Persian / فارسی Fārsī

romanization: UN 1967 + KNAB modifications

Persian uses the Arabic script with some extensions and is written from right to left. The characters have up to four different shapes depending on their position in a word. For the initial, medial and final forms of the principal characters in the romanization table below see the Arabic romanization.

I. Consonant characters

1	ا	-[A]	10	د	d	19	ط	ṭ	28	م	m
2	ب	b	11	ذ	ẕ	20	ظ	ẕ	29	ن	n
3	پ	p	12	ر	r	21	ع	ʿ	30	و	v
4	ت	t	13	ز	z	22	غ	gh	31	ه	h
5	ث	s̄	14	ژ	zh	23	ف	f	32	ى	y
6	ج	j	15	س	s	24	ق	q			
7	چ	ch	16	ش	sh	25	ک	k			
8	ح	ḥ	17	ص	ş	26	گ	g			
9	خ	kh	18	ض	ẕ	27	ل	l			

[A] Not romanized, but see romanizations accompanying *alef* (ا) in Table II.

II. Vowel characters and diacritical marks (ب stands for any consonant character)

1	بَ	a	6	آ	ʼā, ā[A]	11	بْ		(B)
2	بِ	e	7	بِى	ī	12	ء	ʼ	
3	بُ	o	8	بُو	ū	13	بّ		(C)
4	بَا بَا	ā	9	بَى بِى	ey	14	ـﮥى	-e[D], -ye[E]
5	بَى بٰى	á	10	بَو بُو	ow				

[A] Word-initially.
[B] Not romanized; marks absence of the vowel.
[C] Doubling of the consonant letter.
[D] After a consonant (excl. *-ah*).
[E] After a vowel (see also note 2).

Notes

1. The adjectival ending of Arabic origin ‑يه in Persian is romanized *‑īyeh*. In romanizing the definite article the same rules of assimilation of consonants are applied as in Arabic, e.g. زين الدين *Zeyn od Dīn*. At the beginning of the name the article is *al*, otherwise usually *ol*.

2. The relational suffix (*eẕāfeh*) *-e* is usually not expressed in Persian writing after a consonant, e.g. كوه مرغوب *Kūh-e Marghūb*. After final ا or و it is written with ى, e.g. پاى آب *Pā-ye Āb*. After final ى and ه it is expressed by writing *hamzeh* over the character, e.g. دهانة ممبر *Dahāneh-ye Mambar*.

3. To point Persian vowels two systems are in use that are separated by a column in the table. The first system is a Persian one while the other adheres to the Arabic tradition. In normal spelling vowel points are not used.

4. The combination خو before a vowel is romanized *khv* although *v* is not pronounced. The word denoting 'two' is romanized *dō*.

5. The Persian language in Afghanistan is called **Dari** or Farsi-Kabuli. This language which is close to Tajik, is often used by peoples other than the Pashto. Different to Persian, Dari contains a character borrowed from Pashto څ *ts*. In romanization there might also occur *i* (Persian *e*), *ō* (*ow, ū*), *u* (*o*).

6. To convert into the romanization approved by the UN in 2012: á → ā, ch → č, eh (word-finally) → e, gh → q, ḥ → h, ī → i, kh → x, s̄/ş → s, sh → š, ṭ → t, ū → u, z̄/ẕ/z̧ → z, zh → ž, ' → '.

7. For other modifications to ensure reversibility see the Arabic document.

Pronunciation

a	[æ]	h	[h], [-][B]	q	[ɣ], [ɢ][A]	y	[j]
ā, á	[ɒː]	ḥ	[h]	sh	[ʃ]	z̄, ẕ, z̧	[z]
ch	[tʃ]	ī	[i]	s̄, ş	[s]	zh	[ʒ]
e	[ɛ]	j	[dʒ]	ṭ	[t]	'	[ʔ]
gh	[ɣ], [ɢ][A]	kh	[x]	ū	[u]	'	[ʔ], [-][C]

[A] In positions other than between vowels.
[B] Not pronounced word-finally.
[C] When before a vowel, it lengthens the vowel.