

ROMANIZATION OF TAJIK

BGN/PCGN 1994 System

The BGN/PCGN system for Tajik was designed for use in romanizing names written in the Tajik Cyrillic alphabet. The Tajik Cyrillic alphabet contains six characters not present in the Russian alphabet: *ғ* , *ӣ* , *қ* , *ӯ* , *ҳ* and *ҷ* . An orthographic reform of the Tajik Cyrillic alphabet was implemented under the auspices of the Academy of Sciences of Tajikistan. This reform was promulgated in a decree of 3 September 1998 by the government of the Republic of Tajikistan. The reform abolished the characters *ц* , *щ* , *ь* and *ы* (see notes 2 through 5).

	Tajik	Tajik italics/ handwriting ¹	Tajik Unicode	Romanization	Example (Tajik)	Example (romanization)
1	А а	<i>А а</i>	0410; 0430	a	Омударё	Omudaryo
2	Б б	<i>Б б</i>	0411; 0431	b	Бохтар	Bokhtar
3	В в	<i>В в</i>	0412; 0432	v	Ворух	Vorukh
4	Г г	<i>Г г</i>	0413; 0433	g	Зигар	Zigar
5	Ғ ғ	<i>Ғ ғ</i>	0492; 0493	gh	Суғд	Sughd
6	Д д	<i>Д д</i>	0414; 0434	d	Дарбанд	Darband
7	Е е	<i>Е е</i>	0415; 0435	e	Душанбе	Dushanbe
8	Ё ё	<i>Ё ё</i>	0401; 0451	yo	Сирдарё	Sirdaryo
9	Ж ж	<i>Ж ж</i>	0416; 0436	zh	<i>used only in loan words</i>	
10	З з	<i>З з</i>	0417; 0437	z	Зафаробод	Zafarobod
11	И и	<i>И и</i>	0418; 0438	i	Ҷумҳурии Тоҷикистон	Jumhurii Tojikiston
12	Ӣ ӣ	<i>Ӣ ӣ</i>	04E2; 04E3	í	Қуллай Исмоили Сомонӣ	Qullai Ismoilii Somoní
13	Ӯ у	<i>Ӯ у</i>	0419; 0439	y	Шайдон	Shaydon
14	К к	<i>К к</i>	041A; 043A	k	Кӯлоб	Kūlob
15	Қ қ	<i>Қ қ</i>	049A; 049B	q	Қайроққум	Qayroqqum
16	Л л	<i>Л л</i>	041B; 043B	l	Хатлон	Khatlon
17	М м	<i>М м</i>	041C; 043C	m	Мурғоб	Murghob

¹ Handwritten or Italicized Cyrillic characters may differ from the printed form.

	Tajik	Tajik italics/ handwriting ¹	Tajik Unicode	Romanization	Example (Tajik)	Example (romanization)
18	Н н	<i>Н н</i>	041D; 043D	n	Гулистон	Guliston
19	О о	<i>О о</i>	041E; 043E	o	Кӯҳистони Бадахшон	Kūhistoni Badakhshon
20	П п	<i>П п</i>	041F; 043F	p	Помир	Pomir
21	Р р	<i>Р р</i>	0420; 0440	r	Фархор	Farkhor
22	С с	<i>С с</i>	0421; 0441	s	Шаҳритус	Shahritus
23	Т т	<i>Т т</i>	0422; 0442	t	Панҷакент	Panjakent
24	У у	<i>У у</i>	0423; 0443	u	Хуҷанд	Khujand
25	Ӯ ў	<i>Ӯ ў</i>	04EE; 04EF	ǔ	Қарокӯл	Qarokǔl
26	Ф ф	<i>Ф ф</i>	0424; 0444	f	Фарғона	Farghona
27	Х х	<i>Х х</i>	0425; 0445	kh	Хоруғ	Khorugh
28	Ҳ ҳ	<i>Ҳ ҳ</i>	04B2; 04B3	h	Кӯҳҳои Помир	Kūhhoi Pomir
29	Ч ч	<i>Ч ч</i>	0427; 0447	ch	Сомонҷӣ	Somonchí
30	Ҷ ҷ	<i>Ҷ ҷ</i>	04B6; 04B7	j	Панҷ	Panj
31	Ш ш	<i>Ш ш</i>	0428; 0448	sh	Шаҳристон	Shahriston
32	Ъ ъ	<i>Ъ ъ</i>	042A; 044A	'	Роштқалъа	Roshtqal'a
33	Э э	<i>Э э</i>	042D; 044D	é	Энгелс	Éngels
34	Ю ю	<i>Ю ю</i>	042E; 044E	yu	Юқори Полвон	Yuqori Polvon
35	Я я	<i>Я я</i>	042F; 044F	ya	Янгиобод	Yangiobod

NOTES

1. The character sequences *rx*, *zx*, *kx*, and *cx* may be romanized **g·h**, **z·h**, **k·h**, and **s·h** in order to differentiate those romanizations from the digraphs **gh**, **zh**, **kh**, and **sh** which are used to render the characters *ғ*, *ж*, *х*, and *ш*.
2. The obsolete character *ц*, abolished in 1998, should be romanized **s** (before a vowel and/or after a consonant within a word) or **ts** intervocalically.

3. The obsolete character щ, replaced by ш in 1998, should be romanized **sh**.
4. The obsolete character ь, abolished in 1998, should not be romanized.
5. The obsolete character ы, replaced by и in 1998, should be romanized **i**.
6. Unicode values are shown with the uppercase Cyrillic character first, followed by the lowercase character.
7. An inventory of letter-diacritic combinations, with their Unicode encoding, in addition to the unmodified letters of the basic Roman script is:

All apostrophes appearing in romanization are U+2019

Í (U+00CD)	í (U+00ED)
Ů (U+016C)	ů (U+016D)
Ě (U+0116)	ě (U+0117)

8. The Romanization column shows only lowercase forms but, when romanizing, uppercase and lowercase Roman letters as appropriate should be used, as demonstrated in the examples.