

ROMANIZATION OF BALUCHI

BGN/PCGN 2008 System

The following is the BGN/PCGN-approved romanization system for deriving standard spellings of Baluchi geographic names. The romanization system is based on the Hunterian system of romanization, which has been used by the Surveys of India and Pakistan for romanizing Baluchi geographic names for more than one hundred years. The romanization system is compatible with all dialects of Baluchi, including Eastern Baluchi, Western Baluchi, and Southern Baluchi¹.

The BGN/PCGN system laid out below includes diacritical marks in order that the original script can be derived from the romanized form (i.e. it is reversible). For desk users requiring a diacritic-free form, these diacritics can simply be removed. In almost every case the same basic Roman-script characters are kept as are used in the Hunterian system. The BGN/PCGN forms have further been designed to harmonize with the BGN/PCGN Urdu romanization system. In rigorous romanization (i.e. including diacritics), retroflexion is marked by a sub-dot, and aspiration is marked by an apostrophe, where confusion with fricative digraphs could arise. For letters used only in Arabic loan words, the rigorous forms have further been designed to harmonize with the BGN/PCGN Persian romanization system.

1. Consonants and Special Characters

No.	Final	Medial	Initial	Inde- pendent	Unicode Value	Hunterian	BGN/PCGN	Baluchi Example	BGN/PCGN Romanization	Notes
1	ب	ڀ	پ	پ	0628	b	b	بينگ	benṭag ('cultivable patch of riverbed')	
2	پ	ڀ	پ	پ	067E	p	p	جاپان	Jāpān ('Japan')	
3	ت	ڌ	ت	ت	062A	t	t	اربتان	Arabistān ('Arabia')	

¹ Identified as Eastern Western and Southern Balochi in ISO 639-3

No.	Final	Medial	Initial	Independent	Unicode Value	Hunterian	BGN/PCGN	Baluchi Example	BGN/PCGN Romanization	Notes
4	ٹ _ Note8	ٹ _ Note8	ٹ ت	ٹ ت	0679 067C	t	ṭ	بینگ	bentag ('cultivable patch of riverbed')	
5	ث _ Note8	ث _ Note8	ث _ Note8	ث ت	062B 067F	th	ṭh			Represents [θ] in Eastern Baluchi, intervocalically and word-finally. Underbar distinguishes from aspirated [tʰ] (see digraphs table).
6	ج	ج	ج	ج	062C	j	j	بنجاء	bunjāh ('capital city')	
7	چ	چ	چ	چ	0686	ch	ch	بلوچستان	Balochistān	
8	ح	ح	ح	ح	062D	h	ḥ	حلق	ḥalq ('village')	Also seen ملک
9	خ	خ	خ	خ	062E	kh	kh			Largely identical to ح in pronunciation
10	د	د	د	د	062F	d	d	دامان	dāmān ('foothills or skirts of a mountain')	
11	ڈ _ Note8	ڈ _ Note8	ڈ _ Note8	ڈ ڊ	0688 0689	d	ḍ	ڊاڊر	Ḍaḍar	
12	ذ	ذ	ذ	ذ	0630	dh	ḏh	گمبذ	gumbudh ('tomb')	Represents [ð] in Eastern Baluchi, intervocalically and word-finally.
13	ر	ر	ر	ر	0631	r	r	چار راه	chār rāh ('crossroads')	

No.	Final	Medial	Initial	Independent	Unicode Value	Hunterian	BGN/PCGN	Baluchi Example	BGN/PCGN Romanization	Notes
14	ڙ _Note8	ڙ _Note8	ڙ _Note8	ڙ ړ	0691 0693	r	ṛ			
15	ز	ز	ز	ز	0632	z	z	بازار	bāzār ('market')	
16	ڙ	ڙ	ڙ	ڙ	0698	zh	zh			
17	س	س	س	س	0633	s	s	سبې	Sebī	
18	ش	ش	ش	ش	0634	sh	sh	ايشيا	Eshyā ('Asia')	
19	ص	ص	ص	ص	0635	s	ṣ			
20	ض	ض	ض	ض	0636	z	ẓ			
21	ط	ط	ط	ط	0637	t	ṭ	وطن	waṭan ('homeland')	Also spelled وطن
22	ظ	ظ	ظ	ظ	0638	z	ẓ			
23	ع	ع	ع	ع	0639	‘	‘	عباس	‘Abbās ('Bandar Abbas')	Usually not pronounced.
24	غ	غ	غ	غ	063A	gh	gh	غدار	ghaddār ('waterless desert')	
25	ف	ف	ف	ف	0641	f	f	فارموسا	Fārmosā ('Taiwan')	Common in Eastern Baluchi, occurs only sporadically in Western and Southern Baluchi,

