This paper presents the first illustration of the vowel system of the Samagaun dialect of Nubri. We present acoustic data for the nine oral vowels and illustrate the contrastive nasalisation that may be added to three of these vowels. We provide examples of length and non-modal phonation (breathiness and creak), also found in Nubri, but assume that these are a function of tonal contrasts, and not orthogonal to the differences in pitch. We then position the Nubri vowel system in the context of a range of Tibetan languages and discuss challenges for classifying Nubri as Tibetic in the context of possible developments and etymologies of some key words.

Keywords: vowels, phonation, Sama, Nubri, Tibetan

1. Vowel systems

All languages contrast vowels. Globally, the number of contrastive underlying vowel qualities in languages varies from 1 to 17, though there are questions of analysis at both ends (Donohue et al., 2013). The smallest systems of vowel quality contrasts consist solely of a low vowel with wide allophonic variation, conditioned by the surrounding consonants. The largest systems, mostly found in Germanic languages of (northern) Europe, have 14 or more contrastive vowel qualities, though mainland Southeast Asia and south-west China is another region with languages containing large numbers of vowel quality contrasts, with some languages contrasting 9–14 vowel qualities. The distributions of languages with three or less vowels, or with 11 or more vowels, are shown in Maps 1 and 2. In Map 1 white circles represent languages with 3 or fewer contrastive vowel qualities, and we can clearly see that it is uncommon for a language of Eurasia to have so few vowels.

Map 1. Languages with very small vowel inventories

In Map 2 the circles represent languages with 11 or more contrastive vowel qualities; these languages are not common, except in parts of Eurasia.

Map 2. Languages with very large vowel inventories

There are 409 languages with less than four contrastive vowel qualities, displayed in Map 1 and 184 languages with more than ten contrastive vowel qualities, shown in Map 2. While these are large groups, the languages in these two maps represent only 9.4% of the languages in the sample. The remaining languages in the sample have between four and ten contrastive vowel qualities, ZECP). Many thanks also to our primary consultant, Lhakpa Norbu Lama, for his assistance. We would also like to thank Christine Gu for her contributions to the acoustic processing of the vowels.

* The work described in this paper was substantially supported by grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 1760020, AADO CBS-2022-002-K/RIO U-ZECP, AADO CBS-2022-001-K/RIO U-ZECP). Many thanks also to our primary consultant, Lhakpa Norbu Lama, for his assistance. We would also like to thank Christine Gu for her contributions to the acoustic processing of the vowels.
with two thirds of the languages having five, six, or seven contrastive vowel qualities, with all other totals listed in Table 1 containing less than 10% of the sample total.

Figure 1 and Table 1 describe the distribution of numbers of contrastive vowel qualities across the 6,330 languages and dialects in the global sample (Donohue et al., 2013).

Figure 1. Number of contrastive vowel qualities

These data allow us to evaluate the local distribution of vowel quality systems in Tibetan languages (see Tournadre, 2013) from the perspective of their place in Eurasia, in a region where both very large and very small contrastive vowel quality inventories are rare.

Table 1. Languages with different numbers of contrastive vowel qualities (n = 6330)

<table>
<thead>
<tr>
<th>Contrastive vowels</th>
<th>Number of languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>387</td>
</tr>
<tr>
<td>4</td>
<td>389</td>
</tr>
<tr>
<td>5</td>
<td>1,973</td>
</tr>
<tr>
<td>6</td>
<td>1,166</td>
</tr>
<tr>
<td>7</td>
<td>1,018</td>
</tr>
<tr>
<td>8</td>
<td>535</td>
</tr>
<tr>
<td>9</td>
<td>442</td>
</tr>
<tr>
<td>10</td>
<td>214</td>
</tr>
<tr>
<td>11</td>
<td>94</td>
</tr>
<tr>
<td>12</td>
<td>44</td>
</tr>
<tr>
<td>13+</td>
<td>46</td>
</tr>
</tbody>
</table>

Only two languages of Nepal are displayed on Maps 1 and 2; Bahing, in eastern Nepal, is described as contrasting 11 vowel qualities, while Kusunda, in western Nepal, can be analysed as only contrasting 3 underlying vowels (with wide, but conditioned, allophonic variation) (Donohue and 1).

