj* + *šumbur* ‘five’ + ‘day of week’.

Phonemic contrast between closely related vowels is demonstrated in the following pairs:

¹⁶ International Phonetic Alphabet

¹⁷ The Kumzari writing system was developed by phonologist Erik Anonby in 2009 at the behest of Kumzari community representatives and leaders, and the resulting alphabet chart given in a report to Sultan Qaboos University, Muscat.

/i/ and /ī/	<i>išōwō</i> ‘tonight’ <i>īšū</i> ‘living’
/u/ and /ū/	<i>umr</i> ‘age, soul’ <i>ūmat</i> ‘sardine’
/a/ and /ā/	<i>aft</i> ‘seven’ <i>āf</i> ‘short pants’
/ū/ and /ō/	<i>ūd</i> ‘oud’ <i>ōd</i> ‘knock’
/ī/ and /ē/	<i>īran</i> ‘Iran’ <i>ērarağ</i> ‘Bream (fish species)’

2.2 Consonants

Twenty-eight consonants are distinguished in Kumzari. They are listed in Table 3; their phonetic values are described in §2.2.1 and following sections below.

Table 3. Consonant phonemes in Kumzari

	labial (-dental)	alveolar	velarised alveolar	(alveo-) palatal	velar	uvular	pharyngeal	glottal
stop/affricate	p b	t d	ṭ ḍ	č j	k g	q		ʔ
fricative	f	s	š ẓ	š		x ġ	ħ	h
nasal	m	n						
approximant	w	l r	ɭ	y				

The Kumzari consonant inventory is drawn from both of its heritage languages,¹⁸ including the Persian sounds /p/, /č/, and /g/¹⁹, as well as a set of emphatics reminiscent of Semitic: velarised alveolars /ṭ/, /ḍ/, /š/, /ẓ/, and /ɭ/, uvular /q/, and pharyngeal /ħ/. Non-emphatic counterparts are attested for all of the alveolars except /z/, which is always velarised, including in words of Indo-European²⁰ origin, e.g. *rōz* ‘day’, *ẓwān* ‘tongue’, *rēz* ‘pour’. Emphatics occur even in words derived from non-Semitic sources, e.g. *širx* ‘red’, *ṭāzağ* ‘freshness’, *pānda* ‘fifteen’, *bōš* ‘kiss’, *ṭahl* ‘bitter’, *čāz* ‘lunch’, *tēz* ‘sharp’.

Like Bakhtiari, Baluchi, and Mamasani Luri²¹, Kumzari has a bilabial velar approximant /w/, rather than /v/ as in New Persian spoken in Iran, e.g. *wēkil* ‘guardian’ cf. Persian *vakīl*; *wēzīr* ‘government minister’ cf. Persian *vazīr*. The same phoneme can correspond to /b/ of Persian²², as in *swak* ‘lightweight’ cf. Persian *sabok*; *wustin* ‘pregnant’ cf. Persian *ābestān*; *šaw* ‘night’ cf. Persian *šab*. Like Shihhi Arabic (Bernabela 2011:26) and some South

¹⁸ The term ‘heritage language’ refers to Kumzari’s ancestral languages: Middle Persian and a Semitic language (see Classification §1.2.1).

¹⁹ These are also found in Gulf Arabic (Holes 1984:8-9, Holes 1990:260ff).

²⁰ The term ‘Indo-European’ is used throughout this book to refer to origins that are thought to be Persian but not necessarily a specific variety of that family of languages; see further explanation of the term in the list of abbreviations.

²¹ Specifically, Bakhtiari and Mamasani Luri also have labio-dental allophones of /w/ (cf. Jahani & Korn 2009:645-646; Lorimer 1922:16-17; MacKinnon 2011).

²² Skjærvø 2009:201

Arabian languages (Simeone-Senelle 1997:381-382²³), Kumzari does not have certain sounds as are found in Classical or north-central peninsular Arabic varieties: voiced pharyngeal fricative [ʕ] and interdental fricatives [θ], [ð], and [ð̪].

Uniquely for the wider region, but like Shihhi Arabic (Bernabela 2011:23-25) and the South Arabian languages Mehri and Hobyot (Simeone-Senelle 1997:383), the Kumzari /r/ is a retroflex approximant [ɻ]. In certain environments outlined below, /r/ is realised as an alveolar or retroflex flap or as a trill.

Kumzari's retention of phonological elements from both ancestor language families is further evidence of its identity as an entirely mixed language. In languages with high proportions of borrowings, loanwords are adapted to the existing language's phonological system (Bakker 1997:10). Kumzari phonology, in contrast, is not wholly copied from either of its predecessors.

All consonants are found in word-initial position, where they are contrastive. Listed below are consonants in morpheme-initial onsets.

p	<i>pā</i>	'foot, leg'
t	<i>tā</i>	'one'
ʈ	<i>ṭāma</i>	'commission'
k	<i>kāra</i>	'mouth'
q	<i>qāba</i>	'shirt'
'	<i>'ā</i>	3S.ANA (anaphoric pronoun)
b	<i>bā</i>	'armspan'
d	<i>dāmar</i>	'groom'
ḍ	<i>ḍālum</i>	'tyrant'
g	<i>gā</i>	'bull'
č	<i>čādir</i>	'bedsheets'
j	<i>jā</i>	'barley'
f	<i>fālaj</i>	'water channel'
s	<i>sā</i>	'now'
ʃ	<i>šābun</i>	'soap'
š	<i>šā</i>	3PL.EMPH
x	<i>xāna</i>	'marriage'
z	<i>zā</i>	'give birth'
ǧ	<i>ǧāna</i>	'jaw'
ḥ	<i>ḥāra</i>	'quarter (of a city)'
h	<i>hā</i>	'yes?'
m	<i>mā</i>	'month'
n	<i>nābī</i>	'gull sp.'
l	<i>lābit</i>	'certainly'
ʎ	<i>aʎla</i>	'God'
r	<i>rā</i>	'way'
w	<i>wā</i>	'woe'
y	<i>yā</i>	'this'

²³ Soqotri does not have interdentals; in Harsusi and Dhofar Mehri the voiced pharyngeal fricative is marginal.

2.2.1 Stops and Affricates

The phoneme /p/ is a voiceless bilabial stop [p] represented by the word *pāk* ‘clean’.

The phoneme /b/ is a voiced bilabial stop [b] as in the word *bukr* ‘firstborn’.

The phoneme /t/ is a voiceless alveolar stop [t] seen in the word *tak* ‘date syrup basket’.

The phoneme /d/ is a voiced alveolar stop [d] exemplified in the word *dōl* ‘mast’.

The phoneme /k/ is a voiceless velar stop [k] seen in the word *kaf* ‘sole (of foot), palm (of hand)’.

The phoneme /g/ is a voiced velar stop [g] represented by the word *gōz* ‘walnut’.

