

ROMANIZATION OF URDU

BGN/PCGN 2007 System

The following is the approved romanization system for deriving standard spellings of Urdu geographical names for Pakistan. It was jointly adopted by BGN and PCGN at the 23rd BGN/PCGN Conference in Washington, DC, in 2007 and it is based on the Hunterian romanization system for Urdu, which has been used by the Surveys of India and Pakistan for romanizing Urdu geographical names for more than one hundred years. 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 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 Persian romanization system.

1. Consonants and Special Characters

No.	Final	Medial	Initial	Inde- pendent	Unicode Value	Hunterian	BGN/PCGN	Urdu Example	BGN/PCGN Romanization
1	ب	ب	ب	ب	0628	b	b	بوغدی	Boghdī
2	پ	پ	پ	پ	067E	p	p	پالیر	Pālīr
3	ت	ت	ت	ت	062A	t	t	بیزوت کله	Bezot Kale
4	ٹ	ٹ	ٹ	ٹ	0679	t	ṭ	عمل کوٹ	‘Amal Koṭ
5	ث	ث	ث	ث	062B	s	ṣ	شاہ نثار میلہ	Shāh Naṣār Melah
6	ج	ج	ج	ج	062C	j	j	برجو میلہ	Barjū Melah
7	چ	چ	چ	چ	0686	ch	ch	چپری	Chaprī
8	ح	ح	ح	ح	062D	h	ḥ	احمد خان کله	Aḥmad Khān Kale

No.	Final	Medial	Initial	Inde- pendent	Unicode Value	Hunterian	BGN/PCGN	Urdu Example	BGN/PCGN Romanization
9	خ	خ	خ	خ	062E	kh	kh	آکا خیل	Ākā Khel
10	د	د	د	د	062F	d	d	درانی	Durānī
11	ڌ	ڌ	ڌ	ڌ	0688	d	ḍ	ڌنگیلا	Ḍangīlā
12	ذ	ذ	ذ	ذ	0630	z	z̄	ذرائی	Ẓarānī
13	ر	ر	ر	ر	0631	r	r	برکی	Burkī
14	ڑ	ڑ	ڑ	ڑ	0691	r	ṛ	گیدڑہ	Gīḍarāh
15	ز	ز	ز	ز	0632	z	z	علی زائی	‘Alī Zā’ī
16	ژ	ژ	ژ	ژ	0698	zh	zh	ژوب	Zhob
17	س	س	س	س	0633	s	s	بساتو	Bisātū
18	ش	ش	ش	ش	0634	sh	sh	احمدی شاما	Aḥmadī Shāmā
19	ص	ص	ص	ص	0635	s	ṣ	اصالت کله	Aṣālat Kale
20	ض	ض	ض	ض	0636	z	ẓ	خضر خان	Khazār Khān
21	ط	ط	ط	ط	0637	t	ṭ	سلطان	Sulṭān
22	ظ	ظ	ظ	ظ	0638	z	ẓ	عظم سید نور کله	‘Aẓam Sayyid Nūr Kale
23	ع	ع	ع	ع	0639	‘	‘	علم شیر	‘Alam Sher
24	غ	غ	غ	غ	063A	gh	gh	بغاکی	Bghākī
25	ف	ف	ف	ف	0641	f	f	مظفر کوٹ	Muẓafar Koṭ
26	ق	ق	ق	ق	0642	q	q	حقدرہ	Ḥaqdarah
27	ک	ک	ک	ک	06A9	k	k	کچکینہ	Kachkīnah

No.	Final	Medial	Initial	Inde- pendent	Unicode Value	Hunterian	BGN/PCGN	Urdu Example	BGN/PCGN Romanization
28	گ	گ	گ	گ	06AF	g	g	باگن	Bāgan
29	ل	ل	ل	ل	0644	l	l	بلبلک	Bulbalak
30	م	م	م	م	0645	m	m	بلیامین	Bilyāmīn
31	ن	ن	ن	ن	0646	n	n	نہر	Nahr
32	س	س	-	س	06BA	n	ñ	جوکالیان	Jokālīāñ
33	و	و	و	و	0648	w	w	اروالی	Arawālī
34	ہ	ہ	ہ	ہ	0647	h	h	ہیروشاہ	Heroshāh
34A	ہ	ہ	ہ	ہ	06C1	h	h	مہردی	Mahrdī
35	ہ	ہ	-	ہ	06BE	h	h	بڑھ	Baṛh
36	ء	ئ ؤ	ء	ء	0621/ 0626	-	,	شیواؤ	Shīwā'o
37	ی	ی	ی	ی	0649	y	y	یارداکے	Yārdā Kale

2. Digraphs

No.	Final	Medial	Initial	Inde- pendent	Unicode Value	Hunterian	BGN/PCGN	Urdu Example	BGN/PCGN Romanization
38	بھ	بھ	بھ	بھ	0628+06BE	bh	bh	بھائی خان	Bhā'ī Khān
39	پھ	پھ	پھ	پھ	067E+06BE	ph	ph	پھاشک	Phāshk
40	تھ	تھ	تھ	تھ	062A+06BE	th	th	تھل	Thall
41	ٹھ	ٹھ	ٹھ	ٹھ	0679+06BE	th	ṭh	پٹھان ریا	Paṭhān Ryā