Figure 2. Very large and very small vowel quality inventories attested in Nepal: Bahing (left) and Kusunda (right)

Classical Tibetan is written with five vowel symbols (DeLancey, 2003), and Hill (2010) reconstructs six contrastive vowels (i e a o u and i) for Old Tibetan, placing older varieties of Tibetan firmly in the middle of Figure 1, with an average number of contrastive vowel qualities. In this paper we examine the kinds of vowel systems in Central Tibetan languages, and place the complexity in Nubri vocalism in perspective with respect to the other Central Tibetan languages, Tibetan languages more generally, and the languages of Nepal.

2. Vowel contrasts in Tibetan

As mentioned above, Old Tibetan is analysed as being a language with six vowel contrasts, i e a o u and i. Variants of this kind of six-vowel system are found in non-Central Tibetan languages, such as the north-western and western languages of Ladakhi and Jad, and the north-eastern languages of Amdo and Themchen (Makley et al., 1999; Haller, 2004). In these languages the vowel system is reported as being (i e a o u a), a perhaps only notational variant of the Old Tibetan system. A large number of other north-western Tibetic languages have a simple five-vowel system, i e a o u; this is attested in Balti, Kargil, Purik and Spiti (Read, 1934; Zemp, 2006; Rangan, 1979; Sharma, 1992). This five-vowel system, which is also the most commonly attested vowel system around the world, is also present in Central Tibetan languages: Brokpa (from Bhutan), and Helambu Yohlmo, Lamjung Yohlmo, Jirel and Dolpo, from Nepal (Hari, 2010; Gawne, 2016; Maibaum & Strahm, 1973; Watters, 2002).

The different small vowel systems described to this point are represented in Figure 3.
The vowels of Nubri... 

Additional variation, without involving front rounded vowels, is found by adding vowel height contrasts (Hile Sherpa: Graves, 2007), contrasts in the low vowels (Sherpa: Kelly, 2004), or both (Syuwa: Höhlig & Hari, 1976). Figure 4 illustrates these vowel systems. These are not widely found in the Central Tibetan languages, though the Amdo Tibetan variety of Mabzhi has a similar vowel system (i.e., a o o u a), with height contrasts elaborated at the back, rather than the front (Samdrup & Suzuki, 2018).

The most commonly attested vowel system in the Central Tibetan languages does not involve a mid-vowel (i or a), but rather involves elaborating on the basic five-vowel system through the addition of two front rounded vowels, historically the result of back rounded vowels occurring before coronal consonants. Languages with this kind of system, which is not found (in Tibetan languages) outside Central Tibetan, include Gyalsumdo, Chócangacakha, Dzongkha, Lhomi, Mustang, Denjongkha, Humli, Tsum and Shigatse (Hildebrandt & Perry, 2011; Donohue, nd 2; Michailovsky, 1988; van Driem & Tshering, 1998; Watters, 1996, 2018; Downs, 2011; Versalainen & Versalainen, 1976; Kretschmar, 1995; Yliniemi, 2005; Wilde, 2001; Donohue & Dhakal, 2016; Haller, 2000). With the addition of a contrast between /e/ and /ø/ we are also describing the vocalic system of Mugom, Dingri, Drokpa and Kyirong (Lama & Japola, 2002; Herrmann, 1989; Kretschmar, 1986; Huber, 2005). When an additional height contrast is introduced in the back vowels as well, we are describing the vowel system of Lhasa Tibetan (e.g., Tournadre & Dorje, 2003). These systems are shown in Figure 5.

Front rounded vowels are found in approximately 12% of the languages of the world, but their distribution is not even. Map 3 shows the global distribution of languages with front rounded vowels; it is clear that these languages are most common in Eurasia, with the exception of South and Southeast Asia. Map 4 shows languages in and near Nepal with front rounded vowels. These segment types are not restricted to Tibetic languages, though they are frequent in that group along the border between Nepal and Tibet. Other languages with front rounded vowels in Map 4 include some varieties of Kham in western Nepal, and the Kiranti languages Bahing, Thulung, Kana and Dumi in eastern Nepal.

3. Vowel contrasts in Nubri

The western variety of Nubri described here is spoken in an around Samagaun, at the western end of the Nubri valley in northern Gorkha district
It will henceforth be referred to simply as ‘Nubri’. The population of Samagaun identify ethnically as Tibetan, and their language contains a large number of lexical items that are characteristic of Tibetan languages.

The vowel system involves nine distinct vowel qualities as illustrated in Figure 6.

Figure 6. Vowel quality contrasts in Nubri

<table>
<thead>
<tr>
<th>Vowel</th>
<th>i</th>
<th>y</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>ø</td>
<td>o</td>
<td></td>
</tr>
<tr>
<td>ε</td>
<td>ç</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

Sample words illustrating these vowels are given in (1).