The phoneme /q/ is a voiceless uvular stop [q] seen in the word *qōq* ‘tantrum’. The segment *q* was rare or non-existent in Middle Persian (Skjærvø 2009:200), but some modern Iranian languages and dialects retain it (Windfuhr & Perry 2009:423; McCarus 2009:592; Paul 2009:547). While Kumzari has kept all instances of *q* from Old Arabic or Semitic, in adjacent Shihhi Arabic, *q* is often spirantised intervocalically to a fricative (Bernabela 2011:24), as it is in standard New Persian.

The phoneme /t̪/ is a voiceless velarised alveolar stop [t̪] represented by the word *tāf* ‘twenty-four-hour gale’. It contrasts with the voiceless non-velarised alveolar stop /t/ [t]: *tēr* ‘bird’, *tēra* ‘way’.

The phoneme /d̪/ is a voiced velarised alveolar stop [d̪]; it is represented by the word *ḍaby* ‘oryx’. It contrasts with the voiced non-velarised alveolar stop /d/ [d]: *ḍalama* ‘being maligned’, *darama* ‘curing with medicine’.

The phoneme /t͡ʃ/ is a voiceless palato-alveolar affricate [t͡ʃ] as found in the words *čāf* ‘shore’ and *čap* ‘paddle’. Among younger speakers and those under more influence from Arabic, this phoneme is sometimes realised as /š/ [ʃ], e.g. *čihhī* / *šihhī* ‘Shihhi [Arab ethnic group]’, *čupš* / *šupš* ‘lobster’, *pāčar* / *pāšar* ‘raised half-deck (boat)’.

The phoneme /d͡ʒ/ is a voiced alveopalatal affricate [d͡ʒ] as in the word *jāmağ* ‘man’s undershirt’.

Several stops and fricatives have irregular dialectal alternations: [b] ~ [g]: [ʔʷboɪd̪ə] ~ [ʔʷgoɪd̪ə] ‘sheer strake’, [p] ~ [b] ~ [f]: [ʔʷe:pɪl̪əʁ] ~ [ʔʷe:bɪl̪əʁ] ~ [ʔʷe:fil̪əʁ] ‘wooden chest’ and [ʔ] ~ [h]: [hɪɟk] ~ [ʔɪɟk] ‘dry’.

2.2.2 Fricatives

The phoneme /f/ is a voiceless labiodental fricative [f] as found in the word *fīmē* ‘black dolphin’.

The phoneme /s/ is a voiceless alveolar fricative [s] represented by the words *salq* ‘large boat sp.’ and *sist* ‘loose’.

The phoneme /š/ is a voiceless alveopalatal fricative [ʃ] represented by the word *šubr* ‘handspan’. It is in alternation with the affricate /t͡ʃ/ [t͡ʃ]: *šarrax* / *čarrax* ‘straddled’.

The phoneme /x/ is a voiceless uvular fricative [χ] seen in the word *xall* ‘seaweed’.

The phoneme /ğ/ is a voiced uvular fricative [ʁ] seen in the word *ğas* ‘post’.

The phoneme /s̪/ is a voiceless velarised alveolar fricative [s̪] represented by the word *šawz* ‘green’. It contrasts with the voiceless non-velarised alveolar fricative /s/ [s]: *šām* ‘handle’, *sāma* ‘heaven’.

The phoneme /z̪/ is a voiced velarised alveolar fricative [z̪] represented by the word *zīn* ‘thief’. It has no non-velarised counterpart in Kumzari.

2.2.3 Nasals

The phoneme /m/ is a voiced bilabial nasal [m] as in the word *muxx* ‘head’.

The phoneme /n/ is a voiced alveolar nasal [n] represented by the words *nām* ‘name’ and *nēt* ‘charity food’. There is insufficient evidence to regard the velar nasal [ŋ] as a phoneme

distinct from /n/, since it occurs only before voiced and voiceless velar stops /g/ [g]: *dang* [dæŋg] ‘cyst’ and /k/ [k]: *linkit* [lɪŋkɪt] ‘finger’.

2.2.4 Laterals

The phoneme /l/ is a voiced alveolar lateral approximant [l] represented by the words *lupp* ‘marrow’ and *langal* ‘anchor’.

The marginal phoneme /l̥/ is a voiced velarised lateral approximant [l̥] represented by the word *afalla* ‘blessing’. It occurs infrequently, and exclusively in words of Semitic origin. It contrasts with the voiced non-velarised lateral approximant /l/ [l]: *waʔa* ‘or’, *walama* ‘readying’.

2.2.5 Approximants

The phoneme /w/ is a voiced labial-velar approximant [w] shown by the word *waqt* ‘time’.

The phoneme /r/ is a retroflex alveolar approximant [ɻ] represented by the word *raff* ‘niche’. It has the following allophones:

Following a long vowel, the alveolar approximant may be pronounced as not retroflexed [ɻ]:

kōr [ko:ɻ] ‘whale’

brār [bɪa:ɻ] ‘brother’

sūr [su:ɻ] ‘wedding’

It may be pronounced as a flap rather than as an approximant; that is, as a retroflex alveolar flap [ɽ] after a stop in an onset cluster:

krāh [kɽa:h] ‘sandal’

brišt [bɽɪʃt] ‘cooked’

drāz [dɽa:zʷ] ‘length’

and as a non-retroflex alveolar flap [ɾ] after a short vowel:

kara [ˈkɛɾɐ] ‘time’

nēgura [ˈne:ɡɔɾɐ] ‘two days ago’

As a geminate, it is a non-retroflex alveolar trill [r]:

mirr [mɪrr] ‘myrrh’

barra [ˈbɛrrɐ] ‘outside’

qarraʃ [ˈqɛrrɐsʷ] ‘mosquito’

The phoneme /y/ is a voiced palatal approximant [j] seen in the word *yirz* ‘long-handled axe’.

2.2.6 Pharyngeal

The phoneme /ħ/ is a voiceless pharyngeal fricative [ħ] as in the word *ħabb* ‘seed’. It contrasts with the voiceless glottal fricative /h/: *ħazza* ‘wrapping’, *ħazza* ‘angry’, and is more common than *h* across the lexicon. Kumzari words of Semitic origin retain pharyngealisation of *ħ*, e.g. *ħuwwil* ‘spring migration’, *ħaqq* ‘justice’, *ħazana* ‘sad’.

