

ROMANIZATION SYSTEM FOR PASHTO

BGN/PCGN 1968 System, 2006 Revision

This system was developed in 1968 for the romanization of Pashto geographic names. Dari names in Afghanistan are romanized in accordance with the romanization System for Persian (BGN/PCGN 1958 System).

The Pashto alphabet uses a modified form of the Perso-Arabic script, and contains eight additional consonants not present in standard Arabic, as well as two additional vowel characters and an additional vowel point. Since both Pashto and Dari are official languages in Afghanistan, the same conversion values are used for all consonant characters which are common to the two languages. The eight additional Pashto consonant characters (shown below) are represented by Roman-script conversion values that suggest approximate pronunciations, and which, in six instances, bear sub-macrons.

Consonants: پ څ خ ډ ږ ښ ږ ښ

The two additional vowels and the one vowel point found in Pashto but not in Arabic are shown below:

Vowels: ی ی

Vowel Point: (ٚ)

The Pashto system of vowel pointing resembles that of both Arabic and Persian but contains additional features found in neither language. As in other Perso-Arabic-script writing systems, the vowel points and some of the diacritical marks are generally omitted from both manual and machine writing and where used are rarely fully applied. Consequently, a reference source may sometimes be required to ensure correct identification of the standard spellings and proper vowels. It is also necessary to ensure elimination of dialectal and idiosyncratic variations. In the interest of clarity, full vowel pointing has been applied to the examples below.

(Note: it is recommended that a font such as Scheherezade, available from www.sil.org, which includes the entire Unicode extended Arabic subrange, be used to view this system.)

Table 1: Consonant Characters

	Pashto				Romanization	Examples and Remarks	
	Final	Medial	Initial	Independent			
1.	ا			ا	<i>not romanized in word-initial position</i>	<i>See note 1</i>	
2.	ب	ب	ب	ب	b	بَغْلان	Baghlān
3.	پ	پ	پ	پ	p	پَوټکي	Powṭakay
4.	ت	ت	ت	ت	t	شیرین تگاب	Shirīn Tagāb
5.	ټ	ټ	ټ	ټ	ṭ	ټيټي	Taṅī
6.	ث	ث	ث	ث	s̄	ثاير	Sābir

	Pashto				Romanization	Examples and Remarks
	Final	Medial	Initial	Independent		
7.	ج	ج	ج	ج	j	جَلال آباد Jalālābād
8.	چ	چ	چ	چ	ch	چَارِيکار Chārīkār
9.	څ	څ	څ	څ	ts	څوکی Tsawkêy
10.	ځ	ځ	ځ	ځ	Dz	ځدران Dzadrān
11.	ښ	ښ	ښ	ښ	Dz	ښدران Dzadrān
12.	ح	ح	ح	ح	ḥ	حَضْرَتِ اِمَام Ḥazarat Imām ^{1a}
13.	خ	خ	خ	خ	kh	خوست Khowst
14.	د			د	d	سپين بولدک Spīn Bowldak
15.				د	<u>d</u>	دند و پتان Dand wa Patān
16.	ذ			ذ	z	ذوبی غر Zowbay Ghar
17.	ر			ر	r	کندهار Kandahār
18.				ر	r̥	اندر Andar ^{1a}
19.	ز			ز	z	کندز Kunduz
20.	ژ			ژ	zh	میر اسلم ژرنده Mīr Aslam Zhrandah ^{1a}
21.	ه			ه	<u>zh</u>	زیرین Zhīyarayn
22.	س	سد	سد	س	s	دسليم کلی Dê Salīm Kêlay ⁶
23.	ش	شد	شد	ش	sh	مزار شریف Mazār-i Sharīf ¹⁰
24.	ښ	ښد	ښد	ښ	<u>sh</u>	کښته کالا Kshatah Kalā
25.	ص	صد	صد	ص	ṣ	قیصار Qayṣār
26.	ض	ضد	ضد	ض	<u>z</u>	فیض آباد Fayzābād ⁵
27.	ط	ط	ط	ط	ṭ	قُطسِي سُفلی Quṭsī-yi Suflá ^{4,10}
28.	ظ	ظ	ظ	ظ	ẓ	ظاهر کالا Zāhir Kalā
29.	ع	ع	ع	ع	‘	پُل علم Pul-i ‘Alam ^{4,10}
30.	غ	غ	غ	غ	gh	غزنی Ghaznī
31.	ف	ف	ف	ف	f	مزار شریف Mazār-i Sharīf ¹⁰
32.	ق	ق	ق	ق	q	قُطسِي سُفلی Quṭsī-yi Suflá ^{4,10}
33.	ک	ک	ک	ک	k	کندهار Kandahār
34.	گ	گ	گ	گ	g	گردیز Gardēz
35.	گی	گی	گی	گی	g	گیزاب Gīzāb
36.	ل	ل	ل	ل	l	کابل Kābēl

37.	م	م	م	م	m	مَيْمَنَه	Maymanah
38.	ن	ن	ن	ن	n	خَانَآبَاد	Khānābād ⁵
39.	ن	ن	ن	ن	n	تَنِي	Tanī
40.	و			و	w	وَآخَانَ	Wākhān
41.	ه	ه	ه	ه	h	هَرَات	Hirāt
42.	ی	ی	ی	ی	y	يُوسُفِ خَبَل	Yūsuf Khēl
43.	ی	ی	ی	ی	y	پَكْتِيَا	Paktyā

