

Urdu Romanization Table

1997 version

Urdu

Urdu (in Arabic Script)

Letters of the Alphabet

Initial	Medial	Final	Alone	Romanization
ا	ا	ا	ا	omit (see Note 1)
ب	ب	ب	ب	b
پ	پ	پ	پ	p
ت	ت	ت	ت	t
ٿ	ٿ	ٿ	ٿ	kh
د	د	د	د	d
ڌ	ڌ	ڌ	ڌ	dh
ز	ز	ز	ز	z
ڙ	ڙ	ڙ	ڙ	zh
س	س	س	س	s
ڦ	ڦ	ڦ	ڦ	sh
ڻ	ڻ	ڻ	ڻ	sh
ڻ	ڻ	ڻ	ڻ	z
ڻ	ڻ	ڻ	ڻ	z'
ڻ	ڻ	ڻ	ڻ	gh
ڻ	ڻ	ڻ	ڻ	f
ڻ	ڻ	ڻ	ڻ	q
ڻ	ڻ	ڻ	ڻ	k
ڻ	ڻ	ڻ	ڻ	g
ڻ	ڻ	ڻ	ڻ	l
				(ayn)

Urdu Romanization Table

1997 version

Urdu

۔	۔	۔	۔	n
۔	۔	۔	۔	n (see Note 2)
و	و	و	و	v
ہ	ہ	ہ	ہ	h
۔	۔	۔	۔	t (see Rule 10)
۔	۔ (ء، ے)	۔ (ء، ے)	۔	y (see Note 3)

Digraphs Representing Urdu Aspirates (see Note 4)

Value

bh	ٻ
ph	ٻ
th	ٿ
th	ٿ
jh	ڇ
ch	ڇ
dh	ڏ
dh	ڏ
rh	ڙ
kh	ڙ
gh	ڙ

Urdu Vowels and Diphthongs (see Note 5)

Value

a	ا
u	ਊ
i	ਊ
ا	ا
ا	ا
u	و
ي	ي
o	و
e	ي
au	ؤ
ai	ئ

Notes to the Tables

- For the use of ۱ (alif) to support ۴ (hamzah) and ۵ (maddah) see rules 1 and 2, respectively. For the romanization of ۴ by ۱ (alif) see rule 12. For other orthographic uses of ۱ see rules 3-4.
- For the distinction between ۶ and ۷ see rule 6.
- For the distinction between ۸ and ۹ see rule 11(c) and (e).
- For the form of the letter ۰ in these digraphs, see rule 9.
- Vowel points are used sparingly, and for romanization must be supplied from a dictionary.

Urdu Romanization Table

1997 version

Urdu

RULES OF APPLICATION

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

1. ا (alif), و and ی are used to support ه (hamzah); see rule 12. When so used, these letters are not represented in romanization.
2. ا (alif) is used to support م (maddah); see rule 13. When so used, it is not represented in romanization.
3. ا (alif) is used after a consonant to indicate the long vowel romanized ā.
rāj راج
karnā کرنا

In some words of Arabic origin this *alif* appears as a superscript letter over ی representing the *alif maqṣūrah*.

da‘vá

دعویٰ

The *alif* is sometimes omitted in writing. It is always represented in romanization.

‘Abdurrahmān عبد الرحمن

When the long vowel ā is initial, it is written ت. See rule 13(a).

4. ا (alif) may be used as an orthographic sign without phonetic significance. In these cases it is not represented in romanization. See rule 16.
‘ilmān علماء
5. ل appears as a superscript letter over ت, د, and ر when the latter represent the cerebral sounds romanized t̄, d̄, and r̄, respectively.
6. Regardless of pronunciation, undotted forms of the letter و are romanized u and dotted forms are romanized n.
jahān جهان
7. و is used:
 - (a) To represent the consonant sound romanized v.
dev دیو
vujūd وجود

In some words of Persian origin this consonant, though written, has ceased to be pronounced. It is retained in romanization.

khwīsh

خوبیش

Urdu Romanization Table

1997 version

Urdu

(b) To represent the long vowel romanized *ū*.

ūkh	اُخ
Urdū	اردو

(c) To represent the long vowel romanized *o*.

os	اوُس
dost	دوست

For the romanization of the conjunction *و* as *o* see rule 19.

(d) To represent the diphthong romanized *au*.

aur	اور
qaumī	قومی

(e) To support *ء* (*hamzah*). See rule 12.

For the use of *؎* (*shaddah*) with *ء* see rule 14.

8. ◦ is used to represent the consonantal sound romanized *h*.

ham	ہم
gāh	گاہ

Final ◦, though often not pronounced, is normally retained in romanization.

kih	کہ
guldastah	گلستانہ

Exception is made in the case of words whose final syllable ends in an aspirated consonant. When final ◦ is added to the letter ◦ in this position, it is not represented in romanization.

mukh	مکھ
------	-----

9. ◦ (usually written in the form ◦) is used to represent the aspirated element of the sounds romanized *bh*, *ph*, *th*, *t̤h*, *jh*, *ch*, *dh*, *q̤h*, *r̤h*, *kh*, *gh*.

phūl	پھول
acchā	اچھا

For the writing and romanization of words ending in an aspirated consonant, see rule 8.

