

Arabic

Letters of the Alphabet

Initial	Medial	Final	Alone	Romanization
ا	ا	ا	ا	omit (see Note 1)
ب	ب	ب	ب	b
ت	ت	ت	ت	t
ث	ث	ث	ث	th
ج	ج	ج	ج	j
ح	ح	ح	ح	h
خ	خ	خ	خ	kh
د	د	د	د	d
ذ	ذ	ذ	ذ	dh
ر	ر	ر	ر	r
ز	ز	ز	ز	z
س	س	س	س	s
ش	ش	ش	ش	sh
ص	ص	ص	ص	ṣ
ض	ض	ض	ض	ḍ
ط	ط	ط	ط	ṭ
ظ	ظ	ظ	ظ	ẓ
ع	ع	ع	ع	' (ayn)
غ	غ	غ	غ	gh
ف	ف	ف	ف	f (see Note 2)
ق	ق	ق	ق	q (see Note 2)
ك	ك	ك	ك	k
ل	ل	ل	ل	l
م	م	م	م	m
ن	ن	ن	ن	n
ه	ه	ه , هـ	ه , هـ	h (see Note 3)
و	و	و	و	w
ي	ي	ي	ي	y

Vowels and Diphthongs

ā	a	ā	ā (see Rule 5)	ā	ī
ū	u	ū	ū (see Rule 6(a))	ū	aw
ī	i	ī	ī	ī	ay

Letters Representing Non-Arabic Consonants

This list is not exhaustive. It should be noted that a letter in this group may have more than one phonetic value, depending on the country or area where it is used, and that the romanization will vary accordingly.

گ	g	چ	ch	ق	v
ك	g	ج	zh	ق	v
پ	p	ز	zh	پ	v

Notes

1. For the use of *alif* to support *hamzah*, see rule 2. For the romanization of *hamzah* by the consonantal sign ' (alif) see rule 8(a). For other orthographic uses of *alif* see rules 3-5.
2. The *Maghribi* variations پ and ق are romanized *f* and *q* respectively.
3. ð in a word in the construct state is romanized *t*. See rule 7(b).

RULES OF APPLICATION

Arabic Letters Romanized in Different Ways Depending on Their Context

1. As indicated in the table, و and ي may represent:

- (a) The consonants romanized *w* and *y*, respectively

waq'	وضع
'iwaq	عوض
dalw	دلو
yad	يد
hiyal	حبل
tahy	طهي

- (b) The long vowels romanized *ū*, *ī*, and *ā* respectively

ūlā	اولى
ṣūrah	صورة
dhū	ذو
īmān	ايمان
jīl	جيل
fī	في
kitāb	كتاب
saḥāb	سحاب
jumān	جمان

See also rules 11(a) and 11(b1-2).

- (c) The diphthongs romanized *aw* and *ay*, respectively

awj	أوج
nawm	نوم
law	لو
aysar	ايسر
shaykh	شيخ
'aynay	عيني

See also rules 11(a)(2) and 11(b)(3).

2. | (*alif*), و and ى when used to support ة (*hamzah*) are not represented in romanization. See rule 8(a).
3. | (*alif*) when used to support ِ (*waslah*) and ّ (*maddah*) is not represented in romanization. See rules 9 and 10.
4. | (*alif*) and و when used as orthographic signs without phonetic significance are not represented in romanization.

fa'alū	فعلوا
ulā'ika	اولائك
'ilman wa-'amalan	علماء وعملاء

See also rule 12 and examples cited in rules 23-26.

5. | (*alif*) is used to represent the long vowel romanized *ā*, as indicated in the table.

fā'il	فاعل
riḡā	رضا

This | when medial, is sometimes omitted in Arabic; it is always indicated in romanization. See rule 19.

6. Final ى appears in the following special cases:

- (a) As ى (*alif maqṣūrah*) used in place of ِ to represent the long vowel romanized *ā*.

ḡatā	حَتَّى
maqā	مَتَّى
kubrā	كَبْرَى
Yaḡyā	يَحْيَى
musammā	مَسْمَى
Muṡṡafā	مصطفى

- (b) As قِي in nouns and adjectives of the form *fa'īl* which are derived from defective roots. This ending is romanized *ī*, not *īy*, without regard to the presence of ذ (*shaddah*). See rule 11(b2).

Raḍī al-Dīn

رضى الدين

Compare the *fa'īl* form of the same root الرضى [with out *shaddah*] *al-Raḍī*.

- (c) As قِي in the relative adjective (*nisbah*). The ending, like (b) above, is romanized *ī*, not *īy*.

al-Miṣrī

المصري

Compare المصريّة *al-Miṣrīyah* and see rule 11(b1).

7. ة (*tā' marbūṭah*)

- (a) When the noun or adjective ending in ة is indefinite, or is preceded by the definite article, ة is romanized *h*. The ة in such positions is often replaced by ة .

ṣalāh

al-Risālah al-bahīyah

mir'āh

Urjūzah fī al-ṭibb

صلاة

الرسالة البهية

مرآة

ارجوزة في الطب

- (b) When the word ending in ة is in the construct state, ة is romanized *t*.

Wizārat al-Tarbiyah

Mir'āt al-zamān

وزارة التربية

مرآة الزمان

- (c) When the word ending in ة is used adverbally, ة (vocalized ة) is romanized *tan*. See rule 12(b).