2.2.7 Glottals

The phoneme /ʔ/ is a glottal stop [ʔ]; an example is in the word *ʔaʔa* ‘no’. The glottal stop can occur in any position in a word: *ʔōzar* ‘sail’, *bʔām* ‘thumb’, *ābā* ‘trap fish’, *daʔbar* ‘chatter’, *āʔilit* ‘family’, *paʔn* ‘wide’. The minimal pair *qarra* ‘admit’, *qarʔa* ‘gourd’ sets it out as a separate phoneme. The glottal stop is inserted epenthetically between two vowels at

a morpheme boundary, e.g. *hajabē'ō* ‘tholepin [boat]’, and as a word-initial consonant preceding a vowel, e.g. *'ēbē* ‘woman’s black cloak’. In words borrowed from Semitic, the glottal stop [ʔ] is generally retained, e.g. *s'āl* ‘question’. It is also the reflex in Kumzari for Semitic-origin words with the voiced pharyngeal fricative [ʕ], as in the words *arafa* ‘knowing’ [ʕərəf] > [ʔərəfə], *wa'rit* [wəʔɪt] ‘stiflingly hot place’ cf. Old Arabic [wəʕr], *zara'a* ‘cultivating’ [zərəʕə] > [zʔərəʔə]. Among some Kumzari speakers, the glottal stop [ʔ] is the reflex for the glottal fricative [ħ] in words of Indo-European origin, e.g. *ba'tar* ~ *bātar* < *behtar* ‘better’. In the Laraki dialect, the glottal stop has shifted to the glottal fricative /h/ [ħ] in all positions, e.g. Lrk. *kahnağ* Kmz. *ka'nağ* ‘old, worn out’.

The phoneme /h/ is a voiceless glottal fricative [ħ] seen in the word *hē* ‘yes’. This phoneme is rare; usually in words of both Semitic and Indo-European origin the sound has been diachronically replaced with a voiceless glottal stop [ʔ], e.g. Kmz. *la'aba* ‘roaring (fire)’ cf. Arabic *lihīb*, Kmz. *tā'ir* ‘purify’ cf. Arabic *taher*. In at least one word it became pharyngealised [ħ]: *qāħwē* ‘coffee’. Where it occurs, which is for the most part in ideophones and synchronic borrowings, the voiceless glottal fricative has a breathy quality. The phoneme /h/ is more widely established in the Laraki dialect, e.g. Lrk. *hātiš* Kmz. *ātiš* ‘fire’, where it may also occur instead of a glottal stop as a word-initial consonant preceding a vowel, e.g. Lrk. *hāmad* [ħa:məd] ‘come:3sREAL’ (cf. Kmz. [ʔa:məd]).

2.3 Consonant distribution

Most consonants are attested in all positions in a word. Exceptions in each category tend to be peripheral phonemes, glottals, and semivowels. In the following sections consonants are laid out as they appear in word-final position, intervocalically, then in clusters (initial and second consonant in an inset cluster, and initial and second consonant in a coda cluster), in word-internal sequences, and as geminates.

2.3.1 Consonants in word-final position

All consonants except the peripheral phonemes /h/ and /l/ are attested in word-final position in the data.

p	<i>čap</i> ‘paddle’
b	<i>kasib</i> ‘livelihood’
f	<i>tāf</i> ‘head rope of fishing net’
t	<i>mast</i> ‘tired’
ʈ	<i>maqfaʈ</i> ‘trap’
d	<i>qad</i> ‘dimension’
ɖ	<i>ħaɖɖ</i> ‘lucky’
č	<i>gačč</i> ‘mortar’
j	<i>tāj</i> ‘crown’
s	<i>xus</i> ‘dear’
ʃ	<i>šaxʃ</i> ‘person’
z	<i>tēz</i> ‘sharp’
š	<i>gōš</i> ‘ear’
k	<i>tak</i> ‘date syrup basket’
x	<i>lēx</i> ‘fishing net’
g	<i>mayg</i> ‘shrimp’
ğ	<i>muğ</i> ‘date palm’
ʾ	<i>naʾ</i> ‘nine’

h (not attested in this position)

ḥ *lōḥ* ‘wood’

q *rīq* ‘throat’

m *gīm* ‘bait’

n *qarn* ‘horn’

l *gil* ‘dirt’

l̥ (not attested in this position)

r *šīr* ‘milk’

w *taw* ‘illness’

y *say* ‘traditional fish net’

2.3.2 Intervocalic consonants

All consonants occur intervocalically:

p *zēpilaḡ* ‘wooden chest’

b *adaba* ‘suffering’

f *adafa* ‘injuring’

t *sakata* ‘silent’

ṭ *naḡaṭa* ‘shaking’

d *qādaḥ* ‘container’

ḍ *aḍalat* ‘muscle’

č *kāčak* ‘underarm’

j *šaja* ‘a’ ‘support’

s *šakasa* ‘cutting’

ʃ *naʃara* ‘winning’

z *ṭāzaḡ* ‘fresh’

š *drīša* ‘window’

k *adaka* ‘worn out’

x *kašaxa* ‘super’

g *wālagō* ‘deck hatch (boat)’

ḡ *sabaḡa* ‘decorating’

’ *qaša* ‘a’ ‘tipping over’

h *fahama* ‘understanding’

ḥ *saḡara* ‘enchantment’

q *alaqa* ‘hanging’

m *ramaqa* ‘winking’

n *fanana* ‘super’

l *qalaba* ‘turning over’

l̥ *waḷa* ‘or’

r *amara* ‘sealing’

w *adawa* ‘avoidance’

y *rayaḡa* ‘resting’

2.3.3 Consonant clusters and sequences

One factor contributing to Kumzari’s unique phonological structure is its lack of restrictions on sonority in consonant clusters. Many of the examples in the lists below violate the sonority sequencing principle. This distinguishes Kumzari from Persian, Bakhtiari, and standard Arabic, which do not allow initial consonant clusters. Even adjacent Shihhi Arabic limits initial consonant clusters to a few with increasing sonority (i.e. the second consonant is a semivowel, liquid, or nasal) (Bernabela 2011:40-41). One of Kumzari’s presumed ancestor languages, Middle Persian, did allow initial consonant clusters.

The following sections list examples for each phoneme as initial and second consonants in onset clusters and initial and second consonants in coda clusters.