Table 2: Vowel Characters

	Pashto	Romanization	Examples and Remarks	
1.	اَ	a	زَوَر	Zawar
2.	اَ	ā	هَرَات	Hirāt
3.	اِ	i	هَرَات	Hirāt
4.	اِي	ī	غَزَنِي	Ghaznī
5.	اِي	ī	گيزاب	Gīzāb
6.	اِي	ay	شكُور كَلِي	Shakūr Kēlay
7.	اِي	(ey)		<i>(deprecated)</i>
8.	اِي	ē	گَرْدېز	Gardēz
9.	او	aw	خوکی	Tsawkêy
10.	او	āw	يکاولنگ	Yakāwlang
11.	او	(ew)		<i>(deprecated)</i>
12.	اِ	ê	شلگر	Shêlgar
13.	اِي	êy	سَرکانِي	Sarkānêy
14.	اِي	êy	کوئی	kawêy, <i>you do</i> ^{4,10}
15.	اِ	u	مُرغَاب	Murghāb
16.	او	ū	بَالَا بُلُوك	Bālā Bulūk
17.	او	ow	لُوكَر	Lowgar
18.	اِي	ūy	شَاه جُوي	Shāh Jūy
19.	اِي	āy	دَايْمِير دَاد	Dāymīrdād
20.	آ	ā	جَلَال آباد	Jalālābād
21.	ء	'	خَدِير پَاتِين	Khidīr-i Pā'in
22.	اِ	<i>doubling of consonant letter</i>	خَانَشِين	Khānshīn
23.	اِي	á	قُطْسِي سُفَلِي	Qutsī-yi Suflá

NUMERALS									
٠	١	٢	٣	٤,٤	٥,٥	٦,٦	٧	٨	٩
0	1	2	3	4	5	6	7	8	9

Although Perso-Arabic script is written from right to left, numerical expressions, e.g. ١٩٦٨ → 1968, are written from left to right.

NOTES

1. Alīf (ا) should be romanized as follows:
 - a. In word-initial position, alīf should not be romanized.
 - b. In word-initial position, alīf with maddah (ā) should be romanized ā. (See the vowel table.)
 - c. In word-medial and word-final position, alīf should be romanized ā. (See the vowel table.)
 - d. In word-medial and word-final position in words of Arabic origin, alīf may serve as the bearer of hamzah, e.g. رأس - **ra's** (See note 4.)

2. The characters tsē (تس) and dzē (دز) may be romanized \widehat{ts} and \widehat{dz} when for special reasons it is desired that confusion be avoided between ت plus س and between د plus ز, respectively.

3. Occasionally the character sequences زه, زه, شه, and گه occur. They may be romanized **k•h**, **z•h**, **s•h**, and **g•h** in order to differentiate these romanizations from the digraphs **kh**, **zh**, **sh**, and **gh**, which are used to represent the characters ش, ز, خ, and غ.

4. Hamzah (ء) should be romanized as follows:
 - a. In word-initial position, it should not be romanized. In other positions, however, it should be romanized by an apostrophe, e.g. جزء → **juz'**.
 - b. Yē with hamzah (ي) should be romanized **êy**, unless it represents the compound (izāfah) morpheme, in which case it is romanized according to rule 10 below.

5. The division of words utilized in Pashto writing is followed in romanization, except that the elements **-ābād**, **-khwā**, **-shahr**, **-zay** and **ullāh** are always romanized as part of the preceding word, e.g.

رَحْمَتِ أَبَاد → Raḥmatābād and رَحْمَتِ اللَّهِ → Raḥmatullāh. Also, Persian derivational endings such as **-vand** and endings of Turkish origin such as **-lar, -lī, -lū, -i, -u, -si,** and **-su**, should be written together with the preceding word.

6. The Pashto preposition د should be romanized **dê** in agreement with its pronunciation, despite the fact that it is sometimes pointed with zīr.
7. The word الله, meaning God, should always be romanized **Allāh**, except as specified in note 5.
8. In names of Arabic origin, the **l** of the definite article **al** is assimilated before the ‘sun letters’ t, s, d, z, r, z, s, sh, ṣ, z, ṭ, ḏ, and n. In romanization, the article should be separated from the name it precedes and should not be capitalized, except at the beginning of a name, e.g. جَبَلُ السَّرَاجِ → **Jabal as Sarāj**.
9. In names of Arabic origin, a consonant character which would ordinarily be written with a *shaddah* (◌ّ) in Arabic should be romanized by writing the equivalent consonant letter twice, e.g., مُحَمَّد → Muḥammad.
10. In Pashto names which contain the Persian compound (*izafah*) morpheme, the *izafah* should be romanized as **-i** if the preceding word ends with a consonant other than yeh (ی) or silent heh (ه), and as **-yi** otherwise.
11. The character sequence خو, when followed by ې or ی or when pointed by the vowel character (◌ُ), should be romanized **khw**, although the **w** is either not pronounced, or only weakly pronounced.