Xumi (part 2)

Katia Chirkova, Yiya Chen, Tanja Kocjančič Antolík

Xumi, Part 2: Upper Xumi, the Variety of the Upper Reaches of the Shuiluo River

Katia Chirkova (CNRS-CRLAO), Yiya Chen (Leiden University), and Tanja Kocjančič Antolík (CNRS-LPP)

This overview focuses on the variety of the Xumi language (旭米, /^{LP}ʂu ketɕɜ/ or /^{EP}ʂu-hĩ ketɕɜ/) that is spoken in the upper reaches of the Shuiluo River (水洛河) in Shuiluo Township (水洛乡) (hereafter Upper Xumi).[1] The township is located in Muli Tibetan Autonomous County (木里藏族自治县, *smi li rang skyong rdzong* in Written Tibetan [hereafter, WT]), in the South-West of Sichuan Province (四川省) in the People's Republic of China.

Upper Xumi is one of the two dialects of the Xumi language, the other one being Lower Xumi. That dialect is spoken in the lower and middle reaches of the Shuiluo river, as we discussed in our phonetic and phonological sketch of Lower Xumi (Chirkova & Chen, submitted). The present overview of Upper Xumi complements our overview of Lower Xumi.

The two dialects of Xumi are closely related to each other, but they are distinctly different, due to their contact with different languages. Upper Xumi is essentially influenced by the local dialect of Tibetan (Gami Tibetan), as well as, to a lesser extent, by Pumi. By contrast, Lower Xumi is mostly influenced by the local dialects of the Pumi, Mosuo, and Naxi languages. The observed diversity is highly remarkable given:

(a) the small geographical area occupied by the group (there is a mere 15.25 km beeline between the northernmost village of Lanman (兰满村) in the upper reaches of Shuiluo river and the southernmost village of Mianbang (免邦村) in the lower reaches of Shuiluo river);

(b) the small amount of speakers (less than 2,000 people);

(c) the continuing contact and exchange between villages in the upper, middle, and lower reaches of Shuiluo river.

The comparative data on the two varieties are of interest for studies on language change in progress, and they may provide insight into the mechanisms at work in the implementation of language change and the diffusion of linguistic innovations.

The present sketch provides an overview of Upper Xumi with an emphasis on those features that distinguish it from Lower Xumi. It also contains an instrumental study of the voiceless-voiced contrast in nasals and laterals (nasal and oral airflow and electroglottographic measures), made possible by the fact that we could work with our principal Upper Xumi language consultant in a phonetics laboratory.

The description is based on the first author's fieldwork. The word list and the text provided with this paper were recited by a fifty-nine year old male native speaker of Xumi, who was born and raised in Lanman village. To facilitate comparison with Lower Xumi, the present description uses (whenever available) cognates in both varieties as illustrative examples. (In the sound archive, sound files with Lower Xumi forms are marked with "LX".)

Consonants

The consonant inventory of Upper Xumi comprises of 44 consonants, as listed in the table below. (Low frequency phonemes are given in parentheses. See section "Prosodic organization" for tonal variation on monosyllabic words.) Appendix 1 provides the distributional patterns of consonants and vowels that include permissible sequences in syllables as attested in our corpus of ca. 2,800 words. Upper Xumi is characterized by a reduced segmental inventory (44 consonants instead of 50 in Lower Xumi), different phonotactic constraints, and the presence of prenasalized clusters.

	Bilabial	**Alveolar**	**Retroflex**	**Alveolo-palatal**	**Velar**	**Uvular**	**Glottal**
Plosive	p pʰ b	t tʰ d			k kʰ g	q qʰ	
Affricate		ts tsʰ dz	tʂ tʂʰ dʐ	tɕ tɕʰ dʑ			
Nasal	(m̥) m	(n̥) n		ɲ	ŋ		
Fricative		s z	ʂ ʐ	ɕ ʑ	x	(χ) ʁ	h (ɦ)
Approximant	w	ɹ		j			
Lateral approximant		l̥ l		ʎ̥ ʎ			

p	ᴴpɜ	‘to speak’	ʂ	ᴴʂɻ̩	‘fishing net’
pʰ	ᴴpʰɜ	‘price’	ʐ	ᴿᴾlɐ-ʐɻ	‘to sleep’
b	ᴴbɜ	‘(house) beam’	tɕ	ᴴtɕɜ	‘star’
m	ᴿmɜ	‘sky’	tɕʰ	ᴿtɕʰɜ	‘muntjac’
m̥	ᴴm̥ɛ̃	‘medicine’	dʑ	ᴴdʑɜ	‘water’
w	ᴴwɜ	‘tooth’	ɲ	ᴿɲɜ	‘fire’
t	ᴴtjɛ	‘cutting board’	ɕ	ᴿᴾdʑi-ɕɜ	‘one hundred’
tʰ	ᴴtʰjɛ	‘to wear (a hat)’	ʑ	ᴴʑɜ	‘beer, wine’
d	ᴴdjɛ	‘fox’	j	ᴴjɜ	‘tent’
ts	ᴿᴾtsɐ-tsɜ	‘to bundle’	ʎ	ᴴʎɛ	‘correct, right’
tsʰ	ᴴtsʰɜ	‘salt’	ʎ̥	ᴴʎ̥ɛ	‘flavorless’

dz	^{R}dzɜ	‘to eat’	k	^{H}ku	‘to be able’
n	^{R}nɔ	‘whole; complete’	k^{h}	$^{R}k^{h}u$	‘want’
n̥	^{R}n̥õ	‘fur, animal hair’	g	^{H}gu	‘clothes’
ɹ	^{R}ɹɐ	‘chicken’	ŋ	^{H}ŋɜ	‘I’
s	^{H}sɜ	‘to know’	x	^{R}xu	‘to go (IMP)’
z	^{R}zɜ	‘to wash’	q	^{H}qu	‘hearth’
l	^{H}lɐ	‘tiger’	q^{h}	$^{H}q^{h}u$	‘year’
l̥	^{EP}bə-l̥ɐ	‘to open a lock’[2]	χ	^{RP}lɐ-χɛ̃	‘to cause to break’
tʂ	^{RP}dʑi-tʂu	‘unit of length’[3]	ʁ	^{RP}lɐ-ʁɛ̃	‘to break, to be broken’
$tʂ^{h}$	$^{LP}tʂ^{h}u$-$tʂ^{h}u$	‘to suck’	h	^{H}hõ	‘vegetable’
dz̩	^{RP}lɐ-dz̩u	‘to be tired’	ɦ	^{H}ɦõ	‘to dare’

There is a general three-way contrast in stops and affricates: voiceless aspirated, voiceless unaspirated, and voiced. Velars and uvular stops are in near complementary distribution. Uvular stops are found before non-high vowels, and velar stops are found elsewhere. The two series contrast before /e, ɜ, u, õ/, e.g. /$^{R}k^{h}$e/ ‘tax’ (WT *khral*?) vs. /$^{R}q^{h}$e/ ‘lime’, /$^{H}k^{h}$ɜ/ ‘foot’ vs. /$^{H}q^{h}$ɜ/ ‘feces’, /^{H}ku/ ‘to be able’ vs. /^{H}qu/ ‘hearth’, /^{RP}ɲɜ-kõ/ ‘to warm oneself’ vs. /^{H}qõ/ ‘fate, life’. Dissimilar to Lower Xumi, Upper Xumi does not have a voiced uvular stop.

Words with the voiced uvular stop in Lower Xumi correspond to words with the cluster /Ng/ in Upper Xumi (see on prenasalized clusters below).

Fricatives are pronounced at six different places of articulation: alveolar, retroflex, alveolopalatal, velar, uvular, and glottal. Voiceless velar and glottal fricatives are in complementary distribution, the former only occur with oral vowels, whereas the latter only occur with nasal vowels. For example, /^{H}xɔ/ 'cooked rice, food', /^{H}hɔ̃/ 'to stretch'. /ɦ/ has the most restricted distribution of all fricatives. It co-occurs only with the nasal vowels /ɔ̃/ and /ɛ̃/. Examples include: /^{H}ɦɔ̃/ 'to dare' (compare with /^{H}ɔ̃/ [^{H}ʔɔ̃] 'self (first person subject)'), /^{R}ɦɛ̃/ 'self (third person subject)' (compare with /^{H}ɛ̃/ [^{H}ʔɛ̃] 'sheep'). Dissimilar to Lower Xumi, Upper Xumi does not have a voiced velar fricative, but it has the voiceless uvular fricative /χ/ as an independent phoneme, which can occur word-initially, e.g. /^{LP}χɜ-χɐ/ 'to itch' (cf. Lower Xumi /^{LP}qʰɐ-χɐ/), /^{H}χɐ/ 'difficulty' (WT *dka'*).

Xumi nasals and laterals contrast (1) alveolar and alveolopalatal places of articulation and (2) voiced and voiceless counterparts.

