

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
ٹ	ٹ	ٹ	ٹ	ṭ
ث	ث	ث	ث	<u>s</u>
ج	ج	ج	ج	j
چ	چ	چ	چ	c
ح	ح	ح	ح	ḥ
خ	خ	خ	خ	<u>kh</u>
د	د	د	د	d
ڈ	ڈ	ڈ	ڈ	ḍ
ذ	ذ	ذ	ذ	<u>z</u>
ر	ر	ر	ر	r
ڑ	ڑ	ڑ	ڑ	ṛ
ز	ز	ز	ز	z
ژ	ژ	ژ	ژ	zh
س	س	س	س	s
ش	ش	ش	ش	sh
ص	ص	ص	ص	ṣ
ض	ض	ض	ض	ẓ
ط	ط	ط	ط	ṭ
ظ	ظ	ظ	ظ	ẓ
ع	ع	ع	ع	' (ayn)
غ	غ	غ	غ	<u>gh</u>
ف	ف	ف	ف	f
ق	ق	ق	ق	q
ک	ک	ک	ک	k
گ	گ	گ	گ	g
ل	ل	ل	ل	l
م	م	م	م	m
ن	ن	ن	ن	n
و	و	و	و	<u>n</u> (see Note 2)
و	و	و	و	v

Initial	Medial	Final	Alone	Romanization
ھ	ھ	ا	و	h
-	-	آ	و	t (see Rule 10)
ی	ی (ے، ی)	ی (ے، ی)	ی	y (see Note 3)

Digraphs Representing Urdu Aspirates (see Note 4) Value

bh	بھ
ph	پھ
th	تھ
ṭh	ٹھ
jh	جھ
ch	چھ
dh	دھ
ḍh	ڈھ
ṛh	ڑھ
kh	کھ
gh	گھ

Urdu Vowels and Diphthongs (see Note 5) Value

a	ا
u	و
i	ی
ā	آ
á	آ، اِ
ū	و
ī	ی
o	و
e	ے، ی
au	و
ai	ے

Notes

1. For the use of | (*aliif*) to support ʾ (*hamzah*) and ̣ (*maddah*), see rules 1 and 2, respectively. For the romanization of ʾ by (*aliif*), see rule 12. For other orthographic uses of | see rules 3-4.
2. For the distinction between ۛ and ۞, see rule 6.

3. For the distinction between ع and ح, see rule 11(c) and (e).
4. For the form of the letter و in these digraphs, see rule 9.
5. Vowel points are used sparingly, and for romanization must be supplied from a dictionary.

RULES OF APPLICATION

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

1. | (*alif*), و and ع are used to support ء (*hamzah*); see rule 12. When so used, these letters are not represented in romanization.
2. | (*alif*) is used to support م (*maddah*); see rule 13. When so used, it is not represented in romanization.
3. | (*alif*) is used after a consonant to indicate the long vowel romanized *ā*.

rāj	راج
karnā	كرنا

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

da'vá	دعوى
-------	------

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

'Abdurrahmān	عبد الرحمن ، عبد الرحمان
--------------	--------------------------

4. | (*alif*) may be used as an orthographic sign without phonetic significance. In these cases it is not represented in romanization. See rule 16.

'amlan	عملاً
--------	-------

5. ط appears as a superscript letter over ت, د, and ر when the latter represent the cerebral sounds romanized *ṭ*, *ḍ*, and *ṛ*, respectively.
6. Regardless of pronunciation, undotted forms of the letter ن are romanized *n* and dotted forms are romanized *n*.

jahān	جهان
-------	------

7. و is used:

- (a) To represent the consonant sound romanized *v*.

dev	ديو
vujūd	وجود

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

<u>kh</u> vīsh	خویش
----------------	------

- (b) To represent the long vowel romanized *ū*.
Urdū اردو
- (c) To represent the long vowel romanized *o*.
os اوس
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 و see rule 14.

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

ham هم
gāh گاه

Final ɔ, though not pronounced, is normally retained in romanization.

kih که
guldastah گلدسته

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

mukh مکھ

9. ɔ (usually written in the form ھ) is used to represent the aspirated element of the sounds romanized *bh, ph, th, ṭh, jh, ch, ḍh, ṛh, kh, gh*.

phūl پھول
acchā اچھا

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

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

ḥikmat حکمت ، حکمت

11. ی is used:

- (a) To represent the consonant romanized *y*.
siyāsat سیاست
diyā دیا
- (b) To represent the long vowel romanized *ī*.
taṣvīr تصویر

- (c) To represent the long vowel romanized *e*.

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 *á*. See rule 3.

da'vá	دعوی
'uqbá	عقبی

- (e) To represent the diphthong romanized *ai*.

maidān	میدان
bail	بیل

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

hai	ھے
-----	----

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

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

For the use of ی 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. ء (*hamzah*)

- (a) In initial position ء is not represented in romanization.
(b) In medial and final position, when ء represents a consonant, it is romanized ' (*aliḥ*).

mu'min	مؤمن
li'e	لئے
bhāṭ	بھائی

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

malikah-yi Inḡlistān	ملکہ انگلستان
----------------------	---------------

13. آ (*maddah*)

- (a) At the beginning of a word, or following the Arabic article ال, آ is romanized *ā*.

āb	آب
----	----

(b) At the beginning of a syllable within a word, $\bar{ā}$ is romanized 'ā.

mir'āt	مرآت
Qur'ān	قرآن

(c) $\bar{ā}$ is otherwise omitted in romanization.

14. $\overset{\circ}{\text{c}}$ (*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	چچا
khattā	کھٹا
makkhī	مکھی

When $\overset{\circ}{\text{c}}$ occurs over ق and س , these letters are regarded as representing consonants. They are romanized *qv* and *yy*, respectively.

quvvat	قوت
sayyid	سید
Zakariyyā	زکریا

15. $\overset{\circ}{\text{c}}$ (*sukūn* or *jazm*) indicates the absence of a vowel following the letter over which it is written. It is not represented in romanization.

16. *Tanvīn* (written $\overset{\circ}{\text{c}}$, $\overset{\circ}{\text{c}}$, $\overset{\circ}{\text{c}}$ ($\bar{\text{ā}}$)) 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
'amlan	عملاً

Romanization as Affected by Grammatical Structure

17. *izāfat*.

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

tārīk _h -i Hindūstān	تاریخ هندوستان
---------------------------------	----------------

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

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

For the use of ء (*hamzah*) to indicate the *muḏāf*, 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”

(خ، ع، غ، ف، ق، ك، م، و، ه، ي، ا، ب، ج، ح) it is romanized *a/*.

al-Qur’ān

القرآن

(b) When it is prefixed to a word beginning with a “sun letter”

(ز، س، ش، ص، ض، ط، ظ، ل، ن، ت، ث، د، ذ، ر) the / 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 العزيز

‘Abdurrashīd

عبد الرشيد

Abūlfazl

ابو الفضل

Zūlqarnain

ذو القرنين

Faḍlullāh

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 *a/* 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 *muḥāf* with the following vowel or syllable. See rule 17.

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

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 Heḍ Māṣṭar ṣāḥib-i Gavarnmanṭ Hāī 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.