2.3.3.1 Initial consonants in onset clusters

All consonants other than /h/, /l/, /w/, and /y/ occur in the initial position of an onset cluster. A possible reason that the exceptions are not attested in this position in the data is that these are peripheral consonants and glottals or glides. As a first component the consonant phonemes form the following onset clusters:

p	<i>pxūn</i> ‘approaching rain’
b	<i>blind</i> ‘high’
f	<i>flīta</i> ‘wick’
t	<i>twām</i> ‘pair of twins’
ṭ	<i>ṭyāh</i> ‘sardine season’
d	<i>dnān</i> ‘tooth’
ḍ	<i>ḍlaḡ</i> ‘sock’
č	<i>črā</i> ‘lamp’
j	<i>jwān</i> ‘good’
s	<i>sbō</i> ‘week’
ʃ	<i>ʃnāfē</i> ‘Streaked Rabbitfish’
z	<i>zlaḡ</i> ‘sock’
š	<i>šdūd</i> ‘ongoing rain showers’
k	<i>klīl</i> ‘key’
x	<i>xlēsī</i> ‘tuna sp.’
g	<i>gdar</i> ‘wall’
ḡ	<i>ḡbēb</i> ‘light cloud’
ʾ	ʾmēd (a fish species)
ḥ	<i>ḥkum</i> ‘judgment’
h	(not attested in this position)
q	<i>qraḥ</i> ‘bald’
m	<i>mrād</i> ‘reason’
n	<i>ntōr</i> ‘treats’
l	<i>lbān</i> ‘frankincense’
l̥	(not attested in this position)
r	<i>rkāḥ</i> ‘sandal’
w	(not attested in this position)
y	(not attested in this position)

2.3.3.2 Second consonants in onset clusters

All consonants except the marginal phoneme /l̥/ can be found as the second unit in an onset cluster:

p	<i>spēr</i> ‘white’
b	<i>qbēl</i> ‘sardine net’
f	<i>xfēf</i> ‘cheese sp.’
t	<i>ftāq</i> ‘hernia’
ṭ	<i>mṭār</i> ‘label’
d	<i>mdallē</i> ‘coffeepot’
ḍ	<i>xḍārī</i> ‘taupe’
č	<i>pčāl</i> ‘soiled’
j	<i>mjūm</i> ‘metal stud’

s	<i>ḥsēb</i> ‘calculation’
ṣ	<i>rṣāṣ</i> ‘bullet’
ẓ	<i>rẓāq</i> ‘provisions’
š	<i>ršēš</i> ‘neem tree’
k	<i>mkabb</i> ‘cone’
x	<i>sxaḥya</i> ‘concealing, concealment’
g	<i>ngāl</i> ‘headdress rings’
ğ	<i>šğā</i> ‘robust’
’	<i>m’āš</i> ‘salary’
h	<i>dhōr</i> ‘decade’
ḥ	<i>ṣḥabba</i> ‘loving’
q	<i>mqābalit</i> ‘meeting’
m	<i>qmāṭ</i> ‘swaddling’
n	<i>ẓnān</i> ‘dish’
l	<i>klīl</i> ‘key’
l̥ (not attested in this position)	
r	<i>mrū</i> ‘chicken’
w	<i>nwāxa</i> ‘ship captain’, <i>ḥwēl</i> ‘chatter’
y	<i>xyār</i> ‘cucumber, zucchini’

2.3.3.3 Initial consonants in coda clusters

As the initial consonant of a coda cluster, only /č/²⁴ is not attested in the data:

p	<i>kapš</i> ‘sheep’
b	<i>sabğ</i> ‘dye’
f	<i>ṣufr</i> ‘copper’
t	<i>fitr</i> ‘thumb–index finger span’
ṭ	<i>iṭr</i> ‘perfume’
d	<i>qadr</i> ‘amount’
ḍ	<i>ḥaḍḍ</i> ‘luck’
č (not attested in this position)	
j	<i>ijr</i> ‘good deeds’
s	<i>jisr</i> ‘bridge’
ṣ	<i>uṣb</i> ‘stomachache’
ẓ	<i>ğuzr</i> ‘depth’
š	<i>gišr</i> ‘coral’
k	<i>rikd</i> ‘corner’
x	<i>nixn</i> ‘finger/toe nail’
g	<i>wagz</i> ‘tip of palm frond midrib’
ğ	<i>šiğl</i> ‘things’
’	<i>ra’d</i> ‘thunder’
h	<i>fahd</i> ‘cheetah’
ḥ	<i>ṭaḥl</i> ‘bitterness’
q	<i>ṣuqr</i> ‘osprey’
m	<i>jumr</i> ‘ember’
n	<i>ūling</i> ‘traditional small waterpipe’
l	<i>xilq</i> ‘crowd’
l̥	<i>jilṭ-milṭ</i> ‘stark naked’ (cf. <i>jaḷaṭa</i> ‘skinned, skinning [an animal]’)

²⁴ It is probably not attested as an initial consonant in a cluster because it is already phonologically composite [tʃ].

r	<i>urq</i> ‘root’
w	<i>kawl</i> ‘wind’
y	<i>rayb</i> ‘yoghourt’

2.3.3.4 Second consonants in coda clusters

In the second position of a coda cluster, all consonants except /h/ and /w/ occur in the data.

p	<i>asp</i> ‘horse’
b	<i>jēlumb</i> ‘side, moiety’
f	<i>saqf</i> ‘ceiling’
t	<i>waqt</i> ‘time’
ʈ	<i>faxʈ</i> ‘thigh’
d	<i>sard</i> ‘cold’
ɖ	<i>arɖ</i> ‘land’
č	<i>čarč</i> ‘charged (battery)’ (English loan presumed to be via Gulf Arabic)
j	<i>trinj</i> ‘citron’
s	<i>kups</i> ‘mound’
ʃ	<i>čupʃ</i> ‘lobster’
z	<i>luğz</i> ‘riddle’
š	<i>kapš</i> ‘sheep’
k	<i>mišk</i> ‘mouse’
x	<i>muxx</i> ‘head’
g	<i>xāyg</i> ‘egg’
ğ	<i>rişğ</i> ‘wrist’
ʾ	<i>qišʾ</i> ‘deep-water fish habitation’
h (not attested in this position)	
ħ	<i>ʃulħ</i> ‘peace’
q	<i>xabq</i> ‘small hole’
m	<i>lahm</i> ‘body’
n	<i>sinn</i> ‘net anchor’
l	<i>faql</i> ‘Porcupine fish’
ɭ	<i>aʃɭ</i> ‘origin’
r	<i>giʒr</i> ‘carrot’
w	<i>mʃaww</i> ‘barnacle’
y	<i>sōnty</i> ‘raft’

2.3.4 Word-internal consonant sequences

There are no apparent restrictions on word-internal consonant sequences; that is, contiguous consonants across a syllable boundary. Even combinations of consonants which are not attested in onset or coda clusters are possible in a sequence. Below is a selection of potential sequences.

<i>šawħaʈ</i>	‘whale’
<i>bahlul</i>	‘Potato Grouper (fish species), small stage’
<i>ħaylō</i>	‘swing’
<i>qarʾa</i>	‘squash’
<i>madwax</i>	‘pipe (smoking)’
<i>kaʾnağ</i>	‘old, worn out’
<i>ʃaħra</i>	‘desert’
<i>ğēlbū</i>	‘winning’
<i>bağğa</i>	‘self-congratulations’
<i>axča</i>	‘gold’

qāḥwē ‘coffee’
kan’ad ‘fish sp.’