(1) Vowels in Nubri
[i] mi ‘person’ pi ‘underrobe’
[e] be ‘hide’ me ‘fire’ je ‘right’
[ε] ne ‘barley’ e ‘father’
[a] ma ‘maize’ ja ‘upper’ la ‘God’
[s] do ‘potato’ ço ‘friend’
[o] do ‘stone’ lo ‘south’
[u] øu ‘blow’ nu ‘cry’ lu ‘peace’
[y] øy ‘hit’ ty ‘dig’
[ʊ] kʊ ‘appoint’ ço ‘Sama’ pʊ ‘Tibet’

To precisely characterise the vowel quality, an acoustic measurement was carried out. In a sample of 30 words per token, with three repetitions of each word, the first three formants were measured at the mid point of duration of the vowel using Praat software. The mean of these formant measurements are included in Table 2.

Table 2. Formant frequencies for Nubri vowels

<table>
<thead>
<tr>
<th>Vowel</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>280</td>
<td>2312</td>
<td>2994</td>
</tr>
<tr>
<td>e</td>
<td>344</td>
<td>2193</td>
<td>2841</td>
</tr>
<tr>
<td>ε</td>
<td>470</td>
<td>1980</td>
<td>2696</td>
</tr>
<tr>
<td>a</td>
<td>709</td>
<td>1298</td>
<td>2593</td>
</tr>
<tr>
<td>o</td>
<td>534</td>
<td>1000</td>
<td>2728</td>
</tr>
<tr>
<td>u</td>
<td>419</td>
<td>1020</td>
<td>2715</td>
</tr>
<tr>
<td>y</td>
<td>332</td>
<td>885</td>
<td>2611</td>
</tr>
<tr>
<td>Ø</td>
<td>264</td>
<td>1902</td>
<td>2333</td>
</tr>
<tr>
<td>324</td>
<td>1694</td>
<td>2373</td>
<td></td>
</tr>
</tbody>
</table>

These values represent the average formant frequencies taken from a sample of ~ 20 tokens per vowel. A plot of the three formants for each vowel is given below in Figure 7.

Figure 7. Average frequencies for F1, F2, F3 in Samagaun Nubri vowels.

4. The source of the front rounded vowels

The path for the development of front rounded vowels from back rounded vowels preceding a coronal coda is not unproblematic in Nubri. The development of front rounded vowels in Lhasa Tibetan, and other plateau Tibetan varieties, involved fronting of the vowel preceding a coronal consonant (in written Tibetan, -d, -n, -s or -l, but not -r) and the loss of that consonant (with the nasal surviving as nasalisation on the vowel), thus shifting the contrast from the coda of the syllable to the nucleus. These changes are illustrated in (2).

(2) Vowel fronting in Lhasa Tibetan
a. *khyod ‘2SG’ > ço
b. *sbas ‘hide’ > be
c. *dngul ‘silver’ > ny

e. There are a not insignificant number of words in Nubri for which this same pathway can plausibly be posited. The examples in (3) and (4) illustrate the fronted vowels preceding a coronal consonant; the written Tibetan equivalent is shown in brackets where known, and (plausibly) cognate. In (3) the coronal coda is present, and the vowel is fronted. In (4) it is absent.

(3) Vowel fronting in Nubri: fronted before coronal, coronal retained
a. dyn ‘seven’ (bDun)
b. dynço ‘front teeth’
c. gyn kha ‘winter’ (dGun)
d. phyndi ‘fart’ (? Phyen)
e. pyn ‘family’
f. rylan ‘zombie’ (Ro.sLangs)
g. sön ‘ride’ (bZhon)
h. tön ‘show’ (sTon)
i. søn ‘seed’ (Son)
14 / The vowels of Nubri...

j. tshøn ‘paint, dye’ (Tshon)
k. zonzon ‘young’ (gZhon)
l. qønop ‘guest’ (mGron.po)
m. høntor ‘be shocked’
n. dajøl ‘soup bowl’
o. kajøl ‘noodle bowl’

Note that the coronals in (3) are all sonorants, other than (3f). In (2), the original coronal in the cognate word is an obstruent but there is no coronal coda in modern Nubri, but one is found in the corresponding written Tibetan forms.