No.	Final	Medial	Initial	Independent	Unicode Value	Hunterian	BGN/PCGN	Baluchi Example	BGN/PCGN Romanization	Notes
										where it is often replaced by پ
26	ق	قد	ق	ق	0642	q	q	حلق	ḥalq ('village')	Pronounced identically to ک
27	ک _Note8	ک _Note8	ک _Note8	ک ك	06A9 0643	k	k	ڊاک	Ḍāk	
28	گ	گ	گ	گ	06AF	g	g	رنگراه	rungrāh ('narrow mountain path')	
29	ل	لد	ل	ل	0644	l	l	مل	mall ('stream, irrigated area, pasture')	See note 3
30	م	مد	م	م	0645	m	m	جاندوم	jāndum ('south')	
31	ن	ند	ن	ن	0646	n	n	بندر	bandar ('port, harbour')	
32	س	_Note8	_Note8	ن	06BA	n	ñ	نیلبوین زر	Nīlboeñ Zir ('Arabian Sea')	It is undecided whether this character should form part of the Baluchi alphabet; we follow Jahani in accepting it.
33	و	ود	و	و	0648	w	w	مرو	Marw	
34	ه	هد	ه	ه	0647	h	h	هیرات	Herāt	Final <i>heh</i> following a consonant represents a short vowel. See vowel table for Romanization.

No.	Final	Medial	Initial	Independent	Unicode Value	Hunterian	BGN/PCGN	Baluchi Example	BGN/PCGN Romanization	Notes
34A	ھ	ھـ	ھـ	ه	06C1	h	h			
35	هـ	هـ	هـ ^{Note8}	هـ	06BE	h	h	See Digraphs table for romanization		‘Two-eyed <i>heh</i> ’ used to represent aspirated consonants in Eastern Baluchi.
36	ء	ء ءـ	ء	ء	0621/ 0626	<i>Not romanized</i>	,	فلپائن	Filpā’in (‘Philippines’)	
37	ی	یـ	یـ	ی	0649	y	y	یونان	Yūnān (‘Greece’)	

2. Consonantal Digraphs

No.	Final	Medial	Initial	Independent	Unicode Value	Hunterian	BGN/PCGN	Baluchi Example	BGN/PCGN Romanization	Notes
38	بھ	بھـ	بھـ	بھ	0628+06BE	bh	bh			Aspiration is only contrastive in Eastern Baluchi
39	پھ	پھـ	پھـ	پھ	067E+06BE	ph	ph			Aspiration is only contrastive in Eastern Baluchi
40	تھ	تھـ	تھـ	تھ	062A+06BE	th	th’			Aspiration is only contrastive in Eastern Baluchi. Apostrophe distinguishes from fricative /th/.
41	ٹھ	ٹھـ	ٹھـ	ٹھ	0679+06BE	th	ṭh			Aspiration is only contrastive in Eastern Baluchi
42	جھ	جھـ	جھـ	جھ	062C+06BE	jh	jh			Aspiration is only contrastive in Eastern

No.	Final	Medial	Initial	Inde- pendent	Unicode Value	Hunterian	BGN/PCGN	Baluchi Example	BGN/PCGN Romanization	Notes
										Baluchi
43	چھ	چھ	چھ	چھ	0686+06BE	chh	chh			Aspiration is only contrastive in Eastern Baluchi
44	دھ	دھ	دھ	دھ	062D+06BE	dh	dh'			Aspiration is only contrastive in Eastern Baluchi. Apostrophe distinguishes from fricative /dh/
45	ڊھ	ڊھ	ڊھ	ڊھ	0688+06BE	dh	ḍh			Aspiration is only contrastive in Eastern Baluchi
46	ڙھ	ڙھ	ڙھ	ڙھ	0631+06BE	rh	ṛh			Aspiration is only contrastive in Eastern Baluchi
47	کھ	کھ	کھ	کھ	06A9+06BE	kh	kh'			Aspiration is only contrastive in Eastern Baluchi. Apostrophe distinguishes from fricative /kh/
48	گھ	گھ	گھ	گھ	06AF+06BE	gh	gh'			Aspiration is only contrastive in Eastern Baluchi. Apostrophe distinguishes from fricative /gh/
49	لا	لا	لا	لا	0644+0627	lā	lā	ملائشیا	Malā'ishyā (<i>'Malaysia'</i>)	
50	کا	کا	کا	کا	06A9+0627	kā	kā			
51	گا	گا	گا	گا	06AF+0627	gā	gā	مرگاپ	Murgāp	
52	کل	کل	کل	کل	06A9+0644	kl	kl			
53	گل	گل	گل	گل	06AF+0644	gl	gl			