2.3.5 Geminates

Most consonants may form geminates. When a suffix is added, the second consonant in the geminate joins the following syllable. Not attested as geminates in the data are /ʔ/ /ǧ/ /h/, likely due to the phonetic difficulty of doubling final glottals.

p	<i>lupp</i> ‘marrow’
b	<i>dubb</i> ‘bear’
f	<i>laff</i> ‘bandage’
t	<i>jitt</i> ‘corpse’
ṭ	<i>buṭṭ</i> ‘duck’
d	<i>sadd</i> ‘dam’
ḍ	<i>ḥaḍḍ</i> ‘luck’
č	<i>gačč</i> ‘mortar’
j	<i>ḥijj</i> ‘the Hajj pilgrimage’
s	<i>ḥiss</i> ‘voice’
ṣ	<i>ǧaṣṣ</i> ‘peg, post’
z	<i>mzizz</i> ‘very salty food’
š	<i>gišš</i> ‘shard’
k	<i>čikk</i> ‘little’
x	<i>dixx</i> ‘smoke’
g	<i>bugg</i> (proper name of a wadi in Kumzar)
ǧ	(not attested in this position)
ʔ	(not attested in this position)
h	(not attested in this position)
ḥ	<i>gaḥḥ</i> ‘watermelon’
q	<i>ruqq</i> ‘shallow water’
m	<i>gumm</i> ‘fist’
n	<i>sinn</i> ‘net anchor’
l	<i>nall</i> ‘faucet’
l̥	<i>aḷḷa</i> ‘God’
r	<i>girr</i> ‘heedlessness’
w	<i>daww</i> ‘yoghourt drink’
y	<i>sayy</i> ‘lift up’

2.4 Syllable structure

The types of syllable structure in monomorphemic Kumzari words are these:

CV

CVV²⁵

CVC

CVVC

CVCC

CVVCC

CCVV

CCVVC

²⁵ VV represents a long vowel.

CCVCC
CCVVCC

2.4.1 Syllable shapes in monosyllabic words

The minimal syllable shape in monosyllabic words consists of a consonant and a short vowel (CV):

na ‘not’
ka ‘right away’
pi ‘from’
wa ‘and’

or a short vowel between two consonants (CVC):

dil ‘heart’
kaf ‘palm, sole’
gil ‘dirt’

or a long vowel with a consonant before (CVV):

bā ‘armspan, curse’
čō ‘well’
jā ‘barley’

or a glottal stop or other consonant plus a long vowel with a consonant after (CVVC):

*ōl*²⁶ ‘mountain peak’
āw ‘water’
ēl ‘cardamom’
lēx ‘fishing net’
dūš ‘date syrup’
xōx ‘peach’

or a consonant plus a vowel followed by two consonants (CVCC):

maql ‘myrrh’
binj ‘thigh’
tafš ‘vile’
bukr ‘firstborn’

or a consonant, long vowel, and two consonants (CVVCC):

qāpt ‘white fish sp.’
ḥāwš ‘destruction’

Longer syllables of a single morpheme are:

²⁶ Word-initial glottal stops are not written in Kumzari.

CCVV

xwā ‘salt’
drō ‘lie (untruth)’

CCVC

stağ ‘date pit’
qraḥ ‘bald person’

CCVVC

ṭyāḥ ‘sardine season’
ršēš ‘neem tree’
qbīb ‘narrow’
pxūn ‘approaching rain’

CCVCC

mṣaww ‘barnacle’
brinz ‘rice’

CCVVCC

stārg ‘star’

Monosyllabic words with short vowels must be closed syllables; monosyllabic words with long vowels can be open or closed syllables:

<i>sağ</i> ‘dog’	<i>sā</i> ‘now’	<i>sāq</i> ‘tree trunk’
<i>pis</i> ‘son’	<i>pī</i> ‘camel fat’	<i>pīš</i> ‘date palm leaf’
<i>muxx</i> ‘head’	<i>mū</i> ‘hair’	<i>mūl</i> ‘totally’
<i>par</i> ‘feather’	<i>pā</i> ‘foot, leg’	<i>pāk</i> ‘clean’

2.4.2 Syllable shapes in multisyllabic words

Because of the rule about consonant clusters separating at syllable boundaries (see below), syllable shapes are much more limited in multisyllabic words. Syllables beginning with two consonants (CC) are found only in the initial syllable of single-morpheme words, and syllables ending in two consonants (CC) occur only in the final syllable of single-morpheme words.

The following word-initial syllable shapes are attested in multisyllabic words:

CV	<i>fu.rī</i> ‘clever goat’
CVV	<i>dā.drō</i> ‘short dhow sp.’
CVC	<i>nux.rīt</i> ‘nose’
CVVC	<i>rās.tağ</i> ‘straight direction’
CCVV	<i>škē.zī</i> ‘evening wind’
CCVC	<i>qbay.lī</i> ‘hospitable’
CCVVC	<i>qbāy.wā</i> ‘a little while ago’

Kumzari multisyllabic words have only three word-internal syllable shapes:

CV	<i>ē.ra.rağ</i> ‘bream (fish)’
----	--------------------------------

CVV	<i>daḡ.bē.rit</i> ‘thick, dusty haze’
CVC	<i>di.mis.tan</i> ‘winter’

Word-final syllables of only five shapes are found in Kumzari multisyllabic words:

CV	<i>qyā.sa</i> ‘opinion’
CVV	<i>čir.ba.sē</i> ‘slide, slippery rock used as slide’
CVC	<i>buk.rit</i> ‘newborn goat’
CVVC	<i>bur.qēb</i> ‘bull shark, small stage’
CVCC	<i>bar.zung</i> ‘cradle’

Words with more than three syllables are confined to ideophones, suspected compounds or borrowed words:

CVC.CVC.CVV.CVC	<i>innikdūnik</i> ‘in a flash’
CV.CV.CVV.CVC	<i>balalīyit</i> ‘vermicelli noodles’
CVV.CVC CVV.CVC	<i>hēnis wēnis</i> ‘anyone’
CV.CV.CVC.CV	<i>baladiyya</i> ‘garbage’
CVC.CVC CV.CV.CVC.CV	<i>qambab garagumba</i> ‘fish sp.’

2.5 Non-segmental phenomena

When morphemes are added to a word the new shape complies with syllable limitations through processes of assimilation, insertion, deletion, or resyllabification. Several processes may occur in the same word.