Alveolar and alveolopalatal laterals and nasals are minimally distinguished before /u/; in addition, alveolar and alveolopalatal nasals are also distinguished before /ɐ/. Examples include: /^{LP}ɲu-ɲu/ 'breast' vs. /^{LP}nu-tʂʰu/ 'bean curd', /^{R}lu/ 'again' vs. /^{R}ʎu/ 'come (IMP)', /^{RP}nɐʁɐ/ 'twenty' vs. /^{LP}ɲɐʁɔ/ 'shadow'. For nasals, the contrast is not found before the vowels /i, ɛ, ɜ/, and for laterals, before /ɛ/. In the above cases, we use the symbols for the alveolopalatal nasals and laterals, respectively. Example include: /^{H}ɲi/ 'you, thou', /^{H}ɲe/

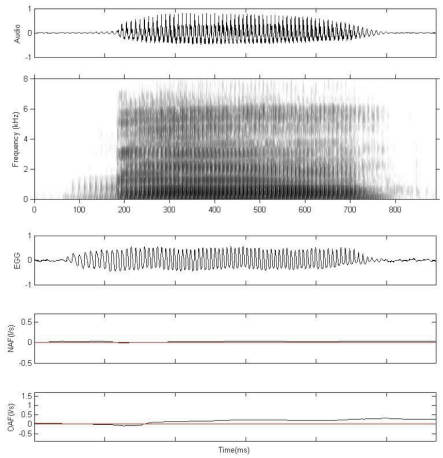
'snivel, snot', /^{H}ɲɛ/ 'milk', /^{H}ɲɜ/ 'to have a craving (for something), to be hungry'; /^{R}ʎɛ/ 'predestined affinity'.

Bilabial and alveolar nasals show a clear voiced vs. voiceless contrast as confirmed with our electroglottography (EGG) data (see Figures 1-2). Voiceless nasals are infrequent. They mostly occur in loanwords from Tibetan, but they are also attested in the native vocabulary, e.g. /^{RP}m̥jɛtsũ/ [^{RP}m̥jætsũ] 'tail', /^{R}n̥ɔ̃/ 'fur, animal hair'.[4] Upper Xumi has two voiceless nasals (/m̥, n̥/), which is two voiceless nasals less than in Lower Xumi (/m̥, n̥, ɲ̥, ŋ̊/). This is surprising, given that Upper Xumi is in closer contact with a local Tibetan dialect than Lower Xumi and that that local Tibetan dialect has four voiceless nasals (/m̥, n̥, ɲ̥, ŋ̊/, Chirkova 2014). Moreover, /^{EP}l̥ɐɻmu/ (WT *lha rnga mo*), the word for 'camel', is distinct in Upper Xumi not only from the corresponding form in Lower Xumi (/^{EP}ŋ̊ɑmõ/ 'camel'), but also from that in Gami Tibetan, /^{H}ŋamɔ̃/ (WT *rnga mong*), suggesting distinct donor languages for this form.

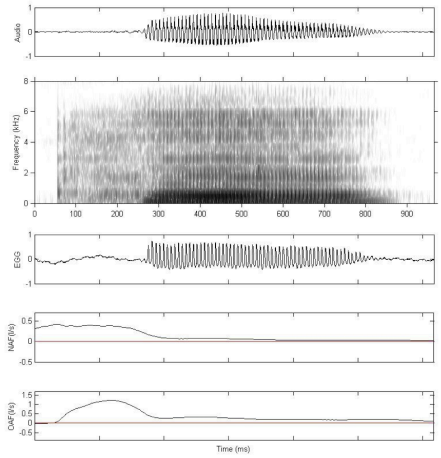
The table below provides the measured values of nasal airflow (NAF) and oral airflow (OAF) for the three (near) minimal pairs: /^{H}mjɛ/ 'bamboo' vs. /^{H}m̥ɛ̃/ 'medicine' (WT *sman*), /^{RP}nɐʁɐ/ 'twenty' vs. /^{H}n̥ɐ/ 'incantation, curse' (WT *sngags?*), and /^{R}nɔ/ 'whole, complete' vs. /^{R}n̥ɔ̃/ 'fur, animal hair'. In the table, NAFmin = minimum NAF; NAFmax = maximum NAF; NAFmean = mean NAF value; the provided measures represent an average of two to three repetitions and are measured during the nasal closure only.

Word	**NAF min (l/s)**	**NAF max (l/s)**	**NAF mean (l/s)**	**OAF min (l/s)**	**OAF max (l/s)**	**OAF mean (l/s)**
/^{H}mjɛ/ ‘bamboo’	0,01311	0,04695	0,02849	-0,05225	0,0166	-0,010466
/^{H}m̥ɛ̃/ ‘medicine’	0,24425	0,38435	0,34294	0,10735	1,1086	0,8391
/^{RP}nɐʁɐ/ ‘twenty’	0,04065	0,1048	0,06417	-0,21485	-0,09365	-0,171984
/^{H}n̥ɐ̃/ ‘incantation’	0,24227	0,43413	0,35542	0,01703	0,75213	0,59644
/^{R}nɔ/ ‘whole, complete’	0,02163	0,0574	0,03928	-0,09573	-0,01457	-0,07147
/^{R}n̥ɔ̃/ ‘fur, animal hair’	0,25163	0,4995	0,40140	0,02893	0,67113	0,31034

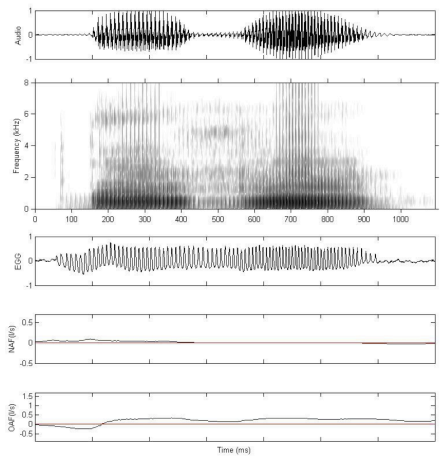
Figure 1 provides waveforms, spectrograms, EGG, NAF, and OAF for the three pairs.



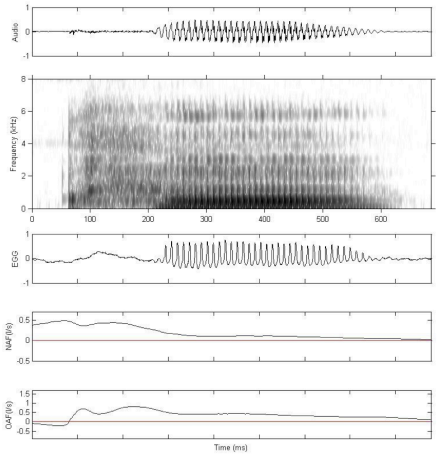
/^{H}mjɛ/ ‘bamboo’



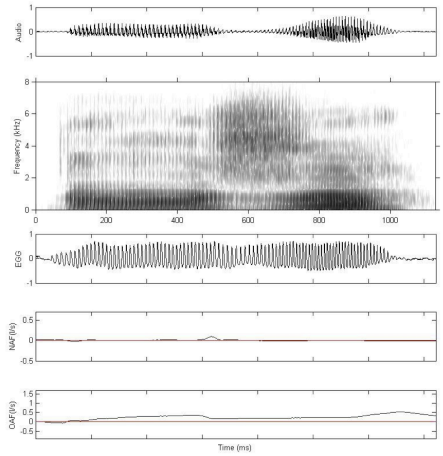
/^{H}m̥ɛ̃/ ‘medicine’



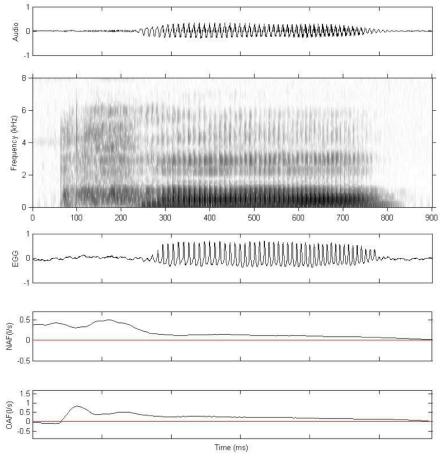
/[RP]nɐʁɐ/ ‘twenty’



/[H]n̥ɐ̃/ ‘incantation’



/[R]nɔ/ ‘whole, complete’



