Persian

Letters of the Alphabet

Initial	Medial	Final	Alone	Romanization
ا	ـا	ـا	ا	omit (see Note 1)
بـ	ـبـ	ـب	ب	b
پـ	ـپـ	ـپ	پ	p
تـ	ـتـ	ـت	ت	t
ثـ	ـثـ	ـث	ث	s̲
جـ	ـجـ	ـج	ج	j
چـ	ـچـ	ـچ	چ	ch
حـ	ـحـ	ـح	ح	ḥ
خـ	ـخـ	ـخ	خ	kh
د	ـد	ـد	د	d
ذ	ـذ	ـذ	ذ	z
ر	ـر	ـر	ر	r
ز	ـز	ـز	ز	z
ژ	ـژ	ـژ	ژ	zh
سـ	ـسـ	ـس	س	s
شـ	ـشـ	ـش	ش	sh
صـ	ـصـ	ـص	ص	ṣ
ضـ	ـضـ	ـض	ض	ẓ
طـ	ـطـ	ـط	ط	ṭ
ظـ	ـظـ	ـظ	ظ	ẓ̤
عـ	ـعـ	ـع	ع	ʻ (ayn)
غـ	ـغـ	ـغ	غ	gh
فـ	ـفـ	ـف	ف	f
قـ	ـقـ	ـق	ق	q
کـ	ـکـ	ـک	ک	k (see Note 2)
گـ	ـگـ	ـگ	گ	g (see Note 3)
لـ	ـلـ	ـل	ل	l
مـ	ـمـ	ـم	م	m
نـ	ـنـ	ـن	ن	n
و	ـو	ـو	و	v (see Note 3)
هـ	ـهـ	ـه ، ة	ه ، ة	h (see Note 4)
یـ	ـیـ	ـی	ی	y (see Note 3)

Vowels and Diphthongs (see Note 5)

◌َ	a	آ، ◌َا ā (see Note 6)	◌ِی	ī
◌ُ	u	◌َی á (see Note 7)	◌َوْ	aw
◌ِ	i	◌ُو ū	◌َی	ay

Notes

1. For the use of ا (*alif*) to support ء (*hamzah*) and آ (*maddah*) see rule 1(a). For the romanization of ء and آ see rules 4 and 5 respectively. For the use of ا to represent the long vowel romanized *ā* see the table of vowels and diphthongs, and rule 1(b).

2. Final ک and گ (often written ك and گ) may have the form ك, without the distinguishing upper stroke or strokes. The two letters are always distinguished in romanization.

3. For other values of و and ی see the table of vowels and diphthongs, and rules 2, 3, and 7.

4. ة (dotted ه) when used as an alternative to ت is romanized *t*.

5. Vowel points are not printed on Library of Congress cards.

6. See rules 1(b) and 5.

7. See rule 3(d).

RULES OF APPLICATION

Letters Which May Be Romanized in Different Ways Depending on Their Context

1. ا (*alif*) is used:

   (a) As a support for ء (*hamzah*) and آ (*maddah*). In these cases it is not represented in romanization. See rules 4 and 5.

   (b) To indicate the long vowel romanized *ā*. For the use of ا in *tanvīn* see rule 6.

   dānā   دانا

2. و is used to represent:

   (a) The consonant romanized *v*.

   varzish   ورزش
   davā   دوا
   sarv   سرو

Silent و following ـخ is retained in romanization.

khvāstan	خواستن
khvud	خود

(b) The long *u*-vowel (and short *u*-vowel in some monosyllables) is romanized *ū*.

dūr	دور
chūn	چون
tū	تو

(c) The diphthong romanized *aw*.

Firdawsī	فردوسی

When the diphthong precedes a consonantal و, the combination is romanized *avv*. See rule 7.

و may be used as a support for ء (*hamzah*); in this case it is not represented in romanization. See rule 4.

3.   ی is used to represent:

(a) The consonant romanized *y*.

yār	یار
siyāh	سیاه
pāy	پای

(b) The long vowel romanized *ī*.

Īrān	ایران
qālī	قالی

(c) The diphthong romanized *ay*.

ayvān	ایوان
ray	ری

(d) The final long vowel romanized *á*.

Muṣṭafá	مصطفیّ

For the use of ی as a mark of *iẓāfah* see rule 8(c).

ی in the medial forms ـیـ, ـیـ, without dots, may be used as a support for ء (*hamzah*); in this case ی is not represented in romanization. See rule 4 below.

Orthographic Symbols Other Than Letters and Vowel Signs

The signs listed below are frequently omitted in Persian writting and printing; their presence must then be inferred. They are represented in romanization according to the following rules:

4.   ٵ (*hamzah*)

    (a)   When initial, ٵ is not represented in romanization.