2.5.1 Assimilation

Words ending in the vowels *ū* and *ī* assimilate to their consonantal counterparts *w* and *y* when a suffix is added:

<i>rū</i> ‘face’	<i>rūw-ē</i> ‘a face’
<i>quḥū</i> ‘cough’	<i>quḥw-ē</i> ‘a cough’
<i>mū</i> ‘hair’	<i>mūw-ē</i> ‘a hair’
<i>gēdū</i> ‘water-pipe’	<i>gēdw-ē</i> ‘a water-pipe’
<i>qrādī</i> ‘bull shark’	<i>qrādy-ē</i> ‘a bull shark’
<i>kabī</i> ‘stay joint (boat)’	<i>kaby-ē</i> ‘a stay joint (boat)’
<i>qōṭī</i> ‘tin can’	<i>qōṭy-ē</i> ‘a tin can’
<i>ḥawī</i> ‘billy goat’	<i>ḥawly-ē</i> ‘a billy goat’

This rule extends to the enclitic subordinator *ā*; on *ā* following a back vowel an epenthetic *w* is inserted, and on *ā* following a front vowel an epenthetic *y* is inserted:

(1) B54

wa	bang-ō	wā
if/when	dusk –the	SUB
‘At dusk,’		

(2) K177

sā āw txōr-in pi ḥasy-ō wā
 now water drink:IMPF-3p from pool -the SUB
 ‘Now drinking water from the pool,’

(3) B382

pi yē čikk-ē yā
 from 3s little.bit -a SUB
 ‘In a little while,’

(4) R697

ba dist asrē yā
 to hand left SUB
 ‘To the left-hand side,’

Vowels in syllables adjacent to an added suffix may change to conform to the added suffix. In the case of the definite –ō, for example, *i* may be lowered to *ē*:

<i>nuxrit</i> ‘nose’	<i>nuxrēt-ō</i> ‘the nose’
<i>battil</i> ‘dhow’	<i>battēl-ō</i> ‘the dhow’
<i>maglis</i> ‘court’	<i>maglēs-ō</i> ‘the court’
<i>maddit</i> ‘while’	<i>maddēt-ō</i> ‘the while’
<i>bukrit</i> ‘newborn goat’	<i>bukrēt-ō</i> ‘the newborn goat’

and *a* raised and *u* lowered to *ō*:

<i>langal</i> ‘anchor’	<i>langōl-ō</i> ‘the anchor’
<i>dar</i> ‘door’	<i>dōr-ō</i> ‘the door’
<i>masbaḥ</i> ‘shower’	<i>masbōh-ō</i> ‘the shower’
<i>xānağ</i> ‘house’	<i>xānōğ-ō</i> ‘the house’
<i>bap</i> ‘father’	<i>bōp-ō</i> ‘the father’
<i>maḥbax</i> ‘kitchen’	<i>maḥbōx-ō</i> ‘the kitchen’
<i>dastur</i> ‘bowsprit’	<i>dastōr-ō</i> ‘the bowsprit’
<i>čum</i> ‘eye’	<i>čōm-ō</i> ‘the eye’
<i>ğambur</i> ‘niche’	<i>ğambōr-ō</i> ‘the niche’

In a single-syllable word, *i* is backed to *u* to assimilate to the definite suffix –ō:

<i>jism</i> ‘body’	<i>jusm-ō</i> ‘the body’
<i>rikd</i> ‘corner’	<i>rukḍ-ō</i> ‘the corner’
<i>sinn</i> ‘net anchor’	<i>sunṇ-ō</i> ‘the net anchor’

In the case of the short vowel in the final syllable being deleted, words of the form *maCCēCit* undergo a second process of lowering the *ē* to *a*, in order to lighten the syllable that now has a final consonant cluster:

<i>maqbērīt</i> ‘grave’	<i>maqbart-ō</i> ‘the grave’
<i>matrēqīt</i> ‘bobbin’	<i>matraqt-ō</i> ‘the bobbin’
<i>madrēsīt</i> ‘school’	<i>madrast-ō</i> ‘the school’

Consonants may assume the voicing qualities of consonants that are joined to them due to prior phonological processes. In the following word, the final short vowel is deleted when a suffix is added, causing devoicing of *b* to *p* next to *t*:

zēribit ‘goat pen’ *zēript-ō* ‘the goat pen’

Similarly, the addition of a suffix to the following word causes the form with an etymologically latent *k* to appear and the devoicing of *d* to *t* before *k*:

mard ‘man’ *martk-ō* ‘the man’

2.5.2 Insertion

An epenthetic vowel is added to alveolar-final verb roots²⁷ before the realis suffix *-d* and the perfect suffix *-s*. As shown in Table 4 and Table 5 below, verb roots ending in a single consonant take *u*, while verb roots ending in a consonant cluster take *ī* and the second consonant of the cluster is transferred to the suffix syllable.

Table 4. Verb roots with final single consonant and epenthetic *u*

verb root	gloss	realis <i>-d</i>	perfect <i>-s</i>
<i>dōz</i>	‘sew’	<i>dōzud</i>	<i>dōzus</i>
<i>kēš</i>	‘pull’	<i>kēšud</i>	<i>kēšus</i>
<i>gnūn</i>	‘believe’	<i>gnūnud</i>	<i>gnūnus</i>
<i>pōr</i>	‘fly’	<i>pōrud</i>	<i>pōrus</i>
<i>rāy</i>	‘be able’	<i>rāyud</i>	<i>rāyus</i>

Table 5. Verb roots with final consonant cluster and epenthetic *ī*

verb root	gloss	realis <i>-d</i>	perfect <i>-s</i>
<i>turs</i>	‘fear’	<i>tursīd</i>	<i>tursīs</i>
<i>gird</i>	‘go around’	<i>girdīd</i>	<i>girdīs</i>
<i>amš</i>	‘sweep’	<i>amšīd</i>	<i>amšīs</i>
<i>fōšn</i>	‘sell’	<i>fōšnīd</i>	<i>fōšnīs</i>
<i>ābn</i>	‘close’	<i>ābnīd</i>	<i>ābnīs</i>

2.5.3 Deletion

In verb roots ending in *r*, the *r* may be dropped before the realis suffix *-d* and the perfect suffix *-s*.

Table 6. Verb roots with *r* deletion

verb root	gloss	realis <i>-d</i>	perfect <i>-s</i>
<i>ambār</i>	‘load’	<i>ambād</i>	<i>ambās</i>
<i>xōr</i>	‘eat’	<i>xōd</i>	<i>xōs</i>
<i>gir</i>	‘take’	<i>gid</i>	<i>gis</i>
<i>ēnar</i>	‘hide’	<i>ēnid</i>	<i>ēnis</i>
<i>wār</i>	‘bring’	<i>wād</i>	<i>wās</i>

In some cases, the *r* is dropped before the perfect *-s* but retained as a flap (lenition) in the realis in the place of *-d*.

²⁷ Nasal- and liquid-final verb roots that are not deleted in realis and perfect forms also generally take an epenthetic vowel.

Table 7. Verb roots with *r* deletion and retention

verb root	gloss	realis – <i>d</i>	perfect – <i>s</i>
<i>būr</i>	‘become’	<i>būr</i>	<i>būs</i>
<i>dār</i>	‘give’	<i>dār</i>	<i>dās</i>
<i>ōdur</i>	‘hold on’	<i>ōdur</i>	<i>ōdus</i>

The imperfect prefix *t-* is deleted before alveolar-initial verb roots.