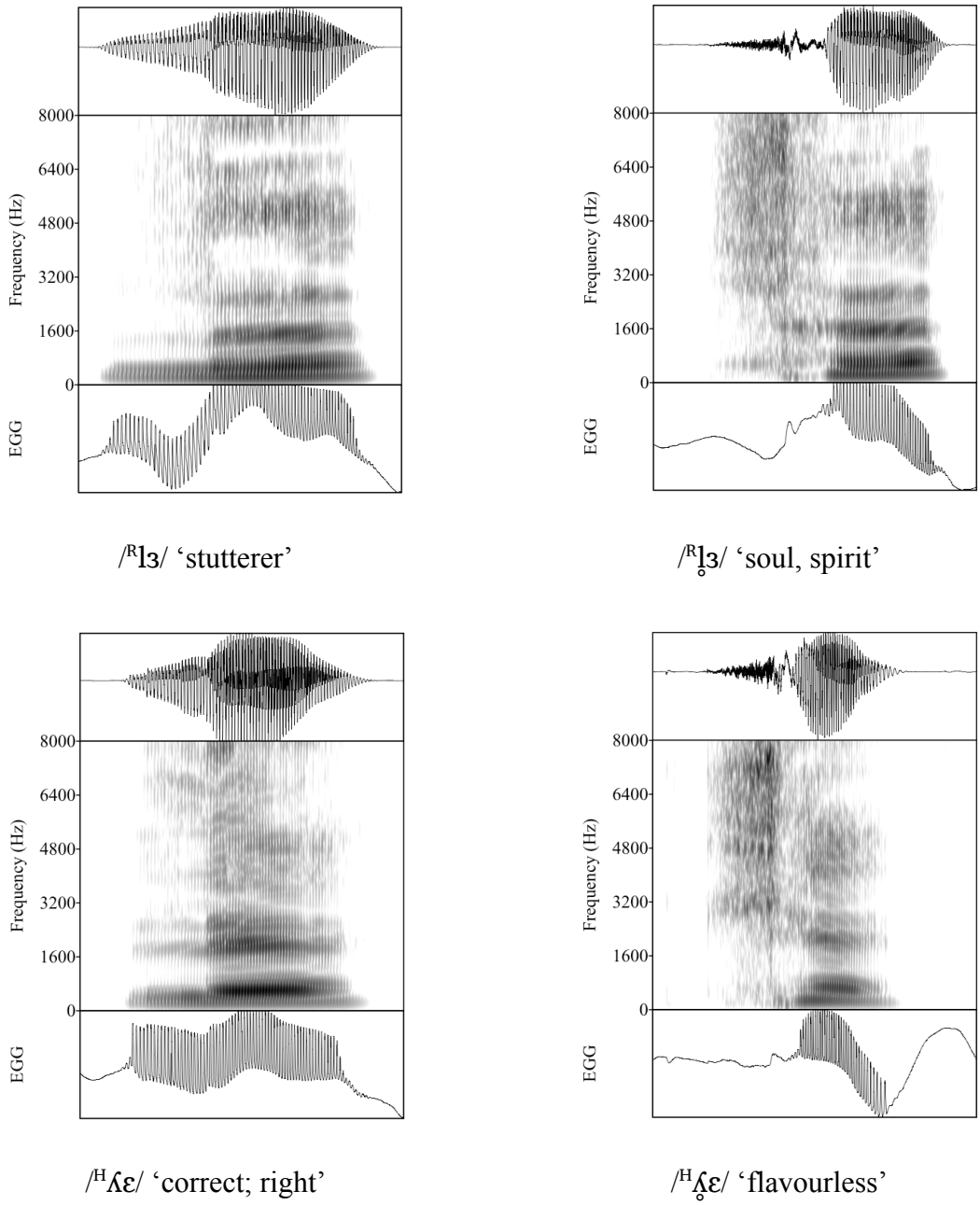
/[R]n̥ɔ̃/ ‘fur, animal hair’

Figure 1. Waveforms, spectrograms, EGG, NAF, and OAF for the pairs /[H]mjɛ/ ‘bamboo’ vs. /[H]m̥ɛ̃/ ‘medicine’, /[RP]nɐʁɐ/ ‘twenty’ vs. /[H]n̥ɐ̃/ ‘incantation’, and /[R]nɔ/ ‘whole, complete’ vs. /[R]n̥ɔ̃/ ‘fur, animal hair’

A surprising finding of this pilot investigation is that all values are greater during the articulation of voiceless nasals. More precisely, while in the articulation of voiceless nasals the air escapes both through the nasal and oral cavity, there is almost no airflow through either nasal or oral cavity for the voiced nasals. Piezoelectric data, recorded for the same corpus, confirms that in voiced nasals, there is resonance recorded in the nose but no nasal

(and no oral) airflow. A more detailed study (in preparation) will hopefully provide a more accurate description of this potentially new kind of nasals.

All examples of the voiceless laterals include frication and an abrupt release of the front closure that occurs slightly before the onset of voicing for the vowel (see Figure 2):



/ᴿlɜ/ ‘stutterer’ /ᴿl̥ɜ/ ‘soul, spirit’

/ᴴʌɛ/ ‘correct; right’ /ᴴʌ̊ɛ/ ‘flavourless’

Figure 2. Waveforms, spectrograms, and EGG of the minimal pairs /ᴿlɜ/ ‘stutterer’ vs. /ᴿl̥ɜ/ ‘soul, spirit’, and /ᴴʌɛ/ ‘correct; right’ vs. /ᴴʌ̊ɛ/ ‘flavourless’

The burst separates the segment into two parts with different spectral structure. The pre-burst part has the greatest energy centered around 3000 Hz and greater amplitude of frequencies above that peak. The post-burst part has the noise present more equally across the frequencies, with minor peaks at the frequencies corresponding to the formant peaks of the following vowel. In addition, the post-burst part resembles the spectral structure of the voiced lateral counterpart. The presence of the burst and the spectral structure of the post-burst part suggest that these segments have developed from clusters.

Syllabic consonant

As a syllabic consonant, /ɻ̩/ may occur with a zero-initial (e.g. /^{RP}kʰu-ɻ̩/ ‘to be in debt’), and after alveolar sibilants and retroflexes (attested in our corpus after /ts, tsʰ, dz, s, z, tʂʰ, ʂ, ʐ/). We take the position that this syllabic consonant is phonologically the same as the fricative vowel after alveolar sibilants and retroflexes (/s, z, ts, tsʰ, dz, ʃ, ʒ, tʃʰ, tʂʰ/). In the latter case, /ɻ̩/ is realized as homorganic to the preceding consonant onset. Consider the following examples: /^{H}tsɻ̩/ [^{H}tsʐ̩] ‘to use’, /^{H}tsʰɻ̩/ [^{H}tsʰʐ̩] ‘to cut with scissors’, /^{H}dzɻ̩/ [^{H}dzʐ̩] ‘wheat’, /^{RP}mɐsɻ̩/ [^{RP}mɐsʐ̩] ‘tomorrow’, /^{RP}mɐzɻ̩/ [^{RP}mɐsʐ̩] ‘cat’, /^{H}tʂʰɻ̩/ [^{H}tʂʰʐ̩] ‘to sell’, /^{H}ʂɻ̩/ [^{H}ʂʐ̩] ‘fishing net’, /^{R}ʐ̩ɻ̩/ [^{R}ʐ̩ʐ̩] ‘to sleep’. Similar to the distribution of the syllabic consonant /ɻ̩/ in Lower Xumi, /ɻ̩/ in Upper Xumi contrasts with the high vowels (/i, u/), as in /^{H}tsi/ ‘lock’, /^{H}tsɻ̩/ ‘to use’, /^{H}tsu/ ‘to pluck (facial hair)’. We therefore analyse it as a separate phoneme in this language.

Consonant clusters

Clusters with /w/

The approximant /w/ occurs in the second position in consonant clusters, where it may be realized as secondary labialization of the first position consonant. /w/ occurs after alveolars, alveolopalatals, retroflexes, velars, uvulars, as well as after /ŋ/ and /ʁ/. It may be followed by /i, e, ɛ, ɜ, ɐ, ɛ̃/ as well as, marginally, also by /ĩ/. If preceded by an alveolopalatal onset or followed by the front vowels /i, e, ɛ, ĩ/, /w/ is realized as [ɥ], e.g. /[H]gwi/ [[H]gɥi] ‘bundle wrapped in cloth’, /[R]dwe/ [[R]dɥe] ‘to ask’, /[R]dʑwɛ/ [[R]dʑɥɛ] ‘bird’, /[R]dʑwɜ/ [[R]dʑɥɜ] ‘honey, sugar’, /[EP]gwɛ-wɜ/ [[EP]gɥɛ-wɜ] ‘to hunt’, /[EP]lɐ-gwɜ/ [[EP]lɐ-gɥɜ] ‘to hide something’, /[R]ʁwɜ/ ‘copper’, /[R]ʁwɐ/ ‘heavy’, /[H]qwɐ/ ‘to cry, to weep’.

