

Urdu

(in Arabic Script)

Letters of the Alphabet

Initial	Medial	Final	Alone	Romanization
ا	ل	ل	ا	omit (see Note 1)
ب	ب	ب	ب	b
پ	پ	پ	پ	p
ت	ت	ت	ت	t
ٹ	ٹ	ٹ	ٹ	t̤
ث	ث	ث	ث	t̤
ق	ق	ق	ق	q
ك	ك	ك	ك	k
خ	خ	خ	خ	kh
ج	ج	ج	ج	j
چ	چ	چ	چ	c
ح	ح	ح	ح	h
ڄ	ڄ	ڄ	ڄ	<u>kh</u>
د	د	د	د	d
ڊ	ڊ	ڊ	ڊ	d̤
ڌ	ڌ	ڌ	ڌ	d̤
ڇ	ڇ	ڇ	ڇ	ç
ڙ	ڙ	ڙ	ڙ	z
ر	ر	ر	ر	r
ڙ	ڙ	ڙ	ڙ	r̤
ڙ	ڙ	ڙ	ڙ	r̤
ز	ز	ز	ز	z
ڙ	ڙ	ڙ	ڙ	zh
س	س	س	س	s
ش	ش	ش	ش	sh
ص	ص	ص	ص	ʃ
ض	ض	ض	ض	ʒ
ط	ط	ط	ط	t̤
ظ	ظ	ظ	ظ	z̤
ع	ع	ع	ع	‘ (ayn)
غ	غ	غ	غ	gh̤
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
ر	ر	ر	ر	r (see Note 2)
و	و	و	و	v
ھ	ھ	ھ	ھ	h
-	-	ة	ة	t (see Rule 10)
ی	ی (ے ، ی)	ی (ے ، ی)	ی (ے ، 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	اَ
u	اُ
i	اِ
ā	آ
á	اَی ، اِی
ū	اُو
ī	اِی
o	اَو
e	اَی ، اِی
au	اَوِی
ai	اَی ، اِی

Notes to the Tables

1. For the use of ^l (alif) to support ء (hamzah) and مَ (maddah) see rules 1 and 2, respectively. For the romanization of ء by ' (alif) see rule 12. For other orthographic uses of ^l see rules 3-4.
2. For the distinction between ن and ر see rule 6.
3. For the distinction between ی and ے see rule 11(c) and (e).
4. For the form of the letter ۞ 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. ^l (*alif*), و and ی are used to support • (*hamzah*); see rule 12. When so used, these letters are not represented in romanization.
2. ^l (*alif*) is used to support ّ (maddah); see rule 13. When so used, it is not represented in romanization.
3. ^l (*alif*) is used after a consonant to indicate the long vowel romanized ā.

rāj
karnā

راج
کرنا

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

da‘vā

دعویٰ

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

‘Abdurrahmān عبد الرحمن ، عبد الرحمان

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

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

‘ilmān

علما

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 ṅ and dotted forms are romanized n.

jahāṅ

جہاں

7. و is used:

(a) To represent the consonant sound romanized v.

dev
vujūd

دیو
وجود

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

khvīsh

خوبیش

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

ūkh	اوخ
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. ◦ is used to represent the consonantal sound romanized
- h*
- .

ham	هم
gāh	گاہ

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

kih	کہ
guldestah	گلدستہ

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

mukh	مکھہ
------	------

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

phūl	پھول
acchā	اچھا

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

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

ḥikmat	حکمت ، حکمت
--------	-------------

11. ی (often written ې) is used:

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

siyāsāt	سیاست
dayā	دیا

- (b) To represent the long vowel romanized *ī*.

taṣvīr
īshvār

تصویر
ایشوار

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

sher
nevā

شیر
نیوا

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

se
laṛke

سے
لڑکے

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

da'vā
'uqbā

دعوی
عقبی

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

maidān
bail

میدان
بیل

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

hai

ہے

- (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 *muḏāf* 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 ' (*alif*).

mu'min
li'e
bhā'ī

مؤمن
لئے
بہائی

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

malikah-yi Inglīstān ملکہ انگلستان

13. ̄ (maddah)

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

- (b) At the beginning of a syllable within a word, ̄ is romanized ā.
mir'āt مرآت
Qur'ān قرآن

- (c) ̄ 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 چچا
khaṭṭā کھٹا
makkhī مکھی

When ̄ occurs over و and ی, these letters are regarded as representing consonants. They are romanized vv and yy, respectively.

quvvat قوت
sayyid سید
Zakariyyā 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. Tanvīn (written ̄, ̄, ̄ (̄)) 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. *Iẓāfat*.

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

tārīkh-i Hindūstān تاریخ ہندوستان

- (b) When the *muzāf* ends in a vowel or in silent ه, -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 ة (*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" (ح ، ج ، ا ، ی ، ہ ، و ، م) it is romanized *al*.

، ک ، ق ، ف ، غ ، ع ، خ ، ع ، ی ، ا ، ب ، ج ، ح)	القرآن
al-Qur'ān	

- (b) When it is prefixed to a word beginning with a "sun letter" (ط ، ل ، ن ، 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.

ط ، ل ، ن ، ت ، ث ، د ، ذ ، ر)	السجل
as-sijill	

- (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
Abūlfazl	ابو الفضل
Zūlqarnain	ذو القرنین
Fazlullāh	فضل اللہ

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.

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 Māsṭar ṣāḥib-i Gavarnmaṭṭ Hā'ī 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>
á	acute	E2
ā	macron	E5
·	dot below	F2
⋈	double dot below	F3
̣	underscore	F6