Table 8. Verb roots with *t* deletion

verb root	gloss	imperfect <i>t-</i>
<i>dān</i>	‘know’	<i>dān</i>
<i>jōr</i>	‘ask’	<i>jōr</i>
<i>sō</i>	‘put’	<i>sō</i>

The weak initial consonant *w* in verb roots tends to be deleted when the verb takes the imperfect prefix.

Table 9. Verb roots with *w* deletion

verb root	gloss	realis – <i>d</i>	perfect – <i>s</i>	imperfect <i>t-</i>
<i>wāt</i>	‘want’	<i>wātud</i>	<i>wātus</i>	<i>tāt</i>
<i>wašt, wēl</i>	‘let’	<i>wašt</i>	<i>wašt</i>	<i>tēl</i>
<i>wār</i>	‘bring’	<i>wād</i>	<i>wās</i>	<i>tār</i>

In initial position in an unstressed word such as the third-person singular pronoun *yē*, following a short-vowel-final word the semivowel *y* may be elided:

ba yē ‘to it’ > *bē*
pi yē ‘from it’ > *pē*
inda yē ‘inside it’ > *indē*

2.5.4 Resyllabification

There are no phonemic syllabic consonants. When an affix or enclitic is added to a morpheme with a consonant cluster, the final consonant joins the affixed syllable:

CVCC	<i>bukr</i> ‘firstborn’	CVC.CVV	<i>bukr -ō</i> ‘the firstborn’
CCVCC	<i>čuḡl</i> ‘thing’	CCVC.CVV	<i>čuḡl -ē</i> ‘a thing’
CVCC	<i>išk</i> ‘dry’	CVC.CVC	<i>išk =in</i> ‘they are dry’

In words of Semitic origin with the source prefix *mu-*, if the syllable following *mu-* has a long vowel or a geminated coda, the *u* is deleted, the *m* becoming part of the initial consonant cluster in the Kumzari word:

mḥāfiḍ ‘governor’
mrād ‘reason’
mqaşş ‘scissors’
mşaww ‘barnacle’

2.5.5 Multiple processes

In many cases, several phonological processes occur with the addition of a suffix in keeping with phonotactic constraints. Rigid stress and syllable rules govern vowels in Kumzari words.

When a suffix is added, penultimate syllable stress is maintained through vowel changes, insertions, and deletions.

All three processes occur in the following word: raising and fronting of the *a* to *ē*, deletion of the *i*, and insertion of an epenthetic *u*:

tarqit ‘wedding poem’ *tēruqt-ō* ‘the wedding poem’

In a process of resyllabification, when labial-initial verb roots are prefixed with the imperfect *t-*, the verb root vowel becomes an epenthetic vowel inserted after the prefix, stress shifts to the prefix, and the verb root vowel is rounded to *u*:

verb root	gloss	imperfect <i>t-</i> verb stem
<i>bēr</i>	‘carry’	<i>tē-bur</i>
<i>mān</i>	‘stay’	<i>tā-mun</i>
<i>mēš</i>	‘look at’	<i>tē-muš</i>
<i>mur</i>	‘die’	<i>tu-mur</i>
<i>pōr</i>	‘fly’	<i>tō-pur</i>

In an open syllable, a final short *a* may inherit stress and be lengthened when a suffix is added to the word:

<i>kahraba</i> ‘electricity’	<i>kahrabā-ō</i> ‘the electricity’
<i>lumba</i> ‘sea urchin’	<i>lumbā-ō</i> ‘the sea urchin’
<i>nwāxa</i> ‘ship captain’	<i>nwāxā-ō</i> ‘the ship captain’
<i>sandaqa</i> ‘goat pen’	<i>sandaqā-ō</i> ‘the goat pen’

or the final vowel may be deleted:

<i>tēra</i> ‘way’	<i>tēr-ō</i> ‘the way’	<i>tēr-ē</i> ‘a way’
<i>xumba</i> ‘clay storage jar’	<i>xumb-ō</i> ‘the clay storage jar’	<i>xumb-ē</i> ‘a clay storage jar’
<i>zīla</i> ‘bailing bucket’	<i>zīl-ō</i> ‘the bailing bucket’	<i>zīl-ē</i> ‘a bailing bucket’

In a closed syllable adjacent to an added suffix, a short vowel may be either lengthened:

<i>surban</i> ‘roof’	<i>surbān-ō</i> ‘the roof’
<i>ṣafṣuf</i> ‘sparrow’	<i>ṣafṣūf-ō</i> ‘the sparrow’
<i>nišbīl</i> ‘fishing line’	<i>nišbīl-ō</i> ‘the fishing line’
<i>intaḥ</i> ‘sun’	<i>intāḥ-ō</i> ‘the sun’
<i>ḥambul</i> ‘newborn’	<i>ḥambūl-ō</i> ‘the newborn’

or deleted:

<i>linkit</i> ‘finger’	<i>linkt-ō</i> ‘the finger’
<i>qāwīl</i> ‘trader’	<i>qāwl-ō</i> ‘the trader’
<i>gōsin</i> ‘goat’	<i>gōsn-ō</i> ‘the goat’
<i>jinjāwir</i> ‘master sorcerer’	<i>jinjāwr-ō</i> ‘the master sorcerer’
<i>ḥējub</i> ‘eyebrow’	<i>ḥējb-ō</i> ‘the eyebrow’

Similar processes are found in the South Arabian languages (in Mehri the definite prefix is *a-* or *ḥə-*):

Mehri Kumzari	gəzáyɾət jəzɾit	a-gzáyɾət jēzɾt-ō	‘the island’ (Rubin 2010:69) ‘the island’
Mehri Kumzari	bəḵār bāqara	a-bḵār bāqarā-ō	‘the cows’ (Rubin 2010:69) ‘the cow’
Mehri Kumzari	məsgēd muzgīt	a-msgēd muzgēt-ō	‘the mosque’ (Rubin 2010:70) ‘the mosque’
Mehri Kumzari	fəɾōḵ famfōxit	ḥə-frōḵ famfōxt-ō	‘the flocks, the camps’ (Rubin 2010:70) ‘the goose egg (swollen bump)’
Mehri Kumzari	səlōb salah	ḥə-slōb slāḥ-an	‘the weapons’ (Rubin 2010:71) ‘the weapons’

Table 10 sets out the phonemic writing system used in this book and their equivalents in the Arabic-based Kumzari script²⁸ and in the International Phonetic Alphabet (IPA).

²⁸ As developed by Erik John Anonby in conjunction with and approved by the Kumzari community, March 2009.

