

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	s
ض	ض	ض	ض	z
ط	ط	ط	ط	ṭ
ظ	ظ	ظ	ظ	z
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	آء ، آء
ū	وو
ī	ی
o	و
e	ے ، ی
au	وو
ai	ے ، ی

Notes

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

3. For the distinction between **ي** and **ء**, 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. **ا** (*alif*), **و** and **ي** are used to support **ء** (*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	كرنا

In some words of Arabic origin this *alif* appears as a superscript letter over **ي** 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. **ا** (*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 *t*, *d*, and *r*, respectively.
6. Regardless of pronunciation, undotted forms of the letter **ن** are romanized *n̄* and dotted forms are romanized *n*.

jahān̄	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.

khvīsh	خویش
--------	------

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

ūkh

Urdū

اوخ

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

گاہ

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

kih

کہ

guldastah

گلدستہ

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

مکھہ

9. ɔ (usually written in the form ه) 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

حکمة ، حکمت

11. ی (often written ی) is used:

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

siyāsat

سیاست

diyā

دیا

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

taṣvīr

تصویر

īshvar

ایشور

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

sher شیر

nevā نیوا

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

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 ى with this value is final, it is generally written ے .

hai ہے

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

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

For the use of ى 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 ' (*alif*).

mu'min مؤمن

li'e لئے

bhāī بهائی

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

malikah-yi Inqlistān ملکہ انگلستان

13. آ (*maddah*)

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

āb آب

- (b) At the beginning of a syllable within a word, Ā is romanized *ā*.

mir'āt مرآت

Qur'ān قرآن

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

14. $\overset{\circ}{\text{}}$ (*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	چچّا
khattā	کھٹّا
makkhī	مگھّی

When $\overset{\circ}{\text{}}$ occurs over g and y , these letters are regarded as representing consonants. They are romanized *vv* and *yy*, respectively.

quvvat	قوّت
sayyid	سیدّ
Zakariyyā	زکریّا

15. $\overset{\circ}{\text{}}$ (*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 $\overset{\circ}{\text{}}$, $\overset{\circ}{\text{}}$, $\overset{\circ}{\text{}}$ ($\overset{\circ}{\text{}}$)) 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 *muzāf*, the first of two words in the grammatical relationship known as *izāfah* ends in a consonant, *-i* is added to it in the romanization.

tārīkh-i Hindūstān	تاریخ هندوستان
--------------------	----------------

(b) When the *muzāf* 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	ملکه انگلستان

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

18. The Arabic article al 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، ک، م، و، ه، ی، ا، ب، ج، ح) it is romanized *al*.

al-Qur'ān	القران
-----------	--------

(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

Faḥlullāh

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 *muzāf* 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 ہید 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.

Pushto

Letters of the Alphabet (see Note 1)

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
ټ	ټ	ټ	ټ	ṭ
ث	ث	ث	ث	s
ج	ج	ج	ج	j
چ	چ	چ	چ	ch
ح	ح	ح	ح	ḥ
ځ	ځ	ځ	ځ	ś
ښ	ښ	ښ	ښ	ž
خ	خ	خ	خ	kh
د	د	د	د	d
ډ	ډ	ډ	ډ	ḍ
ذ	ذ	ذ	ذ	z
ر	ر	ر	ر	r
ړ	ړ	ړ	ړ	ṛ
ز	ز	ز	ز	z
ژ	ژ	ژ	ژ	zh
ښ	ښ	ښ	ښ	zh
س	س	س	س	s
ش	ش	ش	ش	sh
ښ	ښ	ښ	ښ	śh
ص	ص	ص	ص	ṣ
ض	ض	ض	ض	z
ط	ط	ط	ط	t
ظ	ظ	ظ	ظ	z
ع	ع	ع	ع	' (ayn)
غ	غ	غ	غ	gh
ف	ف	ف	ف	f
ق	ق	ق	ق	q
ک	ک	ک	ک	k