3. Vowels, Diphthongs, and Diacritical Marks

No.	Final	Medial	Initial	Independent	Unicode Value	Hunterian	BGN/PCGN	Baluchi Example	BGN/PCGN Romanization	Notes
54	یِ	یِ	ایِ	یِ	0650+0649	ī	ī			
55	ِ ¹	ِ	اِ	ِ	0650	i	i			
56	ے	یِ	ایِ	ے	06D2	e	e	سبئی	Sebī	
57	اَ	اَ	آ	اَ	0627 / 0622	ā	ā			
58	-	َ	اَ	َ	064E	a	a			
59	و	و	او	و	0648	o	o			
60	-	ُ	اُ	ُ	064F	u	u	بۇنجاہ	bunjāh ('capital city')	
61	وُ	وُ	اُو	وُ	064F+0648	ū	ū	یونان	Yūnān ('Greece')	
62	ےَ	یِ	ایِ	ےَ	064E+06D2	ay	ay			
63	وَ	وَ	اَو	وَ	064E+0648	aw	aw			
64	َ	َ	َ	َ	0652	<i>Not Romanized</i>	<i>Not Romanized</i>	مرو	Marw	
65	َ	-	-	َ	0670	a	á			
66	ّ	ّ	-	ّ	0651	<i>Double Consonant</i>	<i>Double Consonant</i>	ملّ	mall ('stream, irrigated area, pasture')	
67				ءَ	0621 + 064E	-ā	-ā			Singular definite suffix, see note 4

No.	Final	Medial	Initial	Independent	Unicode Value	Hunterian	BGN/PCGN	Baluchi Example	BGN/PCGN Romanization	Notes
68				ۛ	0621 + 0650	-ay	-ay			Possessive suffix, see note 4

4. Numerals

۰	۱	۲	۳	۴	۵	۶	۷	۸	۹
0	1	2	3	4	5	6	7	8	9

Although Perso-Arabic script is written from right to left, numerical expressions, e.g. ۱۹۶۸ → 1968, are written from left to right. A comma is inserted into longer sequences, either after thousands, millions, etc.

Notes

- Occasionally, sequences of /z/ or /s/ plus /h/ may be encountered, i.e. **z·h**, **s·h**. These may be romanized with the Unicode 'center dot' (U+00B7) separating the two letters, to distinguish them from the digraphs /zh/ and /sh/.
- The character *tā' marbūṭah* ۛ is found very rarely in Baluchi, principally in certain Arabic religious terms, e.g. زكاة zakāt ('alms'). It should be romanized **t**.
- When the letters ال are found, representing the Arabic definite article, the ل is assimilated to a following 'sun letter' د, ث, ت, ذ, ذ, ر, ز, س, ش, ص, ض, ط, ظ, ل or ن and is romanized t, th, d, dh, r, z, s, sh, ş, z, ț, z, l, n accordingly.
- In romanization, the suffixes ۛ (-ā, singular definite) and ۛ (-ay, possessive) are connected to the previous word by a hyphen, though they are usually written separately.

5. The word for 'and', written as و or ء, should be romanized as –u–, linked by hyphens to the two words it connects; e.g.,
سند و ہند → Sind-u-Hind ('The Gangetic Plain').
6. Except as specified in notes 4 and 5, word division in romanization should follow word division in the Baluchi script.
7. Note that the short vowels in the Baluchi examples are not pointed.
8. Certain initial, medial and final characters are not readily available in a Unicode-encoded font in a standalone form.
9. An inventory of letter-diacritic combinations in addition to the unmodified letters of the basic Roman script is:

‘ (U+2018)	’ (U+2019)
Á (U+00C1)	á (U+00E1)
Ā (U+0100)	ā (U+0101)
Ḑ (U+0044+0323)	ḑ (U+0064+0323)
ḐḐ (D+H+U+035F)	ḑḑ (d+h+U+035F)
Ḧ (U+1E28)	ḧ (U+1E29)
Ī (U+012A)	ī (U+012B)
Ñ (U+00D1)	ñ (U+00F1)
Ṛ (U+0052+0323)	ṛ (U+0072+0323)
Ş (U+015E)	ş (U+015F)
Ṭ (U+0054+0323)	ṭ (U+0074+0323)
Ṛ (U+0162)	ṛ (U+0163)
ṬṬ (T+H+U+035F)	ṭṭ (t+h+U+035F)
Ū (U+016A)	ū (U+016B)
Ẑ (U+005A+0327)	ẑ (U+007A+0327)
Ẓ (U+005A+0331)	ẓ (U+007A+0331)

10. The Romanization columns show only lowercase forms but, when romanizing, uppercase and lowercase Roman letters as appropriate should be used.