Table 10. Kumzari sounds, transcription, and writing system

IPA	phonemic transcription	Kumzari script
p	p	پ
b	b	ب
t	t	ت
d	d	د
tʰ	tʰ	ط
dʰ	dʰ	ض
k	k	ک
g	g	چ
q	q	ق
ʔ	ʔ	ء
tʃ	tʃ	چھ
dʒ	dʒ	جھ
f	f	ف
s	s	س
sʰ	sʰ	ص
zʰ	zʰ	ظ
ʃ	ʃ	ش
χ	x	خ
ʁ	ʁ	ځ
ħ	ħ	ح
ɦ	ɦ	ه
m	m	م
n	n	ن
w	w	و
l	l	ل
ɭ	r	ر
ɭʰ	ɭʰ	لھ
j	y	ي
i:	ī	ي
u:	ū	و
e:	ē	ي
o:	o	و
a:	ā	ا
ɪ	i	ي
ʊ	u	و
ə	a	و

3 Noun

3.1 Noun: Introduction

Nouns in Kumzari prototypically express time-stable concepts. They can be inflected for definiteness and number, and may be derived to form items of other word classes. Nouns can function as the head of a noun phrase, as the subject or object of a clause, as a modifier in a possessive noun phrase, or as the complement of a prepositional phrase or of a predicate with an existential enclitic.

3.1.1 Structural properties of nouns

3.1.1.1 Noun root structure

The noun in Kumzari consists of one, two, or three syllables.²⁹ There are sixty noun root shapes: monosyllabic nouns have ten different root shapes, disyllabic nouns have twenty-four, and trisyllabic nouns have twenty-six different root shapes.³⁰

Root shapes in monosyllabic nouns:

CVC	<i>dil</i>	‘heart’
CCVC	<i>stağ</i>	‘date pit’
CVCC	<i>čupş</i>	‘lobster’
CCVCC	<i>brinž</i>	‘rice’
CVV	<i>jā</i>	‘barley’
CCVV	<i>črā</i>	‘lamp’
CVVC	<i>čāf</i>	‘beach’
CCVVC	<i>ğmūt</i>	‘gills’
CVVCC	<i>qāpṭ</i>	‘white fish sp.’
CCVVCC	<i>stārg</i>	‘star’

Root shapes in disyllabic nouns :

CV.CVV	<i>furī</i>	‘clever goat’
CV.CVC	<i>bağal</i>	‘upper arm’
CVC.CV	<i>lumba</i>	‘sea urchin’
CVC.CVV	<i>farzē</i>	‘nostril-stud’
CVC.CVC	<i>nišbil</i>	‘fishing line’
CVC.CVVC	<i>burqēb</i>	‘small bull shark’
CVC.CCV	<i>sinsla</i>	‘metal chain’
CVC.CVCC	<i>baržung</i>	‘cradle’
CCVC.CV	<i>zğurda</i>	‘sheer strake (boat plank sp.)’
CCVC.CVV	<i>mdallē</i>	‘coffee pot’
CCVC.CVC	<i>qrambiş</i>	‘needlefish sp.’
CVCC.CVC	<i>sanksar</i> ³¹	‘red bream fish sp.’

²⁹ In addition, there is a single example of a quadrisyllabic noun represented in the data: CV.CV.CVC.CV *garagumba* ‘fish sp.’ This word is probably a historical compound.

³⁰ Root shapes are ordered in these lists firstly by vowel length (V before VV), secondly by syllable type (open before closed), thirdly by coda (C before CC), and fourthly by onset (C before CC).

³¹ It is rare to have three consecutive consonants in Kumzari; this form is most likely the result of a historical compound, perhaps of *sangī* ‘heavy’ + *sar* ‘head.’

CVV.CV	<i>tēxa</i>	‘goat hair rope’
CVV.CVV	<i>jīrī</i>	‘sand’
CVV.CVC	<i>lētab</i>	‘wild fig’
CVV.CVCC	<i>bālišṭ</i>	‘pillow’
CVV.CVVC	<i>dāzūd</i>	‘turmeric’
CCVV.CV	<i>twāra</i>	‘shelter’
CCVV.CVV	<i>škēzī</i>	‘evening wind’
CCVV.CVC	<i>nwāšam</i>	‘evening’
CVVC.CV	<i>jērda</i>	‘sluice irrigation’
CVVC.CVV	<i>jāzrī</i>	‘northeast wind’
CVVC.CVC	<i>kāsrit</i>	‘small drum’
CVVC.CVVC	<i>bānnāt</i>	‘tradition’

Root shapes in trisyllabic nouns:

CV.CV.CV	<i>lagana</i>	‘bowl’
CV.CV.CVVC	<i>xanaḥṭis</i>	‘sideburn’
CV.CVV.CV	<i>čigāra</i> ³²	‘cigarette’
CV.CVV.CVV	<i>tiḥādī</i>	‘rainbow runner fish sp.’
CV.CVC.CV	<i>kuruxda</i> ³³	‘old person’
CV.CVV.CVC	<i>mazāraq</i>	‘fish sp.’
CV.CVC.CVC	<i>dimistan</i>	‘winter’
CV.CVCC.CVV	<i>aruḥṭē</i>	‘smell’
CVC.CV.CV	<i>qandaḥa</i>	‘rainbow’
CVC.CV.CVV	<i>čirbasē</i>	‘slide, slippery rock’
CVC.CV.CVC	<i>silsilit</i>	‘descendant’
CVC.CVV.CV	<i>bandēra</i> ³⁴	‘flag (of ship)’
CVC.CVV.CVV	<i>tambūlī</i>	‘goat wattles’
CVC.CVV.CVC	<i>zangērīr</i>	‘slave’
CVC.CVC.CV	<i>mayšaṭṭa</i>	‘palm frond broom’
CVC.CVC.CVV	<i>purwandō</i>	‘lateen yard’
CVC.CVC.CVC	<i>kilwiskit</i>	‘date sp.’
CVV.CV.CV	<i>ḡābana</i>	‘inlet’
CVV.CV.CVV	<i>māšuwē</i>	‘skiff’
CVV.CV.CVC	<i>zēpilaḡ</i>	‘wooden chest’
CVV.CVV.CVC	<i>šādī’it</i>	‘argument’
CCVV.CV.CVV	<i>mwāṭanī</i>	‘cherished person’
CCVV.CV.CVC	<i>mqābalit</i>	‘meeting’
CVVC.CV.CVV	<i>dīsdimī</i>	‘latter keel’
CVVC.CVV.CVV	<i>‘ēlyē’ō</i>	‘foremost portside tholepin’
CVVC.CVC.CV	<i>sīflindō</i>	‘silver eel’

3.1.1.2 Noun inflection

The Kumzari noun is optionally marked for definiteness (by the suffix *–ō*)³⁵ or indefiniteness (the suffix *–ē*), where a lack of any such suffix denotes a generic noun. Plurality is marked

³² This word’s origin is English presumed to be via Gulf Arabic.

³³ This word’s origin is probably the Persian word *kadxōdā* ‘chieftain.’

³⁴ This word’s origin is most likely Portuguese *bandeira* ‘flag.’