Initial	Medial	Final	Alone	Romanization
گ	گ	گ	گ	g
ک	ک	ک	ک	g
ل	ل	ل	ل	l
م	م	م	م	m
ن	ن	ن	ن	n
ڼ	ڼ	ڼ	ڼ	ṅ
ڼر	ڼر	ڼر	ڼر	ṅ
و	و	و	و	w (see Note 2)
ه	ه	ه، ة	ه، ة	h
ي	ي	ي، ى	ي، ى	y (see Note 3)

Vowels and Diphthongs (see Note 4)

Romanization	Initial	Medial	Final
a (see Note 5)	ا	-	-
u	أ	و	و
i	إ	ي	ي
ā	آ	آ	آ
á	-	-	آ (see Note 6)
ū	أو	و	و
ī	إي	ي	ي، ى
o	آو	و	و
e	إي، اې	ي، ې	ي، ى، ې، ى
aw	آو	و	و
ay	آي	ي	ي، ى
ay	-	-	ي، ى، ى

Notes

1. For the use of ا (*aliif*) see the table of vowels and diphthongs, and rules 1-4.
2. For other uses of و see the table of vowels and diphthongs, and rule 5.
3. For other forms and uses of ى see the table of vowels and diphthongs, and rules 6 and 7.
4. Vowel points are rarely found in Pushto writing and printing. Vowels not indicated in the script are supplied in romanization.

5. In addition to the three short vowels recognized by the orthography of all languages using Arabic script, Pushto possesses a short, central vowel for which the sign ˘ (differentiated from ˚) has sometimes been used. Both signs, whether written or inferred, may be romanized *a*. When this central vowel is indicated by ء (*hamzah*), the vowel is romanized *ạ*; the *hamzah* in this case is not represented in romanization. See rule 8(c) and (d). When it is desired for any reason to show the presence of a central vowel in a particular word, the romanization *ạ* may be used.
6. See rule 7(a).
7. See rule 7(b).

RULES OF APPLICATION

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

1. ʾ (*alif*) is used:
- (a) To indicate the presence of any of the short vowels *a*, *u*, *i*.
- | | |
|-------|-------|
| atəh | اتۀ |
| ūlas | اولس |
| in‘ām | انعام |
- (b) To represent the long vowel romanized *ā*. When so used, ʾ is usually written with ̄ (*maddah*).
- | | |
|------|-----|
| ādam | آدم |
|------|-----|
- (c) To represent, in combination with و or ی :
- (1) The long vowels romanized *ū* and *ī* respectively.
- | | |
|------|-------|
| ūṣh | اوبس |
| Īrān | ايران |
- (2) The long vowels romanized *o* and *e* respectively.
- | | |
|-------|-------|
| os | اوس |
| eṣhal | اينسل |
- (3) The diphthongs romanized *aw* and *ay* respectively.
- | | |
|-------|------|
| awwal | اول |
| ayyām | ايام |
2. ʾ (*alif*), when medial or final, is used to represent the long vowel romanized *ā*.
- | | |
|------|------|
| bābā | بابا |
|------|------|
3. ʾ (*alif*), when final, sometimes represents the combination romanized *an*. This value is confined to a few words of Arabic origin. See rule 10.

- (c) When the *muzāf* is marked by the addition of ی, it is followed by *-yi*.
 khānah-yi buzurg خانہ ی بزرگ

Affixes and Compounds

13. Affixes.

When the affix and the word with which it is connected grammatically are written separately in Pushto, the two are separated in romanization by a single prime (').

nā'ashnā نااشنا
 war'astawī وراستوی

Note 1: The particle د (*da*), meaning "of," is always written as a separate word in romanization.

da dāh د ده

Note 2: The Arabic article *a/* is separated from what follows by a hyphen. The / of the article is retained in romanization, regardless of whether or not this / is assimilated in pronunciation to the sound of the following letter.