No.	Final	Medial	Initial	Inde- pendent	Unicode Value	Hunterian	BGN/PCGN	Urdu Example	BGN/PCGN Romanization
42	جھ	جھ	جھ	جھ	062C+06BE	jh	jh	جھیل	Jhīl
43	چھ	چھ	چھ	چھ	0686+06BE	chh	chh	بادشاہ چھم	Bādshāh Chhum
44	دھ	دھ	دھ	دھ	062D+06BE	dh	dh	سندھ	Sindh
45	ڈھ	ڈھ	ڈھ	ڈھ	0688+06BE	dh	ḍh	ڈھنڈ	Ḍhandḍ
46	ڑھ	ڑھ	ڑھ	ڑھ	0631+06BE	rh	ṛh	غوز گڑھی	Ghozgarḥī
47	کھ	کھ	کھ	کھ	06A9+06BE	kh	kh	دوغل گاکھر	Doghal Gākhar
48	گھ	گھ	گھ	گھ	06AF+06BE	gh	gh	خان گھڑی	Khān Gharī
49	لا	لا	لا	لا	0644+0627	lā	lā	غلامک کله	Ghulāmak Kale
50	کا	کا	کا	کا	06A9+0627	kā	kā	کاراخیل	Kārākhel
51	گا	گا	گا	گا	06AF+0627	gā	gā	خپینگا	Khapyangā
52	کل	کل	کل	کل	06A9+0644	kl	kl	گندہ کله	Gandah Kale
53	گل	گل	گل	گل	06AF+0644	gl	gl	گلونا ڈھیری	Gulonā Ḍherī

3. Vowels, Diphthongs, and Diacritical Marks

No.	Final	Medial	Initial	Inde- pendent	Unicode Value	Hunterian	BGN/PCGN	Urdu Example	BGN/PCGN Romanization
54	یِ	یِ	ایِ	یِ	0650+0649	ī	ī	خیرہ دین	Kherah Dīn
55	یِ ¹	یِ	ایِ	یِ	0650	i	i	مورپتھی	Maurpithī
56	ے	ے	ایِ	ے	06D2	e	e	درے پلاری	Dre Plārī

No.	Final	Medial	Initial	Inde- pendent	Unicode Value	Hunterian	BGN/PCGN	Urdu Example	BGN/PCGN Romanization
57	اَ	اَ	آ	اَ	0627 / 0622	ā	ā	آگرہ	Āgrah
58	-	اِ	اِ	اِ	064E	a	a	ڈنڈر	Ḍanḍar
59	و	و	او	و	0648	o	o	گیدو	Gīdo
60	-	اُ	اُ	اُ	064F	u	u	گبازانہ	Gubāzānah
61	وُ	وُ	اُو	وُ	064F+0648	ū	ū	اوشو	Ūsho
62	اِے	اِے	اِے	اِے	064E+06D2	ai	ai	حیدر علی کالے	Ḥaidar ‘Alī Kale
63	اِو	اِو	اِو	اِو	064E+0648	au	au	تودہ چینہ	Taudah Chīnah
64	اِ	اِ	اِ	اِ	0652	<i>Not Romanized</i>	<i>Not Romanized</i>	بوغدی	Boghdī
65	اِ	-	-	اِ	0670	a	á	موسیٰ خان کالے	Mūsá Khān Kale
66	اِ	اِ	-	اِ	0651	<i>Double Consonant</i>	<i>Double Consonant</i>	ملا باغ	Mullā Bāgh

Notes

1. When the vowel sign zīr (ِ) occurs word-finally in the first element of a compound, it is assumed to mark the Persian izafat morpheme, and is romanized **-e**, not **i**.
2. The source of almost all example names is the *1951 Census of Pakistan, Village List, Northwest Frontier Province, Chitral State*. Office of the Provincial Superintendant of Census, North-West Frontier Province, Peshawar.
3. No examples of aspirated dental r (rh, رھ) were found, though this phoneme is assumed to be part of the phonology of Urdu, and was therefore left out of Table 2.
4. Note that the short vowels in the Urdu examples are not pointed.
5. 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/.

6. 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)
Ḑ (U+1E10)	ḑ (U+1E11)
Ḣ (U+1E28)	ḣ (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)
Š (U+0053+0304)	š (U+0073+0304)
Ṭ (U+0054+0323)	ṭ (U+0074+0323)
Ṭ (U+0162)	ṭ (U+0163)
Ū (U+016A)	ū (U+016B)
Ẑ (U+005A+0327)	ẑ (U+007A+0327)
Ž (U+005A+0304)	ž (U+007A+0304)
Ẓ (U+005A+0331)	ẓ (U+007A+0331)

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