

## 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: ё, ў, Ѯ, ў, ѵ and Ѵ. 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 ѵ, ѿ and Ѵ (see notes 2 through 5).

	Tajik	Tajik italics	Tajik Unicode <sup>1</sup>	Romanization	Example (Tajik)	Example (romanization)
1.	А а	А а	0410;0430	а	Омударё	Omudaryo
2.	Б б	Б б	0411;0431	б	Бохтар	Bokhtar
3.	В в	В в	0412;0432	в	Ворух	Vorukh
4.	Г г	Г г	0413;0433	г	Зигар	Zigar
5.	Ҙ ғ	Ҙ ғ	0492;0493	gh	Суғд	Sughd
6.	ҏ д	ҏ д	0414;0434	d	Дарбанд	Darband
7.	Қ е	Қ е	0415;0435	е	Душанбе	Dushanbe
8.	҃ ё	҃ ё	0401;0451	yo	Сирдарё	Sirdaryo
9.	Җ ж	Җ ж	0416;0436	zh	<i>used only in loan words</i>	
10.	Ҕ з	Ҕ з	0417;0437	z	Зафаробод	Zafarobod
11.	ҙ и	ҙ и	0418;0438	i	Ҙумҳурии Тоҷикистон	Jumhuri Tojikiston
12.	Ҏ ў	Ҏ ў	04E2;04E3	í	Қуллаи Исломили Сомонӣ	Qullai Ismoili Somoni
13.	ҏ ў	ҏ ў	0419;0439	ӯ	Шайдон	Shaydon
14.	Ҍ к	Ҍ к	041A;043A	k	Қўлоб	Kulob
15.	҆ қ	҆ қ	049A;049B	q	Қайроққум	Qayroqqum
16.	҈ л	҈ л	041B;043B	l	Хатлон	Khatlon

<sup>1</sup> <https://www.unicode.org/charts/PDF/U0400.pdf>.

	Tajik	Tajik italics	Tajik Unicode <sup>1</sup>	Romanization	Example (Tajik)	Example (romanization)
17.	М м	<i>M m</i>	041C;043C	m	Мурғоб	Murghob
18.	Н н	<i>N n</i>	041D;043D	n	Гулистон	Guliston
19.	О о	<i>O o</i>	041E;043E	o	Кӯҳистони Бадаҳшон	Kӯhistoni Badakhshon
20.	П п	<i>P p</i>	041F;043F	p	Помир	Pomir
21.	Р р	<i>R r</i>	0420;0440	r	Фарҳор	Farkhor
22.	С с	<i>C c</i>	0421;0441	s	Шаҳритус	Shahritus
23.	Т т	<i>T t</i>	0422;0442	t	Панҷакент	Panjakent
24.	Ү ү	<i>Ү ү</i>	0423;0443	u	Хӯҷанд	Khujand
25.	Ӯ ӯ	<i>Ӯ ӯ</i>	04EE;04EF	ў	Қароқӯл	Qarokӯl
26.	Φ φ	<i>Φ φ</i>	0424;0444	f	Фарғона	Farghona
27.	Ҳ ҳ	<i>Ҳ ҳ</i>	0425;0445	kh	Хорӯғ	Khorugh
28.	Ӯ ӹ	<i>Ӯ ӹ</i>	04B2;04B3	h	Қӯҳҳои Помир	Kӯhhoi Pomir
29.	Ҷ ҷ	<i>Ҷ ҷ</i>	0427;0447	ch	Сомонҷӣ	Somonchī
30.	Ҵ ҵ	<i>Ҵ ҵ</i>	04B6;04B7	j	Панҷ	Panj
31.	Ш շ	<i>Ш շ</i>	0428;0448	sh	Шаҳристон	Shahriston
32.	Ҷ Ҷ	<i>Ҷ Ҷ</i>	042A;044A	'	Роштқалъа	Roshtqal'a
33.	Ҹ Ҹ	<i>Ҹ Ҹ</i>	042D;044D	é	Энгелс	Engels
34.	Ұ Ұ	<i>Ұ Ұ</i>	042E;044E	yu	Ұоқори Полвон	Yuqori Polvon
35.	ҵ ҵ	<i>ҵ ҵ</i>	042F;044F	ya	Янгиобод	Yangiobod

## NOTES

- The character sequences **Ҵ ҵ**, **Ҹ Ҹ**, **Ұ Ұ**, and **ҵ ҵ** 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 **Ҷ**, **ҵ**, **Ұ**, and **ҵ**.

2. The obsolete character **ң**, abolished in 1998, should be romanized **s** (before a vowel and/or after a consonant within a word) or **ts** intervocally.
3. The obsolete character **ҹ**, replaced by **ҹ** in 1998, should be romanized **sh**.
4. The obsolete character **ԡ**, abolished in 1998, should not be romanized.
5. The obsolete character **ԡ**, replaced by **и** in 1998, should be romanized **i**.
6. An inventory of letter-diacritic combinations, with their Unicode encoding, in addition to the unmodified letters of the basic Roman script is:

	<b>Latin Extended</b>	<b>Unicode</b>		<b>Latin Extended</b>	<b>Unicode</b>
1.	É	0116	5.	è	0117
2.	Í	00CD	6.	í	00ED
3.	Ü	016C	7.	ü	016D
4.	'	2019			

7. 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.