10. ◦ and ◦, which are sometimes used interchangeably, are both romanized *t*.

hikmat	حکمت
--------	------

11. ◦ (often written ◦) is used:

(a) To represent the consonant romanized *y*.

siyāsat	سیاست
dayā	دیا

Urdu Romanization Table

1997 version

Urdu

(b) To represent the long vowel romanized ī.

taşyīr	تصویر
īshvār	ایشوار

(c) To represent the long vowel romanized ē.

sher	شیر
nevā	نیوا

When ی with this value is final, the form ے generally replaces ی.

se	سے
laṛke	لڑکے

(d) To represent the long vowel romanized ā. See rule 3.

da'vā	دعوی
'uqbā	عقبہ

(e) To represent the diphthong romanized ai.

maidān	میدان
bail	بیل

When ی with this value is final, it is sometimes written ے.

hai	ے
-----	---

(f) To support ے (hamzah). In this position ی is usually undotted. See rule 12.

For the use of ۽ (shaddah) with ے see rule 14.

For the use of ے in a *muzāf* see rule 17.

Romanization of Orthographic Symbols Other Than Letters and Vowel Signs

Although vowel signs are frequently omitted in printed texts, they are always taken into consideration in romanization. The rules for other symbols vary.

12. ے (hamzah)

(a) In initial position ے is not represented in romanization.

(b) In medial and final position, when ے represents a consonant, it is romanized ' (alif).

mu'min	مؤمن
li'e	لنے
bhā'ī	بھائی

Urdu Romanization Table 1997 version

Urdu

(c) When σ represents the connective syllable joining a *muzāf* to what follows, it is romanized *-yi*. See rule 17.

ملکہ انگلستان malikah-vi Inglistān

13. $\bar{\square}$ (*maddah*)

(a) At the beginning of a word, or following the Arabic article **الـ**, **أـ** is romanized **ā**.

(b) At the beginning of a syllable within a word, \mathfrak{T} is romanized 'ā.

mir'āt	مِرَّاتٍ
Our'ān	وَرَآنَ

(c) \square is otherwise omitted in romanization.

14. ☐ (*shaddah* or *tashdīd*) indicates the doubling in pronunciation of the letter over which it is written. It is represented in romanization by doubling the letter or digraph concerned.

caccā
khaṭṭā
makkhī

When **ڻ** occurs over **ڙ** and **ڙ**, these letters are regarded as representing consonants. They are romanized *vv* and *yy*, respectively.

quvvat	قوّت
sayyid	سید
Zakariyā	زکریا

15. ◊ (*sukūn* or *jazm*) indicates the absence of a vowel following the letter over which it is written. It is not represented in romanization.

16. *Tanvin* (written ئ, ة, ئـ (ئـ)) is romanized *un*, *in*, *an*, respectively, when it occurs in a word or expression borrowed from Arabic. Otherwise it is not represented in romanization.

fauran فوراً

Romanization as Affected by Grammatical Structure

17. *Izāfat,*

(a) When the *muzāf*, the first of two words in the grammatical relationship known as *iżāfah*, ends in a consonant, *-i* is added to it in the romanization.

تاریخ ہندوستان

Urdu Romanization Table

1997 version

Urdu

(b) When the *muzāf* ends in a vowel or in silent *o*, *-yi* is added.

daryā-yi shor	دریائی شور
zabān-i Urdū-yi mu'allā	زبان اردو معلی
malikah-yi Inglistān	ملکہ انگلستان

For the use of *ء* (*hamzah*) to indicate the *muzāf*, see rule 12(c).

18. The Arabic article *ال* is romanized differently depending on the letters and context with which it is associated.

(a) When it is prefixed to a word beginning with a "moon letter" (ء، و، ه، ي، ا، ب، ج، ح) it is romanized *al*.

al-Qur'ān	القرآن
-----------	--------

(b) When it is prefixed to a word beginning with a "sun letter" (ن، ل، ت، ث، د، ذ، ر) the *l* of the article is replaced in romanization by the same letter, or digraph, as that which begins the following word.

as-sijill	السجل
-----------	-------

(c) When it occurs before the second element in a name, the vowel of the article is replaced by the final vowel of the preceding word.

'Abdul'azīz	عبد العزیز
'Abdurrashīd	عبد الرشید
Abūlfażl	ابو الفضل
Zūlqarnain	ذو القرنين
Fażlullāh	فضل الله

19. The conjunction *و*, when used to join two closely associated members of a phrase, is romanized *o*.

māl o asbāb	مال و اسباب
-------------	-------------

Otherwise *و* is romanized *va*.

20. Rules for the capitalization of English are followed, except that the Arabic article *al* is lowercased in all positions.

21. The macron is used with both capital and lowercase letters.

22. The hyphen is used:

(a) To connect a *muzāf* with the following vowel or syllable. See rule 17.

(b) To connect the Arabic article *al* with the following word. See rule 18.

Urdu Romanization Table

1997 version

Urdu

23. Foreign words in an Urdu context, including Arabic and Persian words, are romanized according to the rules for Urdu.

جناب ہیڈ ماسٹر صاحب گورنمنٹ ہائی اسکول
Janāb-i Hed Māstār ṣāḥib-i Gavarmānṭ Hā'i Iskūl

For short vowels not indicated in the script, the Urdu vowels nearest the original pronunciation of the word concerned are supplied in romanization.

24. A quotation in another language using the Arabic script is romanized according to the rules for the language concerned.

SPECIAL CHARACTERS AND CHARACTER MODIFIERS IN ROMANIZATION

<i>Special Characters</i>	<i>Name</i>	<i>USMARC hexadecimal code</i>
,	alif	AE
‘	ayn	B0

<i>Character Modifiers</i>	<i>Name</i>	<i>USMARC hexadecimal code</i>
ˊ	acute	E2
ˉ	macron	E5
˘	dot below	F2
˘˘	double dot below	F3
˘˘˘	underscore	F6