Prenasalized clusters

Similar to its linguistic neighbour, Gami Tibetan (Chirkova 2014), Upper Xumi has prenasalized voiced stops and affricates clusters (/Nb, Nd, Ndz, Ndʑ, Ndʐ, Ng/). The stop component of the cluster may only be voiced, and a sequence of a nasal and a stop must be homorganic.[5] Prenasalization in Upper Xumi is contrastive, as evidenced by the (near) minimal pairs below:

/[LP]bemi/ ‘axe’ /[H]Nbe/ ‘saddle pad’ (WT *'bol* ‘cushion’)

/[R]dze/ ‘to exist (at a fixed location)’ /[R]Ndze/ ‘board for pressing things together’

/[H]dzָɔ/ ‘enemy’ (WT *dgra*)

/[R]Ndzָɔ-Ndzָɔ/ ‘to be similar’ (WT *'dra*)[6]

/[RP]lɐ-gɐ/ ‘to obstruct, to block’

/[RP]lɐ-Ngɐ/ ‘to be scorched’

Prenasalized clusters mostly occur in Tibetan loanwords (corresponding Gami sound files have been provided whenever available), and they are likely to be borrowed due to widespread Xumi-Gami Tibetan bilingualism in the upper reaches of the Shuiluo River. Examples include: /[RP]Ndiwu/ ‘bullet’ (WT *mde'u*, Gami /[H]Ndi/), /[R]Ndzָɔ-Ndzָɔ/ ‘to be similar’ (WT *'dra*, Gami /[H]Ndzָɔ/). (Compare the corresponding words in Lower Xumi: /[RP]diwu/ ‘bullet’, /[EP]dzָe-hĩ/ ‘similar’, with the nominalizing marker /hĩ/.) Prenasalized clusters are also attested in Xumi native vocabulary. We note that such words (invariably with the cluster /Ng/) are cognates with Lower Xumi forms with the voiced uvular stop, as in ‘to stew’: Upper Xumi /[R]Ngu/, Lower Xumi /[R]ɢo/. As suggested by the external reviewer of this paper, the presence of Tibetan loanwords containing the cluster /Ng/ has possibly paved the way for the reinterpretation of the voiced uvular stop as a /Ng/ cluster in Upper Xumi. We note that such a development would be consistent with the correlation between stops and voicing, as discussed in Ohala (1983:194-201). Given that the farther forward in the vocal tract a stop is articulated, the better able it is to accommodate voicing, /ɢ/ is the most difficult stop to voice. The prenasalized stop /Ng/ avoids the problems associated with maintaining voicing, while maintaining the original phonemic contrast between /g/ and /ɢ/.

Consonant lenition

Similar to Lower Xumi, Upper Xumi has a set of productive lenition rules, which transform some intervocalic voiced and voiceless aspirated stops and affricates into spirants. However, compared to the lenition rules in Lower Xumi, those in Upper Xumi are reduced in number and solely yield independent full-fledged phonemes that can occur word-initially. The attested lenition rules are summarized in the table below.

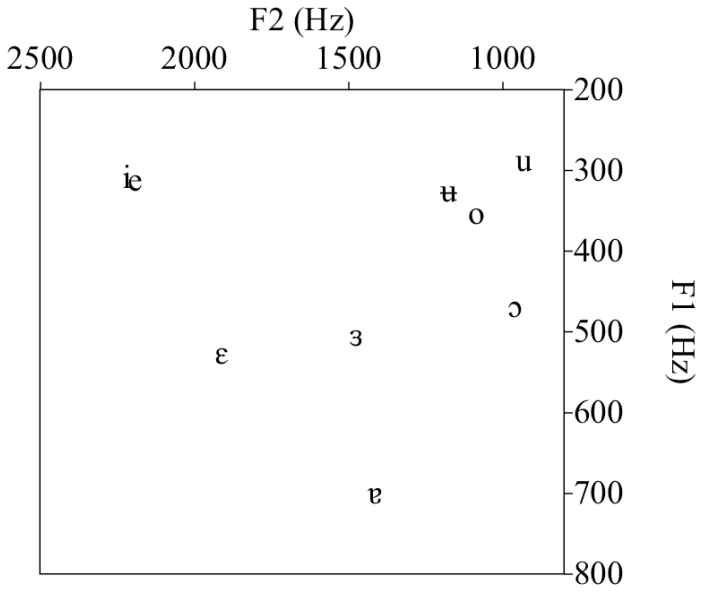
Change	**Examples**
b > w	/^{R}bjɛ/ ‘leaf’ > /^{LP}sĩ-wjɛ/ [^{LP}sĩ-ɥɛ] ‘tree leaves’
p > hw	/^{H}pʰi/ ‘to weave’ > /^{EP}mjɛ hwi/ [hɥi] ‘to weave a bamboo basket’
dz > z	/^{H}dzjɛ/ ‘to ride’ > /^{RP}rɔ̃ zjɛ/ ‘to ride a horse’
tɕʰ > ɕ	/^{R}tɕʰĩ/ ‘to drink’ > /^{EP}tɕwe ɕĩ/ ‘drink tea’
dʑ > ʑ	/^{R}dʑwɐ/ ‘to hit’ > /^{LP}hĩ ʑwɐ/ ‘to hit a man’
kʰ > x	/^{R}kʰu/ ‘to want’ > /^{EP}mu＝xu/ ‘not want’
qʰ > χ	/^{H}qʰɔ/ ‘to slaughter, kill’ > /^{RP}be χɔ/ ‘to kill a pig’

The noted changes are not only less numerous, but are also less regular than the lenition rules in Lower Xumi, and for some initials, also sporadic and marginal. This is the case for /^{R}dʑɔ̃/ ‘to have’, where the initial /dʑ/ irregularly lenites to /ʑ/, as in [^{RP}mu＝ʑɔ̃] ‘not have’.

Vowels

The Xumi vowel system comprises 9 oral vowels and 6 nasal vowels (i.e. one oral vowel more than in Lower Xumi). Vowels exhibit a wide range of surface realization (also among the language consultants that we worked with) with most vowels occurring in a rather narrow range of contexts. The number of hypothesized underlying vowels and the relation between the surface vowel realizations and the phonemic vowel categories have been established on the basis of cognates in both varieties. See also the vowel chart plotted on the relative f1/f2 formant values below. (F1 and f2 were measured at one quarter of the vowel duration (25%), at mid point (50%), and at three quarters of the vowel duration (75%); the mean formant value of a vowel was calculated by averaging over the three measurements).

i	^{R}bi	'flour'	^{H}tsi	'lock'	^{RP}lɐ-ts^{h}i	'to abstain'
e	^{R}be	'pig'	^{EP}tsezi	'to trample'	^{H}ts^{h}e	'to like'
ɛ	^{R}bjɛ	'leaf'	^{EP}ɐ-tsjɛ	'hot'	^{EP}n̥ɐts^{h}jɛ	'earring'[7]
ɜ	^{R}bɜ	'pot, pan'	^{RP}tsɐ-tsɜ	'to bundle'	^{H}ts^{h}ɜ	'salt'
u	^{R}bu	'crops'	^{H}tsu	'to pluck'	^{R}ts^{h}u	'ghost'
ʉ	^{R}bʉ	'to rub with hands'				
o	^{LP}boxu	'gourd'	^{H}tso [^{H}tsɤ]	'crown of a head'	^{H}ts^{h}o	'lungs'
ɔ	^{H}bɔ	'mask'[8]				
ɐ	^{RP}bɐzɻ̃	'cicada'	^{H}tsɐ	'sinew'	^{EP}n̥ɐts^{h}ɐ	'ink'[9]



Upper Xumi nasal vowels include /ĩ, ɛ̃, ũ, ɔ̃, ɐ̃/, as in /^{H}hĩ/ ‘man, person’, /^{H}hɛ̃/ ‘to cut, to slice (meat)’, /^{H}hũ/ ‘to blow’, /^{H}hɔ̃/ ‘vegetable’, /^{RP}lɐ-hɐ̃/ ‘have disappeared’. Marginally, Upper Xumi also has the vowel /ə̃/, which is only attested in one word in our corpus, that is /^{LP}mə̃dɐ/ ‘on the roof, upstairs’ (compare, /^{LP}mĩdɐ/ ‘pitiful’).

Vowel realization

/i/ and /e/ overlap with some initials (see for discussion below).

Front mid vowels contrast two degrees of vowel height: /e/ vs. /ɛ/, as in the following examples (where these vowels are also contrasted with /i/): /^{R}ji/ ‘to exist (of animate beings)’, /^{R}je/ ‘to lick’, /^{R}jɛ/ ‘vegetable oil’; /^{RP}ɐdʑi/ ‘alcohol, spirits’, /^{H}dʑe/ ‘happy’, /^{H}dʑɛ/ ‘rich’.

The front vowel /ɛ/ seems to be limited to occur only with [j]. For example, /ᴴbjɛ/ ‘hoof’, /ᴴdjɛ/ ‘fox’, /ᴿgjɛ/ ‘eggplant’, /ᴴmjɛ/ ‘bamboo’. The combination /jɛ/ is realized as [jæ] for some speakers, e.g. /ᴸᴾmjɛ-tsʰu/ [ᴸᴾmjæ-tsʰu] ‘downstairs’.