(4) Vowel fronting in Nubri: fronted before etymological coronal, coronal lost

a. ɹ̥y ‘avalanche’, < *(gNgas)rud
b. lømpa ‘wet’, < *roln
c. ly ‘overflow’, < *lud
d. ly ‘fertiliser’, < *lud
e. yakeø ‘praise’, < *bsNgags + brJod
f. ø ‘shine, bright’, < *’od
g. po ‘Tibet’, < *bod
h. søpa ‘fresh’, < *sos.pa
i. thy ‘connect’, < *mThud
j. thy ‘read’, < *sGrod (?)
k. thy wash’, < *’khrud (?)
l. tsøpa ‘think, guess’, < *tschod

In addition to the forms in (3) and (4) above, we also find a number of words in which vowel fronting is not attested in an environment in which it would thus be expected. These are shown in (5).

(5) Vowel fronting in Nubri: not fronted before coronal

a. khjo ‘2SG’, (Khyod)
b. gon ‘middle’
c. gonpa ‘monastery’ (dGon.pa)
d. konda ‘where (allative)’
e. yul ‘village, valley’ (Yul)
f. qul ‘snake’ (sBrul)
g. gul ‘move, shake’ (sGul)
h. kul ‘move (transitive)’
i. gølp ‘rotten’ (Tulpa)
j. yul ‘silver’ (dNgul)
k. pul ‘push’ (‘Phul)
l. tøl ‘down’
m. tulto ‘knot’ (mDud? Khru?)

n. khol ‘boil (intransitive)’ (Khol)
o. kol ‘boil (transitive)’
p. col ‘spray’
q. sol ‘cooked rice (honorific)’
r. toldjun ‘ear tag (on yak)’
s. bolbu ‘soft’ (‘Bolpo)
t. colto ‘three-year old horse’
u. kuntso ‘how’
v. lam tun ‘guide’ (Lam.sTon.pa)
w. yönpo ‘blue’ (sNgonpo)
x. jun ‘left’ (gYon)
y. bal ‘wool’ (Bal)
z. gal ‘crowbar’
aa. gjalmo ‘queen’
ab. kal ‘tree (species)’
ac. kalpen ‘scarf’
ad. kul ‘move (transitive)’

Finally, the words listed in (6) show fronted vowels, without any conditioning environment based on the written Tibetan form. For instance, the form -søø, attested in damso ‘molar, back tooth’ and dynsø ‘front teeth’, contains a front rounded vowel. The free form ‘tooth’, sa, has a back vowel, and the written Tibetan form, so, does not contain a coronal coda which would provide the conditioning environment for the vowel to be fronted.

(6) Vowel fronting in Nubri: fronted, not before etymological coronal

a. -sø ‘tooth’ < *so
b. ɹ̥ dig < *bru
c. ko ‘carve’ < *rKo
d. ky ‘steal’ < *rKu
e. ky ‘respect’ < *bKur
f. kjøp ‘protect’ < *sKyob
g. py ‘remove’ < *sPo
h. capø ‘dance’ < *zhab.bro
i. ɹ̥ ‘hit’ < *gZhu

These data imply that at least one source of the front rounded vowels in Nubri is a non-Tibetan element in the language. This, combined with the unpredictability of coronal codas conditioning fronted vowels, raises the possibility that the forms in (3) and (4), which conform to Standard Tibetan forms, might be loans.
5. Nasal vowels in Nubri

In addition to the nine contrastive vowels, Nubri also contrasts three nasal vowels, ē, ã and ō. From the instances where we have a Tibetan etymology, shown in (7), it is clear that these nasal vowels arose as the result of a VNV sequence reducing to a single nasal vowel, as in (8). To judge from the data in (7), which is an exhaustive list of attested words with nasal vowels, the nasal is most likely to be the bilabial nasal m (but cf. (7f). In (8) we see exceptions to the path of development shown in (8) (and here, too, the list is an exhaustive compilation of so-far attested VNV sequences in Nubri).

(7) Sources of nasal vowels in Nubri:
a. ā ‘mother’ < *a.ma
b. pā ‘tree (sp.)’
c. dāā ‘three-year old yak’
d. lāci ‘cucumber’
e. ẽ ‘milk’ < *o.ma
f. ẽ ‘come’ < *ong
g. bĩ ‘girl’ < *bu.mo
h. bĩ ‘container’
i. džĩ ‘hen’ < *bya.mo
j. lō ‘winnowing tray’
k. tsō ‘niece’ < *tsho.ma
l. gĩ ‘intestines’ < *rGyu.ma
m. nĩ ‘younger sister’ < *nome
n. pĩ ‘sun’ < *pīma
o. mē ‘father’s father’ < *mes.mes
p. maihē ‘buffalo’ < *ma.he
q. bē ‘sand’ < *bye.ma

(8) The development of nasal vowels

*V_aNV_b > ō

The resulting nasalized vowel in (8) typically has the quality of the first vowel, but cf. (7g).