    (b)   When medial or final, ٵ is romanized ' (alif) except as noted in (c) and (d) below.

        mu'aṣṣir   مؤثّر
        khulafā'   خلفاء
        pā'īn   پائین

    (c)   When used as a mark of *iẓāfah*, ٵ is romanized -'i.

        astānah-'i dar   استانهٔ در

    (d)   When used to mark the indefinite article, ٵ is romanized 'i.

        khānah'i   خانهٔ

5.   آ (*maddah*)

    (a)   Initial آ is romanized *ā*

        āb   آب
        Kullīyat al-Ādāb   کلیة الآداب

    (b)   Medial آ, when it represents the phonetic combination '*ā*, is so romanized.

        ma'āṣir   مآثر
        Daryā'ābādī   دریاآبادی

    (c)   آ is otherwise not represented in romanization.

        gardāvarandah   گردآورنده

6.   *Tanvīn* (written ٌ, ٍ, ً, ًا), which occurs chiefly in Arabic words, is romanized *un*, *in*, *an*, and *an*, respectively.

7.   ّ (*shaddah* or *tashdīd*) is represented by doubling the letter or digraph concerned.

        khurram   خرّم
        avval   اوّل
        bachchah   بچّه
        Khayyām   خیّام

Note the exceptional case where ة is written over و and ى to represent the combination of long vowel plus consonant.

nashrīyāt   نشرّيات

qūvah   قوّه

Grammatical Structure as It Affects Romanization

8.   *Iẓāfah.* When two words are associated in the relation known as *iẓāfah*, the first (the *muẓāf*) is followed by an additional letter or syllable in romanization. This is added according to the following rules:

     (a)   When the *muẓāf* bears no special mark of *iẓāfah*, it is followed by *-i*.

         dar-i bāgh   در باغ

         qālī-i Īrān   قالى ايران

         khānah-i buzurg   خانه بزرگ

     (b)   When the *muẓāf* is marked by the addition of ة, it is followed by *-'i*.

         qālī-'i Īrān   قالئ ايران

         khānah-'i buzurg   خانة بزرگ

     (c)   When the *muẓāf* is marked by the addition of ى, it is followed by *-yi*.

         rū-yi zamīn   روى زمين

         Daryā-yi Khazar   درياى خزر

         khānah-yi buzurg   خانـهى بزرگ

     (d)   *Iẓāfah* is represented in romanization of personal names only when expressly indicated in the Persian script.

Affixes and Compounds

9.   Affixes.

     (a)   When the affix and the word with which it is connected grammatically are written separately in Persian, the two are separated in romanization by a single prime (ʹ). See also 12(b) below.

         khānahʹhā   خانـهها

         khānahʹam   خانـهام

         khānahʹī   خانـهاى

         mīʹravam   مى روم

    but   mīravam   ميروم

         bihʹgū   بـهگو

         barʹrasīhā   بررسيها

         Kāẓimʹzādah   كاظم زاده

    but   Kāẓimzādah   كاظمزاده

(b) The Arabic article *al* is separated by a hyphen, in romanization, from the word to which it is prefixed.

دار المعلمين   dār al-muʻallimīn
عبد الحسين   ʻAbd al-Ḥusayn

10. Compounds. When the elements of a compound (except a compound personal name) are written separately in Persian, they are separated in romanization by a single prime (ʹ). See also 12(b) below.

مریض خانه   marīẓʹkhānah
مریضخانه   *but*   marīẓkhānah
شاه نامه   Shāhʹnāmah
شاهنامه   *but*   Shāhnāmah

Note the treatment of compound personal names:

غلام علي *or* غلامعلي   Ghulām ʻAlī
شاه جهان *or* شاهجهان   Shāh Jahān
ابن ابي طالب *or* ابن ابيطالب   Ibn Abī Ṭālib

Orthography of Persian in Romanization

11. Capitalization.

(a) Rules for the capitalization of English are followed, except that the Arabic article *al* is lower cased in all positions.

(b) Diacritics are used with both capital and lower case letters.

12. The single prime (ʹ) is used:

(a) To separate two letters representing two distinct consonantal sounds, when the combination might otherwise be read as a digraph.

مرزها   marzʹhā

(b) To mark the use of a letter in its final form when it occurs in the middle of a word. See also rules 9(a) and 10 above.

راهها   rāhʹhā
قایم مقامي   Qāyimʹmaqāmī
بهآذین   Bihʹāẕīn

13. Foreign words in a Persian context, including Arabic words, are romanized according to the rules for Persian. For short vowels not indicated in the script, the Persian vowels nearest the original pronunciation of the word are supplied in romanization.

14. Dictionaries.

In romanizing Persian, the Library of Congress has found it necessary to consult dictionaries as an appendage to the romanization tables, primarily for the purpose of supplying vowels. For Persian, the principle dictionary consulted is:

M. Mu'īn. *Farhang-i Fārsī-i mutavassit.*

SPECIAL CHARACTERS AND CHARACTER MODIFIERS IN ROMANIZATION

Special Characters	Name	USMARC hexadecimal code
′	soft sign (prime)	A7
’	alif	AE
‘	ayn	B0

Character Modifiers	Name	USMARC hexadecimal code
́	acute	E2
̄	macron	E5
̣	dot below	F2
̤	double dot below	F3
̱	underscore	F6