/ʉ/ has a restricted distribution. It only occurs after bilabials (e.g. /ᴿbʉ/ ‘to rub (with hands), to twist’), and velars, where it is realized as [ɯ] (e.g. /ᴴkʉ/ [ᴴkɯ] ‘to bake, to roast, to toast’).

/o/ has many different realizations, for it represents a change in progress, that is, raising to /u/ (see below). After alveolars, it is realized as [o] or [ɵ] (e.g. /ᴴtʰo/ ‘manner, way (to do things)’, /ᴴto/ [ᴴtɵ] ‘to build’). After alveopalatals, it is realized as [əʊ] (e.g. /ᴴtɕʰo/ [ᴴtɕʰəʊ] ‘to insert, to stick in’, /ᴴɕo/ [ᴴɕəʊ] ‘meat’). Occasionally, it can also be realized as [ɤ] (e.g. /ᴴtso/ [ᴴtsɤ] ‘crown of a head’).

/ɔ/ frequently appears in Tibetan loanwords (where it corresponds to WT *a*). For example, /ᴱᴾtʰuwɔ/ ‘hammer’ (WT *tho ba*), /ᴴbɔ/ ‘mask’ (WT *'ba*), /ᴴdʐɔ/ ‘enemy’ (WT *dgra*), /ᴿᴾlɐtɕʰɔ/ ‘instrument, utensil’ (WT *lag cha*), /ᴱᴾmɐzɔ/ ‘peacock’ (WT *rma bya*). In the native vocabulary, /ɔ/ contrasts with /ɐ/, e.g. /ᴴqɔ/ ‘mountain range’ vs. /ᴴqɐ/ ‘to move’, /ᴴʁɔ/ ‘strength’ vs. /ᴿᴾnɐʁɐ/ ‘twenty’.

Comparison with Lower Xumi

The main differences between the vowel systems of Lower and Upper Xumi can be summarized as follows. Compared to their counterparts in the vowel system of Lower Xumi, the vowels of Upper Xumi evidence a chain shift that consists of the raising of most vowels. The chain shift is possibly triggered by the addition to the system of the phoneme /ɔ/. Hence, Lower Xumi /e, o, ɐ, ɑ, ẽ, õ, ɐ̃, ɑ̃/ generally correspond to Upper Xumi /e, u, ɜ, ɔ, ĩ, ũ, ɛ̃, ɔ̃/, respectively. For the oral mid-vowels /e, o/, the change is in progress, yielding a number of intermediate forms (see above on the realization of these vowels) and resulting in near merger for some words, as discussed below.

The Lower Xumi /ʉ/ generally merges with /u/ in Upper Xumi, except after bilabial and velar initials, where it continues to be contrastive with /u/. For example, /ᴴbʉ/ ‘bowels, intestines’ vs. /ᴴbu/ ‘Pumi people’, /ᴿbʉ/ ‘to rub (with hands), to twist’ vs. /ᴿbu/ ‘crops’, /ᴴkʉ/ to bake, to roast, to toast’ vs. /ᴴku/ ‘to be able’. In all other cases, we observe merger with /u/. For example, the words ‘to count’ and ‘to wipe’, which are contrastive in Lower Xumi (/ᴴsu/ and /ᴴsʉ/, respectively), are homophonous in Upper Xumi, /ᴴsu/ ‘to count’ and /ᴴsu/ ‘to wipe’, respectively.

The clearest examples of vowel raising involve the (Lower Xumi) low vowels /ɐ, ɑ, ɐ̃, ɑ̃/, which rise into the empty mid vowel space. Thus, Lower Xumi /ɐ/ generally corresponds to Upper Xumi /ɜ/, Upper Xumi Lower Xumi /ɑ/ generally corresponds to Upper Xumi /ɔ/, Lower Xumi /ɐ̃/ generally corresponds to Upper Xumi /ɛ̃/, and Lower Xumi /ɑ̃/ generally corresponds to Upper Xumi /ɔ̃/, see examples in the table below:

Gloss	Lower Xumi	Upper Xumi
‘to speak’	/ᴴpɐ/	/ᴴpɜ/
‘to arrive’	/ᴴpɑ/	/ᴴpɔ/
‘sheep’	/ᴴɐ̃/	/ᴴɛ̃/
‘self’	/ᴴɑ̃/	/ᴴɔ̃/

The situation is more complex for the (Lower Xumi) oral mid-open vowels /e, o/, which by raising towards (Upper Xumi) /e, u/ cause the overcrowding of the high vowel space, leading in some cases to merger or near merger. The minimal pair ‘now’ and ‘cloth’ (Lower Xumi /ᴴɹi/ ‘now’, /ᴴɹe/ ‘cloth’) is a case in point. A minimal pair test investigating speakers’ intuitions about their pronunciation of these words (cf. Labov 1994: 353-354) suggests complete merger for most Upper Xumi speakers that we consulted: they judged the two words to be homophonous (compare the realization of these words by the speaker DDTR, /ᴴɹi/ ‘now’ and /ᴴɹi/ ‘cloth’). A more careful pronunciation of these words, as in word elicitation and discussion with our principal consultant, yielded a distinction, which our main consultant claimed to be that of tone. He pronounces the word for ‘now’ with a continued falling after the abrupt f0 rising in the initial part of the syllable, which gives the perception of a falling tone (i.e. [ɹi^{451}] ‘now’). Conversely, he pronounces the word for ‘cloth’ with a high-level pitch contour after the abrupt rising, which gives the perception of a high level tone (i.e. [ɹi^{455}] ‘cloth’). We note, however, that our principal language consultant was inconsistent in

his judgements, and that this type of variation in the realization of the high tone is non-contrastive in this variety (see below on tone).

Consider another example, involving the back vowels /u/ and /o/, as in the minimal pair /^{H}qʰu/ ‘year’ and /^{H}qʰo/ ‘bowl’, Lower Xumi /^{H}qʰʉ/ and /^{H}qʰo/, respectively. While our principal language consultant maintains a distinction, the difference is acoustically very minor. Video recording reveals a difference in the relative degree of lip rounding (see the video clips ‘Year’ and ‘Bowl’). Clearly, broader surveys and commutation tests (Labov 1994: 356-357) are needed to provide a complete account of this merger in progress.

Syllable Structure

The canonical syllable of Upper Xumi may contain up to two optional elements in the following two types of linear structures: (1) (C1)(C2)V, and (2) (C2)(C3)V, where C1 stands for nasal (before voiced stops and affricates), C2 can be any consonant, and C3 can only be -w-; V stands for vowel (or the syllable consonant /ɻ̩/), and parentheses indicate optional constituents. A non-phonemic glottal stop can be inserted at the left edge of a vowel-initial stressed syllable (e.g. /^{H}ɛ̃/ [^{H}ʔɛ̃] ‘sheep’, /^{H}ũ/ [^{H}ʔũ] ‘collar’), and at the left edge of a vowel-initial root (e.g. /^{RP}lɐ-ɛ̃/ [^{RP}lɐ-ʔɛ̃] ‘to be drunk’, /^{EP}mjɛ-ũ/ [^{EP}mjæ-ʔũ] ‘to swallow down’).

(1) V /^{RP}ɐ＝ʐ̩̃/ ‘do you have?’

(2) CV /^{H}dʐ̩ɔ/ ‘enemy’ (WT *dgra*), /^{H}sɜ/ ‘to know’

(3) CCV /^{R}Ndʐ̩ɔ-Ndʐ̩ɔ/ ‘to be similar’ (WT *'dra*), /^{H}swɜ/ ‘to whet (a knife)’

Xumi is phonologically monosyllabic with a strong tendency towards disyllabicity in its lexicon. Trisyllabic and quadrisyllabic words are mostly composite, e.g. /^{LP}tʰɔ̃-sĩpɔ̃/ ‘pine tree’ (WT *thang*) (from the word /^{LP}sĩpɔ̃/ ‘tree’), /^{RP}l̥ɜwu-dʑwɜ/ ‘peach’ (from the bound root /dʑwɜ/ ‘fruit’, as in the word /^{LP}dʑwɜ-dʑwɜ/ ‘fruit’), /^{LP}ɲɜmi buxu/ ‘sunflower’ (from /^{LP}ɲɜmi/ ‘sun’, /^{RP}buxu/ ‘flower’). At the same time, Upper Xumi also has a handful monomorphemic trisyllabic words (both native and loanwords), e.g. /^{LP}bɜzɻki/ ‘pillar’, /^{EP}lɔ̃mutɕʰi/ ‘elephant’ (WT *glang po che*).

Prosodic organization

Similar to Lower Xumi, Upper Xumi has a two-way tonal contrast on monosyllabic words and a three-way contrast of tonal melodies on polysyllabic words.