(9) Exceptions to nasal vowel development:
a. tama ‘four-year yak’
b. tehama ‘decorative belt’
c. lama ‘monk’
d. dzama ‘pyjamas’
e. tshamo ‘night’
f. goma ‘mare’
g. soma ‘straw’
h. doma ‘ant’
i. gama ‘ant’
j. loma ‘leaf’
k. komo ‘old woman’
l. momo ‘dumpling’
m. kuma ‘thief’
n. pima ‘Sunday’ (sun: po)
o. nema ‘yarn’
p. tsema ‘game’
q. keme ‘woman’
r. pyem ‘knee’
s. rana ‘Rana village’
t. ciniŋ ‘earlier than last year’
u. ṭaŋa ‘rosary’
v. ḏoŋo ‘morning’
w. dzama ‘face’
x. nema ‘proverb’

These data show that both VNV sequences, which correspond to VNV in written Tibetan forms, as well as ō corresponding to VNV in written Tibetan forms, are attested in Nubri.

6. Other vowel modification in Tibetan

We have seen that Nubri contrasts oral and nasal vowels, though there are only three contrasts amongst the nasal vowels. Other ways in which the vowel inventory of Tibetan languages can be expanded, beyond vowel qualities and nasalisation, include length contrasts, which are found in most Central Tibetan languages, and phonation contrasts (cf. Gordon & Ladefoged, 2001), which are restricted to Khams languages of the south-east of the Tibetan range, probably acquired via contact with the non-Tibetan Qiangic and Ersuic languages of that area. Map 5 illustrates the distribution of vowel prosodies amongst the Tibetan languages.

Map 5. The distribution of vocalic prosodies in Tibetan languages.

Note: White circle: no prosodic contrasts; grey circle: length contrasts; square: nasalisation contrasts; diamond: length and nasalisation contrast; black circle: retroflex, tense or pharyngeal phonations contrast (as well as length and nasalisation)

As noted earlier, there are no phonation contrasts in Nubri, though we do find breathy voice dominating, especially on the low vowels, when they occur in either of the two low tones (Donohue
The Khams region shows complexity in its vowel qualities as well as its vowel prosodies. A great variety of vocalic systems are found in the Khams Tibetan dialects. Consider the examples in Figure 8, showing the contrastive vowel qualities in nine Khams Tibetan varieties, Dongwang (Bartee, 2007), Cone (Jacques, 2014), Yangthang (Suzuki, 2011a); gZhungwa (Suzuki, 2008), Gagatang (Suzuki, 2012a), Sangdam (Suzuki, 2012b), Rgyalthang (Hongladarom, 1996; Wang, 1996), dGudzong (Suzuki, 2011b), and Litang (Chen & Zhou, 2020).

Figure 8. Khams Tibetan vowel quality inventories (sample): A: Dongwang; B: Cone; C: Yangthang; D: gZhungwa; E: Sangdam; F: Gagatang; G: Rgyalthang; H: dGudzong; I: Litang

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
</tr>
<tr>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
</tr>
<tr>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
</tr>
<tr>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
</tr>
<tr>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
</tr>
<tr>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
</tr>
<tr>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
</tr>
<tr>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
</tr>
<tr>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
<td>æ a</td>
</tr>
</tbody>
</table>

7. Discussion and concluding remarks

This paper presents both impressionistic and quantified data for the vowel qualities in Nubri. We have discussed the vowels in the context of a range of Tibetan varieties and have shown that many of the words do not show regular correspondences, calling into question the exact etymologies of the words. We raise the possibility that many, if not most, of the Tibetan etyma in Nubri are present due to contact.

We conclude that the history of the vowels in Nubri cannot be accounted for by exclusive reference to Tibetan languages. The implication is that contact, with a non-Tibetan language, has played a significant role in the history of the language. A detailed etymological study of the development of Old Tibetan or Classical Tibetan words is needed to fully understand Nubri vowels, as well as a study of languages with which Nubri is, or has been, in contact.

References


Donohue, M. nd 1. Kusunda field notes.

Donohue, M. nd 2. Chöcangacakha field notes.