'Abd al-Rashīd عبد الرشید
 Fazl al-Ḥaqq فضل الحق
 Nizām al-Dīn نظام الدین

14. Compounds.

The elements of a compound (except a compound personal name) are separated in romanization by a single prime (').

kitāb'khānah کتابخانه
but marīz'khānah مریضخانه

Orthography of Pushto in Romanization

15. Capitalization.

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

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

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

Yūnīwarsīfī Buk Ejānsī یونیورسٹی بک ایجنسی
 (*not* University Book Agency)
 Dārmastatar (*not* Darmesteter) دارمستتر

Sindhi (in Arabic script)

Letters of the Alphabet (see Note 1)

Initial	Medial	Final	Alone	Romanization
ا	ا	ا	ا	(omit)
ب	ب	ب	ب	b
پ	پ	پ	پ	p
ت	ت	ت	ت	t
ث	ث	ث	ث	th
ٺ	ٺ	ٺ	ٺ	t̪
ڙ	ڙ	ڙ	ڙ	t̪h
س	س	س	س	s
پ	پ	پ	پ	p
ج	ج	ج	ج	j
ڇ	ڇ	ڇ	ڇ	ç
جھ	جھ	جھ	جھ	jh
ڻ	ڻ	ڻ	ڻ	ñ
ڪ	ڪ	ڪ	ڪ	c
ڪھ	ڪھ	ڪھ	ڪھ	ch
ح	ح	ح	ح	ḥ
ڪھ	ڪھ	ڪھ	ڪھ	<u>kh</u>
د	د	د	د	d
دھ	دھ	دھ	دھ	dh
ڍ	ڍ	ڍ	ڍ	ḍ
ڍھ	ڍھ	ڍھ	ڍھ	ḍh
ڙ	ڙ	ڙ	ڙ	z
ر	ر	ر	ر	r
رھ	رھ	رھ	رھ	r̥
ز	ز	ز	ز	z
س	س	س	س	s
ش	ش	ش	ش	sh
ص	ص	ص	ص	ṣ
ڙھ	ڙھ	ڙھ	ڙھ	z̪
ط	ط	ط	ط	t̪

Initial	Medial	Final	Alone	Romanization
ظ	ظ	ظ	ظ	z
ع	ع	ع	ع	' (ayn)
غ	غ	غ	غ	gh
ف	ف	ف	ف	f
ق	ق	ق	ق	ph
ك	ك	ك	ك	q
ڪ	ڪ	ڪ	ڪ	k
گ	گ	گ	گ	kh
گ	گ	گ	گ	g
گ	گ	گ	گ	g̃
گھ	گھ	گھ	گھ	gh
-	گ	گ	گ	n̄
ل	ل	ل	ل	l
م	م	م	م	m
ن	ن	ن	ن	n
ڻ	ڻ	ڻ	ڻ	n̄
و	و	و	و	v
ھ	ھ	ھ ، ھ ، ھ	ھ ، ھ	h
-	-	-	ع	(omit)
ي	ي	ي	ي	y

Vowels and Diphthongs (see Note 2)

اَ	a	آ	ā
اُ	u	اُو	o
اِ	i	اِی ، اِی	e
آ ، اِی ، اِی	ā	اُو	au
اُو	ū	اِی ، اِی	ai

Notes

1. When Sindhi is written in the Devanagari script, it is romanized according to the Hindi table. The letters ख, ख, ख, and ङ are romanized *kh*, *kh̄*, *kḣ*, and *gh̄* respectively.
2. The rules of application for Urdu are to be used for Sindhi. Vocalization is that used for entries in Mewaram's *A Sindhi English Dictionary* (1st ed., Hyderabad, 1910), *Sindhī Ūrdū Lughāt* (Haidarābād, 1959), and *Jāmi'-ī Sindhī Lughāt* (Karācī, 1931).