Tone and tonal melodies on lexical words

On monosyllabic roots, we observe a two-way tonal contrast: (1) rising (R) (e.g. /^{R}je/ ‘to lick’, /^{R}jɛ/ ‘vegetable oil’, /^{R}lɐ/ ‘suspension bridge’, /^{R}wɐ/ ‘cow’), (2) high (H) (e.g. /^{H}je/ ‘tobacco’, /^{H}jɛ/ ‘to buy’, /^{H}lɐ/ ‘tiger’, /^{H}wɐ/ ‘tooth’). The pitch contours of the two tones differ slightly from the corresponding tones in Lower Xumi, as in Figure 3:

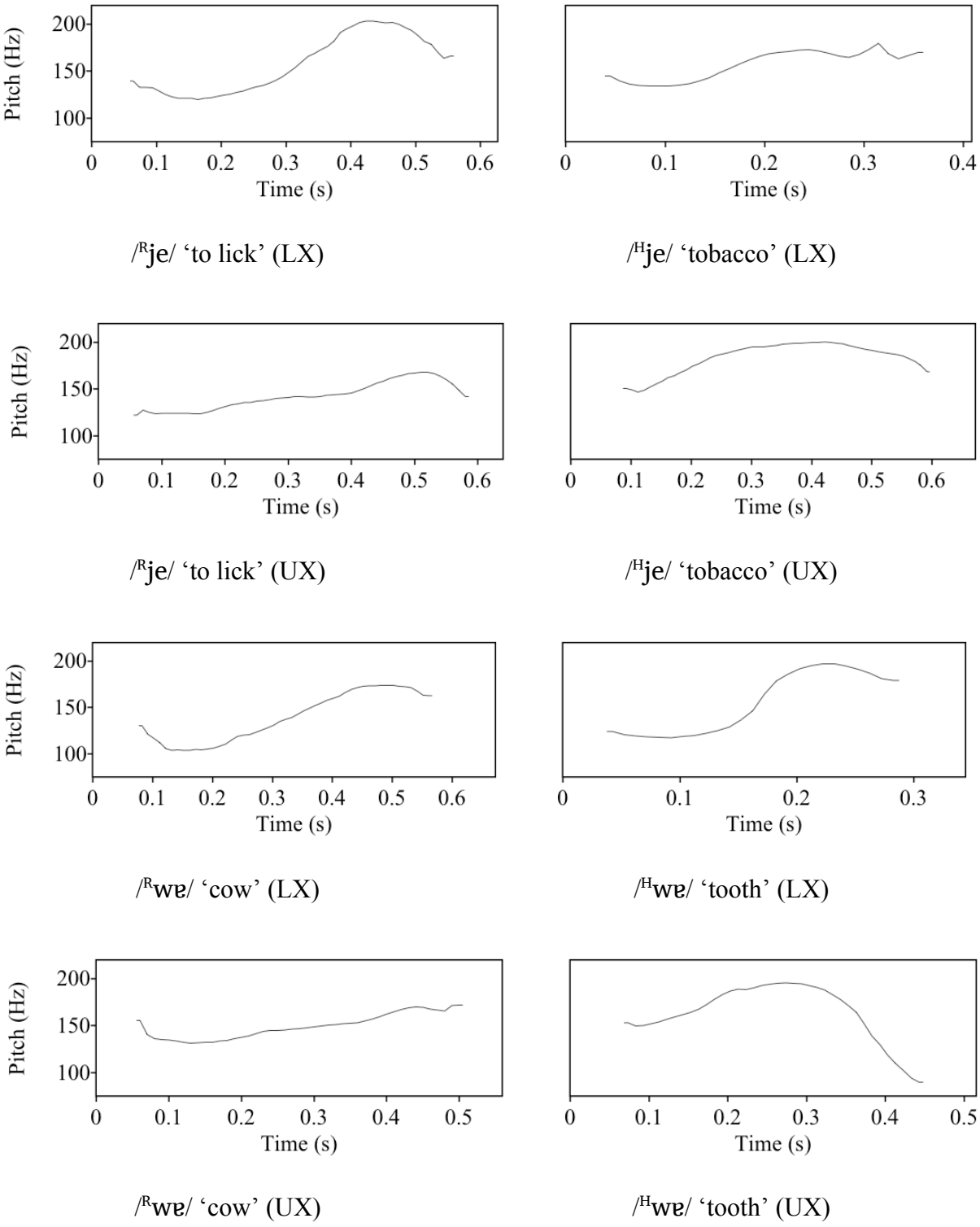


Figure 3. Pitch contours for the pairs /Rje/ ‘to lick’ vs. /Hje/ ‘tobacco’, and /Rwɐ/ ‘cow’ vs. /Hwɐ/ ‘tooth’ in Lower and Upper Xumi

Similar to Lower Xumi, non-contrastive variation abounds in the actual realization of the two lexical tones in Upper Xumi. In particular, the high tone may be realized as a falling tone. For example, the word /^{H}jɜ/ ‘tent’ may be realized as [jɜ^{455}] or as [jɜ^{451}].

In polysyllabic monomorphemic words, we observe a three-way contrast of tonal melodies. Here, again, we adopt the prosodic system developed for Lizu (Chirkova & Chen 2013) but it is important to note that just like Lower Xumi, Upper Xumi does not show a consistent correlation between duration and melody as observed in Lizu. Rather, the three patterns are more consistently signalled via their melodic difference. These tonal melodies are:

(1) Equally-Prominent Contour (EP): There is no salient rise or fall over any of the syllables. Rather, it seems to be high-level pitch contours throughout the two syllables. This pattern is mostly attested in monomorphemic words and in loanwords. For example, /^{EP}mɐmi/ ‘soldier’ (WT *dmag mi*), /^{EP}tɕɐzɐ/ ‘airplane’ (WT *lchags bya*).

(2) Left-Prominent Contour (LP): The high f0 peak is realized before the end of the first syllable, where the pitch starts to fall already and continues to fall in the second syllable. For example, /^{LP}mɐɲe/ ‘fire tongs’, /^{LP}dʑɜɕi/ ‘chopsticks’, /^{LP}bemi/ ‘axe’.

(3) Right-Prominent Contour (RP): The high f0 peak is realized over the last syllable of the word. For example, /^{RP}mɐɲi/ ‘mani pile’ (WT *ma ni*, pile of stones with the Mani Mantra of Avalokiteshvara), /^{RP}tɕɐzɐ/ ‘weeding hoe’ (WT *lcags gzar*), /^{RP}dʑɜɕi/ ‘to stand’, /^{RP}be-mi/ ‘sow’.

Tone patterns in compounds

If the tone of the leftmost monosyllabic root is high, the resulting compound has the left-prominent pattern, as illustrated with the following pair:

/^{H}sĩ/ ‘wood, tree’ + /^{H}q^{h}o/ ‘bowl’ = /^{LP}sĩ-χo/ ‘wooden bowl’

/^{H}sĩ/ ‘wood, tree’ + /^{R}k^{h}ʉ/ ‘root’ = /^{LP}sĩ k^{h}ʉ/ ‘root of a tree’

/^{H}sĩ/ ‘wood’ + /^{RP}dʑɔ̃hɔ̃/ ‘house’ = /^{LP}sĩ dʑɔ̃hɔ̃/ ‘wooden house’

Conversely, if the tone of the leftmost monosyllabic root is rising, the resulting tonal melody is right prominent. For example:

/^{R}ʂɔ̃/ ‘iron’ + /^{H}q^{h}o/ ‘bowl’ = /^{RP}ʂɔ̃-χo/ ‘iron bowl’

/^{R}ɹɔ̃/ ‘horse’ + /^{R}k^{h}wɜ/ ‘shed’ = /^{RP}ɹɔ̃ k^{h}wɜ/ ‘stable, horse shed’

/^{R}ɹɔ̃/ ‘horse’ + /^{RP}m̥jɛtsũ/ ‘tail’ = /^{RP}ɹɔ̃ m̥jɛtsũ/ ‘horse tail’

Transcription of the appended text

Our principal language consultant was recorded at the Laboratoire de Phonétique et Phonologie (LPP) of the Centre National de la Recherche Scientifique (CNRS). The original recording (made with a Digidesign 003 Rack souncard, Sound Studio 3.6 software for iMac, and an AKG C520L microphone) has been made available to the JIPA along with this analysis. In the transcription, only lexical items are marked for tone, whereas function words are not. Xumi function words and discourse particles (e.g. the genitive particle /ji/, the topic

marker /ʐɻ/ in the appended text) are never pronounced in isolation. Their surface tone realization depends on the tone of the preceding (host) lexical word (similar to tone sandhi in compounds).

