

ROMANIZATION SYSTEM FOR KURDISH

BGN/PCGN 2007

The tabulation below is applicable to the Kurdish language as a whole. It is based for the most part on the Hawar Roman alphabet used in the Library of Congress *Standard Kurdish Orthography Table*, but it also incorporates certain non-Hawar elements found in *A Kurdish-English Dictionary* (Taufiq Wahby & C J Edmonds, OUP, 1966). The tabulation covers both major varieties of the Kurdish language: Kurmanji and Sorani. *Kurmanji* is spoken principally in Turkey and in Iraq north of the Great Zab River (Muḥāfazat Dahūk (Dahūk Governorate)). It is usually written in Roman script, employing much of the Roman orthography in the table below. *Sorani* is spoken principally in Iraq south of the Great Zab river (Arbīl and As Sulaymānīyah governorates). It is usually written in Perso-Arabic script, employing much of the Perso-Arabic script orthography below.

Kurdish forms of geographical names in Turkey will usually be found in Roman script, and so no romanization process will be required. The digraph options for consonant letters 6, 15, and 20 will not be encountered for such names. In Iraq, Syria, and Iran, Kurdish will usually be encountered in Perso-Arabic script as in column 1 below, in which case it should be romanized into the corresponding Roman script form in column 2 below, here employing the digraph options for consonant letters 6, 15, and 20. Kurdish geographical names for places and features outside Turkey, found in Roman script form, should, where necessary and if possible, be tailored to fit the orthography of column 2 below and should employ the digraph options for consonant letters 6, 15, and 20.

CONSONANT CHARACTERS

	Kurdish	Romanization	Notes
1.	ء	'	see note 1
2.	ب	b	
3.	پ	p	
4.	ت	t	see note 2
5.	ج	c	
6.	چ	ch / ç	see note 3
7.	ح	h	
8.	خ	x	
9.	د	d	
10.	ر	r	
11.	ژ	ṛ	formerly written ر ژ or زر according to typeface available; may vary on older sources

	Kurdish	Romanization	Notes
12.	ز	z	see note 2
13.	ژ	j	
14.	س	s	see note 2
15.	ش	sh / ş	see note 3
16.	ص	ş	see note 2
17.	ض	ḍ	see note 2
18.	ط	ṭ	see note 2
19.	ع	‘	
20.	غ	gh / ħ	see note 3
21.	ف	f	
22.	ڤ	v	see note 4
23.	ق	q	
24.	ك	k	
25.	گ	g	
26.	ل	l	
27.	ڵ	ł	formerly written ڵ according to type available; may vary on older sources
28.	م	m	
29.	ن	n	
30.	و	w	see note 4
31.	ه	h	see note 5
32.	ي	y	

VOWELS

	Kurdish	Romanization	Notes
1.	ئه/ه	e	see note 5
2.	ئا/ا	a	
3.	ئي/ي	î	see note 6
4.	(no mark) / ئ	i	see note 6
5.	ئى/ئ	ê	
6.	ئو/و	u	

	Kurdish	Romanization	Notes
7.	ئوو / وو	û	
8.	ئۆ / وۆ	o	
9.	و̄	ö	Rare; previously written وي
10.	ۆ	ü	Only appearing in some dialects and only in old sources. Often equated to /û/ (row 7 above). Sometimes written بو

NOTES

1. In pure Kurdish words *hamza* is borne by *yā'* (ئ) and occurs only before initial vowels; it is not romanized. Medial and final *hamza* in Arabic borrowings are romanized by ' (apostrophe).
2. The letters ظ ط ض ص ذ ث do not occur in pure Kurdish words. In Arabic borrowings some writers retain these letters, others substitute ز س ز ت ز respectively. Only the letters ض ط and ص are catered for in the Library of Congress tabulation, as reflected in lines 16-18 of the above table. Words of obvious Arabic origin occurring in a Kurdish toponymic environment will be treated as Kurdish rather than Arabic.
3. The digraph options should be used for Kurdish geographical names in Iraq, Iran, and Syria. The single character options should be used for Kurdish geographical names in Turkey.
4. ڤ is used to represent “v” in foreign words. Some southern Kurdish writers use it to represent the “v” in borrowings from northern Kurdish dialects. و is pronounced as a “v” in the north and as a “w” elsewhere.
5. *Hā'* can be used as a vowel or a consonant. The initial (ه̄) and medial (ه̄) forms are used for the consonant ‘h’, while the final (ه̄) and independent (ه̄) forms are used to represent the vowel ‘e’. Therefore, when used as a consonant, the final and independent forms of *hā'* will be seen as ‘ه̄’ instead of ‘ه̄’ and ‘ه̄’, respectively. For example, مه مه (meh, “month”). When used as ‘e’, the *hā'* behaves like the letters *alif* (ا), *wāw* (و), *dāl* (د), *rā* (ر), in that it never joins to the following letter (*i.e.*, it has no medial form). Consequently, the following letter will display the initial form (unless there is only one following letter, in which case it will be written in the independent form). As with other vowels (see special rules 2 and 3), initial ‘e’ is preceded by the *kursî hamza*, yielding initial ه̄ (e.g., ه̄ ئه̄ enî “forehead”).
6. In pure Kurdish words, the vowel ی is always long î. When it represents *îzafe*, it is also romanized î and joined by means of a hyphen to its preceding word.

SPECIAL RULES

1. The conjunction ‘and’ (و) should be rendered u if the preceding word ends in a consonant, and w if the preceding word ends in a vowel. It should be separated by spaces from the preceding and following words.

2. In the Perso-Arabic orthography for Kurdish, all vowels are written, with the exception of the short i, which is expressed with a *kasrah* under the preceding consonant (ـِ). In Perso-Arabic script, the *kasrah* will rarely be written (e.g., کړن kirin “to do”). Like all Kurdish vowels, the short i will be preceded by a *kursî hamza* (ئ) if it appears at the beginning of a word (see 3. below)
3. In the Perso-Arabic orthography for Kurdish, when a vowel comes at the beginning of a word, or when a vowel directly follows another vowel, a *kursî hamza* precedes it (e.g., ئاگر agir “fire”).
4. A Kurdish word will never start with *alif* (ا). A Kurdish word may begin with a *yā’* (ي) or *wāw* (و), but only when they are used as a consonant, when they will be romanized as y and w, respectively.
5. When preceded by a consonant, *yā’* (ي) and *wāw* (و) should be romanized î and u, respectively. When preceded by a vowel (including short i, which is not written), *yā’* (ي) and *wāw* (و) should be romanized y and w, respectively.
6. The Arabic sign *shaddah* (ّ) denoting a doubled consonant is not used in Kurdish; doubled consonants, which are rare, are written twice. *Shaddah* might be used in Arabic borrowings but, as in unpointed Arabic, would generally be omitted.
7. Particles such as *le* (= at, in, on) and *be* (= to, for, by, with) should be written separately from their following word.
8. Occasionally the character sequences چ, سه and گ, occur. They may be romanized c·h, s·h, and g·h in order to differentiate those romanizations from the digraphs ch, sh, and gh.