North wind and the sun

Semi-narrow phonetic transcription

^{RP}dʑi-bu | ^{EP}dʑɜ-ʁu = ji ^{H}l̥ɔ ɲi | ^{LP}ɲɜmi ^{RP}ɲɜ-ku | ^{H}ɲi χɐ | ^{EP}eɹɔ ʎɛ = ji | ^{RP}tɕ^{h}elɐ li ɲɔ̃ || ^{RP}tɕ^{h}elɐ dʑɜ lɜ | ^{EP}ɹumi nɔ ^{RP}ts^{h}ɻ̩-ɻ̩ p^{h}ɜlɔ ^{H}gu ^{H}hĩ ^{H}hĩ ^{R}dʑĩ | ^{H}tɜ ^{H}pɔ ^{R}tɕ^{h}ũ li ɲɔ̃ || ^{RP}t^{h}e-ʎɛ ^{LP}t^{h}i-dzɻ̩ ^{RP}ɲɜ-ku | ^{RP}lɐ-k^{h}ɔ̃dɛ̃ se | ^{H}t^{h}i ^{R}bɜ ʎu | ^{H}pɜ li ɲɔ̃ || ^{H}ɲi = ɹɛ̃ ^{EP}ɕʁʁɐ nɔ | ^{H}hĩ ^{RP}t^{h}i = sɻ̩ | ^{RP}ts^{h}ɻ̩-ɻ̩ p^{h}ɜlɔ | ^{R}ɦɛ̃ = ɹɛ̃ ^{RP}lɐ-p^{h}ʉ dɔ̃ xi ʑu | ^{RP}t^{h}i = ʐɻ | ^{EP}eɹɔ nɔ ^{H}su ʎu ^{H}pɜ li ɲɔ̃ || ^{RP}t^{h}e-ʎɛ | ^{EP}dʑɜ-ʁu = ji ^{H}l̥ɔ = ɹɛ̃ | ^{EP}ɕɜ mu gi ^{R}bɜ se | ^{H}l̥ɔ ^{H}ɕũ ^{R}xi li ɲɔ̃ || ^{H}l̥ɔ ^{RP}sʁʁɔ ^{EP}duɜ nɔ ɲi | ^{H}hĩ t^{h}i ^{R}dʑɔ̃ se | ^{EP}p^{h}ɜlɔ ^{RP}lisɻ̩pu = ʁɔ̃ | ^{RP}sʁʁɔ ^{RP}lɐ-lu-lu li ɲɔ̃ || ^{EP}dʑɜ-ʁu = ji ^{H}l̥ɔ | ^{H}t^{h}o ^{RP}mu = ʐɔ̃ se | ^{H}tɜ ^{H}tʂ^{h}u ^{R}tɕi li ɲɔ̃ || ^{EP}dʑɐ-χɐ ʎɛ ʎɛ | ^{LP}ɲɜmi = ɹɛ̃ | ^{RP}bə-tsɐ li se | ^{RP}mjɛ-qɛ̃ ^{H}pu ^{R}xɐ li ɲɔ̃ || ^{H}hĩ t^{h}i | ^{EP}ɐ-tsjɛ ^{H}pu li se | ^{EP}p^{h}ɜlɔ | ^{EP}dʑɐ-χɔ ^{RP}dʑi-wu | ^{RP}lɐ-p^{h}ʉ ^{R}tɕi li ɲɔ̃ || ^{RP}t^{h}e-ʎɛ | ^{EP}dʑɜ-ʁu = ji ^{H}l̥ɔ = ɹɛ̃ | ^{H} mɔ ɲɔ̃ ^{H}pɜ lɜ | ^{H}pɜ mɔ hɛ̃ | ^{LP}ɲɜmi ^{H}χɐ ^{EP}eɹɔ ^{R}wɛ̃ ɲɔ̃ | ^{H}pɜ li ɲɔ̃ ||

Interlinear morphemic glossing

Abbreviations used in the gloss below follow the Leipzig Glossing Rules (LGR, http://www.eva.mpg.de/lingua/resources/glossing-rules.php). Non-standard abbreviations (those not included in the LGR) are: ANM = animate, ITSF = intensifying.

^{RP}dʑi-bu | ^{EP}dʑɜ-ʁu = ji ^{H}l̥ɔ ɲi | ^{LP}ɲɜmi ^{LP}ɲɜ-ku | ^{H}ɲi χɐ|

one-time	water-head=GEN	wind	and	sun		two-item	who	only
^{EP}eɹɔ	ʌɛ = ji |	^{RP}tɕ^{h}elɐ	li			ɲɔ̃ ||	^{RP}tɕ^{h}elɐ	
ruthless	become=GEN	discuss	NMLZ.PST			COP	discuss	
dʑɜ	lɜ |	^{EP}ɹumi	nɔ	^{RP}ts^{h}ɻ̩-ɻ̩		p^{h}ɜlɔ	^{H}gu	^{H}hĩ
EXP	come	road	on	ship-skin		coat	clothes	man
^{H}hĩ	^{R}dʑĩ |	^{H}tɜ	^{H}pɔ	^{R}tɕ^{h}ũ	li	ɲɔ̃ ||	^{RP}t^{h}e-ʌɛ |	^{LP}t^{h}i-dzɻ̩
man	one	there	arrive	come	NMLZ.PST	COP	that-become	that-DU
ɲɜ-ku |	^{RP}lɐ-k^{h}ɔ̃dɛ̃		se |	^{H}t^{h}i	^{R}bɜ	ʌu |	^{H}pɜ	li
two-item	PFV-talk.over		then	that	do	come.IMP	speak	NMLZ.PST
ɲɔ̃ ||	^{H}ɲi = ɹɛ̃	^{EP}ɕɐʁɐ	nɔ |	^{H}hĩ	^{RP}t^{h}i = sɻ̩ |	^{RP}ts^{h}ɻ̩-ɻ̩		
COP	who=AGT	first	inside	man	that=ANM.PNT	goat-skin		
p^{h}ɜlɔ |	^{R}ɦɛ̃ = ɹɛ̃	^{RP}lɐ-p^{h}ʉ	dɔ̃	^{R}xi	ʑu |	^{RP}t^{h}i = zɻ̩ |		
coat	self=AGT	PFV-take.off	can	CAUS	if	that=TOP		
^{EP}eɹɔ	nɔ	^{H}su	ʌu |	^{H}pɜ	li	ɲɔ̃ ||	^{LP}t^{h}e-ʌɛ |	
ruthless	inside	count	come.IMP	speak	NMLZ.PST	COP	that-become	
^{EP}dʑɜ-ʁu = ji	^{H}l̥ɔ = ɹɛ̃ |		^{EP}ɕɜ	mu	gi	^{R}bɜ	se |	^{H}l̥ɔ	^{H}ɕũ
water-head=GEN	wind=AGT		die	NEG	care	do	then	wind	blow
^{R}xi	li	ɲɔ̃ ||	^{H}l̥ɔ	^{RP}sɐʁɔ	^{EP}duɜ	nɔ	ɲi |		
CAUS	NMLZ.PST	COP	wind	more	big	inside	and		
^{H}hĩ	t^{h}i |	^{RP}dʑɔ̃	se |	^{EP}p^{h}ɜlɔ |	^{RP}lisɻ̩pu = ʁɔ̃ |		^{RP}sɐʁɔ |	
man	that	cold	then	coat	body=on		more	
^{RP}lɐ-lu-lu		li	ɲɔ̃||	^{EP}dʑɜ-ʁu = ji			^{H}l̥ɔ |	
PFV-wrap-wrap		NMLZ.PST	COP	water-head=GEN			wind	
^{H}t^{h}o	^{LP}mu = ʐɔ̃	se |	^{H}tɜ	^{H}tʂ^{h}u	^{R}tɕi	li	ɲɔ̃||	
manner	NEG=have	then	there	abandon	do	NMLZ.PST	COP	
^{EP}dʑɐ-χɐ	ʌɛ	ʌɛ |		^{LP}ɲɜmi = ɹɛ̃ |		^{RP}bə-tsɐ	li	
one-instant	become	become		sun=AGT		outward-jump	NMLZ.PST	
se |	^{RP}mjɛ-qɛ̃		^{H}pu	^{R}xɐ	li	ɲɔ̃ ||	^{H}hĩ	t^{h}i |

then	downward-roast		carry	go.PST		NMLZ.PST	COP	man	that
^{EP}ɐ-tsjɛ	^{H}pu	li		se |	^{EP}p^{h}ɜlɔ		^{EP}dʑɐ-χɔ	^{RP}dʑi-wu |	
ITSF-hot	carry	NMLZ.PST		then	coat		one-instant	one-return	
^{RP}lɐ-p^{h}ʉ	^{R}tɕi	li		ɲɔ̃ ||	^{RP}t^{h}e-ʎɛ |		^{EP}dʑɜ-ʁu＝ji		
PFV-take.off	do	NMLZ.PST		COP	that-become		water-head=GEN		
^{H}l̥ɔ＝ɹɛ̃ |	^{H}mɔ		ɲɔ̃	^{H}pɜ	lɜ |	^{H}pɜ	mɔ	hɛ̃ |	
wind=AGT	NEG.COP		COP	speak	come	speak	NEG.COP	have	
^{LP}ɲɜmi	^{H}χɐ	^{EP}eɹɔ		^{R}wɛ̃	ɲɔ̃ |	^{H}pɜ	li	ɲɔ̃ ||	
sun	only	ruthless		COP.EMPH	COP	speak	NMLZ.PST	COP	

Acknowledgments

We would like to thank our principal language consultants, Mr. Lurong Duoding 鲁绒多丁, Mr. Mulian Lanbu 木联兰布, and Mr. Duoding Tshi'er 阿果多丁次尔, for their patience and assistance. We are grateful to Jonathan Evans and the anonymous reviewers of this journal for helpful comments and suggestions. Thanks are also due to the directors of the Laboratoire de Phonétique et Phonologie (LPP) of the Centre National de la Recherche Scientifique (CNRS) Jacqueline Vaissière and Annie Rialland for making the phonetics laboratory of the LPP available to us through the project LabEx-EFL (Empirical Foundations of Linguistics); to Angélique Amelot (LPP, CNRS) for helping us with recordings; to Jos Pacilly (Leiden University) for helping us with Praat scripts; and to Wang Dehe 王德和 for helping us with videorecordings. We gratefully acknowledge the financial support of the Agence Nationale de la Recherche (France) as part of the research project "What defines Qiang-ness? Towards a phylogenetic assessment of the Southern Qiangic languages of Muli" (ANR-07-JCJC-0063) to Katia Chirkova and Tanja Kocjančič Antolík. Yiya Chen is currently supported by the Netherlands Organisation for Scientific Research (NWO-VIDI 016084338) and the European Research Council (ERC-Starting Independent Researcher Grant, 206198).

References

Chirkova, Katia. 2009. Shixing, a Sino-Tibetan language of South-West China: A grammatical sketch with two appended texts. *Linguistics of the Tibeto-Burman area* 32(1), 1–90.

Chirkova, Katia. 2014. Kami. *Phonological Profiles of Little-Studied Tibetic Varieties*, ed. by Jackson T.-S. Sun (Language and Linguistics Monograph Series 55), pp. 1-77. Taipei: Academia Sinica.

Chirkova, Katia & Yiya Chen. 2013. Lizu (Illustrations of the IPA). *Journal of the International Phonetic Association* 43(1): 75-86.

Chirkova, Katia & Yiya Chen. 2013. Lower Xumi. *Journal of the International Phonetic Association* 43(3): 363-379.

Huang Bufan & Renzeng Wangmu. 1991. Shixingyu [The Shixing language]. In Dai Qingxia, Huang Bufan, Fu Ailan, Renzeng Wangmu & Liu Juhuang (eds.), *Zang-Mianyu shiwu zhong / Fifteen Tibeto-Burman languages*, 174–197. Beijing: Beijing Yanshan Chubanshe.

Gesang Jumian (skal bzang 'gyur med) & Gesang Yangjing (skal bzang dbyangs chan). 2002. *Zangyu fangyan gailun / Bod kyi yul skad rnamb shad / An Introduction to Tibetan Dialects*. Beijing: Nationalities Press.

Labov, William. 1994. *Principles of linguistic change: Internal Factors*. Cambridge, MA: Blackwell Publishers.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer.

Appendix 1: C-V combinations in Upper Xumi, as based on a vocabulary list of ca. 2,800 words (marginal phonemes restricted to 2 or less occurrences, Chinese loanwords and place names (such as Lanman) are in grey)

	i	e	ɛ	ɨ	ɜ	ʉ	u	o	ɔ	a	ĩ	ɛ̃	ũ	ɔ̃	ɐ̃
Ø	✓	✓		✓			✓	✓		✓		✓	✓	✓	
p	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
pʰ	✓		✓		✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
b	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Nb		✓					✓		✓	✓					
t	✓	✓	✓		✓		✓	✓		✓	✓	✓	✓	✓	✓
tʰ	✓	✓	✓				✓	✓		✓		✓		✓	
d		✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Nd	✓						✓								
k	✓	✓	✓			✓	✓			✓		✓		✓	
kʰ	✓	✓			✓	✓	✓			✓				✓	✓
g	✓	✓	✓		✓		✓			✓	✓	✓		✓	
Ng							✓	✓		✓					
q					✓		✓	✓	✓	✓		✓	✓	✓	
qʰ		✓			✓		✓	✓	✓	✓			✓		
ts	✓	✓	✓	✓	✓		✓	✓		✓		✓	✓	✓	✓
tsʰ	✓	✓	✓	✓	✓		✓	✓		✓	✓		✓	✓	
dz		✓	✓	✓	✓		✓			✓		✓	✓	✓	
Ndz	✓	✓			✓				✓						
tʂ	✓			✓			✓			✓		✓		✓	
tʂʰ	✓	✓		✓			✓			✓		✓		✓	
dʐ	✓			✓			✓	✓	✓	✓					
Ndʐ	✓								✓	✓					
tɕ	✓	✓	✓		✓		✓	✓		✓	✓	✓	✓		
tɕʰ	✓	✓	✓		✓		✓	✓		✓	✓	✓	✓	✓	
dʑ	✓	✓	✓		✓		✓	✓		✓	✓	✓	✓	✓	
Ndʑ		✓													
m̥			✓									✓			
m	✓	✓	✓		✓		✓	✓		✓	✓		✓	✓	✓
n̥										✓				✓	✓
n							✓		✓	✓				✓	✓
	i	e	ɛ	ɨ	ɜ	ʉ	u	o	ɔ	a	ĩ	ɛ̃	ũ	ɔ̃	ɐ̃
ɲ	✓	✓	✓		✓		✓			✓				✓	

ŋ					✓		✓			✓					
s	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
z	✓	✓		✓	✓		✓			✓	✓	✓	✓	✓	
ʂ				✓			✓			✓		✓		✓	
ʐ				✓	✓		✓							✓	
ɕ	✓		✓		✓		✓	✓		✓	✓	✓	✓		
ʑ	✓	✓	✓		✓		✓			✓	✓				
x			✓			✓	✓	✓	✓	✓					
χ					✓		✓			✓					
ʁ			✓				✓	✓	✓	✓		✓	✓	✓	
h											✓	✓	✓	✓	✓
ɦ												✓		✓	
w	✓	✓			✓		✓	✓	✓	✓		✓			✓
ɹ	✓	✓			✓		✓			✓		✓	✓	✓	✓
j	✓	✓	✓		✓		✓			✓	✓	✓	✓	✓	✓
l̥	✓	✓			✓				✓	✓					
l	✓	✓			✓		✓		✓	✓				✓	✓
ʏ			✓				✓			✓					
ʏ̥			✓												

Clusters with *-w-*

	i	**e**	**ɛ**	**ɜ**	**ɐ**	**ĩ**	**ɛ̃**
tw	✓			✓			
t^hw				✓			
dw	✓	✓		✓			
kw	✓	✓		✓			
k^hw				✓			
gw	✓		✓	✓	✓		
qw				✓	✓		
q^hw				✓	✓		
tsw	✓			✓			
ts^hw	✓		✓	✓			
	i	**e**	**ɛ**	**ɜ**	**ɐ**	**ĩ**	**ɛ̃**
dzw				✓			

tʂw					✓		✓
tʂʰw				✓			✓
dʐw					✓		
tɕw	✓	✓	✓	✓			
tɕʰw	✓			✓			
dʑw	✓	✓	✓	✓			
ŋw							✓
sw	✓			✓			✓
zw				✓			
ɕw	✓		✓	✓		✓	✓
ʑw			✓	✓			
ʐw				✓			
xw	✓	✓					
hw						✓	
ʁw					✓		
ɹw				✓	✓		
ʎw	✓	✓		✓			

[1] In transcriptions of Xumi words "-" stands for morpheme boundary, and "=" stands for clitic boundary. See section 'Prosodic organisation' for the adopted system of tone notation (the superscript letters and numbers in transcriptions).

[2] In prefixes (which are prosodically unstressed), such as the directional prefix /bu-/ 'outward', /u/ is reduced to a schwa, e.g. /[RP]bə-ɕe/ 'to stick, to glue', /[RP]bə-ge/ 'to pull outward', /[RP]bə-ɲɔ̃/ 'to push (outward)'.

[3] This is a traditional unit of length, equal to the distance between the thumb and out-stretched middle finger.

[4] The combination /jɛ/ is realized as [jæ] for some speakers, see also below on vowels.

[5] Homorganic nasal-consonant onsets can be described as unit phonemes or as consonant clusters (clusters carrying secondary articulatory gestures). Our analysis of such segments as consonant clusters in Xumi follows that of similar segments in Kham Tibetan dialects (e.g. Gesang & Gesang 2002), which, in most cases, appear to be the donors of words with homorganic nasal-consonant onsets in Upper Xumi.

[6] This word is realized with a rising contour on the first syllable. This is an unusual pattern for disyllabic words, not further attested in other words in the corpus.

[7] This word is likely to be a loan from WT *rna* ? 'ear ?'.

[8] This word is a loan from WT *'ba* 'mask'.

[9] This word is a loan from WT *snag tsha* 'ink'.