Romanization of Arabic Orthographic Symbols Other than Letters and Vowel Signs

The signs listed below are frequently omitted from unvocalized Arabic writing and printing; their presence or absence must then be inferred. They are represented in romanization according to the following rules:

8. ة (*hamzah*)

- (a) In initial position, whether at the beginning of a word, following a prefixed preposition or conjunction, or following the definite article, ة is not represented in romanization. When medial or final, ة is romanized as ' (alif).

asad	أسد
uns	أنس
idha	إذًا
mas'alah	مسألة
mu'tamar	مؤتمر
dā'im	دائم
mala'a	ملا
khaṭī'a	خلى

(b) ة , when replaced by the sign ِ (waslah) and then known as *hamzat al-waṣl*, it is not represented in romanization. See rule 9 below.

9. ِ (waslah), like initial ة , is not represented in romanization. See also rule 8(b) above. When the *alif* which supports *waslah* belongs to the article ال , the initial vowel of the article is romanized *a*. See rule 17(b). In other words, beginning with *hamzat al-waṣl*, the initial vowel is romanized *i*.

Riḥlat Ibn Jubayr	رحلة ابن جبير
al-istidrāk	الاستدراك
kutub iqtanat' hā	كتب اقتنتها
bi-ihtimām 'Abd al-Majīd	بأهتمام عبد المجيد

10. ̣ (maddah)

(a) Initial ̣ is romanized *ā*

ālah	آلة
Kullīyat al-Ādāb	كلية الآداب

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

ta'ālīf	تأليف
ma'āthir	مآثر

(c) ̣ is otherwise not represented in romanization.

khulafā'	خلفاء
----------	-------

11. ̣ (shaddah or tashdīd)

(a) Over و :

(1) و̣, representing the combination of long vowel plus consonant, is romanized *ūw*.

'adūw	عدو
qūwah	قوة

See also rule 1(b).

- (2) شَوَّال , representing the combination of diphthong plus consonant, is romanized *aww*.
 Shawwāl شَوَّال
 ṣawwara صَوَّرَ
 jaww جَوَّ

See also rule 1(c).

- (b) Over ي :

- (1) Medial يَّ , representing the combination of long vowel plus consonant, is romanized *yy*.
 al-Miṣrīyah المِصْرِيَّة

See also rule 1(b)

- (2) Final ي is romanized *ī*. See rules 6(b) and 6(c).

- (3) Medial and final يَّي , representing the combination of diphthong plus consonant, is romanized *ayy*.
 ayyām أَيَّام
 sayyid سَيِّد
 Quṣayy قُصَيِّ

See also rule 1(c)

- (c) Over other letters, ش is represented in romanization by doubling the letter or digraph concerned.

al-Ghazzī الغَزَّزِي
 al-Kashshāf الكَشَّاف

12. *Tanwīn* may take the written form ن , نِ (نَ), or نٍ , romanized *un*, *an*, and *in*, respectively. *Tanwīn* is normally disregarded in romanization, however. It is indicated in the following cases:

- (a) When it occurs in indefinite nouns derived from defective roots.

qāḍīn قَاضٍ
 ma'nan مَعْنَى

- (b) When it indicates the adverbial use of a noun or adjective.

ṭab'an طَبَعًا
 faj'atan فَجَاءَ
 al-Mushtarik waḍ'an wa-al-muftariq والمُشْتَرِكُ وَضَعًا وَالْمُفْتَرِقُ
 ṣuq'an صُقِعًا

Grammatical Structure as It Affects Romanization

13. Final inflections of verbs are retained in romanization, except in pause.
- | | |
|---------------------------------------|------------------------------|
| man waliya Miṣr | من ولي مصر |
| ma'rifat mā jāyibu la-hum | معرفة ما يجب لهم |
| ṣallā Allāh 'alayhi wa-sallam | صلى الله عليه وسلم |
| al-Lu'lu' al-maknūn fī ḥukm al-akhbār | الؤلؤ المكنون في حكم الاخبار |
| 'ammā sayakūn | عما سيكون |
14. Final inflections of nouns and adjectives:
- (a) Vocalic endings are not represented in romanization, except preceding pronominal suffixes, and except when the text being romanized is in verse.
- | | |
|------------------------------|-----------------------------|
| Ma'had Mawlāya al-Ḥasan | معهد مولاي الحسن |
| uṣūluhā al-nafsiyah wa-ṭuruq | اصولها النفسية وطرق تدريسها |
| tadrīsihā | |
| ilā yawminā hādhā | الى يومنا هذا |
- (b) *Tanwīn* is not represented in romanization, except as specified in rule 12.
- (c) *ō* (*tā' marbūṭah*) is romanized *h* or *t* as specified in rule 7.
- (d) For the romanization of the relative adjective (*nisbah*) see rule 6(c).
15. Pronouns, pronominal suffixes, and demonstratives:
- (a) Vocalic endings are retained in romanization.
- | | |
|----------------------------|-------------------|
| anā wa-anta | انا وانت |
| hādhihi al-ḥāl | هذه الحال |
| mu'allafātuhu wa-shurūḥuhā | مؤلفاته وشروحاتها |
- (b) At the close of a phrase or sentence, the ending is romanized in its pausal form.
- | | |
|--------------------|-------------------------|
| ḥayātuhu wa-'aṣruh | حياته وعصره |
| Tawfīq al-Ḥakīm, | توفيق الحكيم ، أفكاره ، |
| alkāruh, āthāruh | آثاره |
16. Prepositions and conjunctions:
- (a) Final vowels of separable prepositions and conjunctions are retained in romanization.
- | | |
|---------------|----------|
| anna | ان |
| annahu | انه |
| bayna yadayhu | بين يديه |

Note the special cases: مما *mimmā*, ممن *mimman*.

- (b) Inseparable prepositions, conjunctions, and other prefixes are connected with what follows by a hyphen.

bi-hi	به
wa-ma'ahu	ومعه
lā-silkī	لاسلكى

17. The definite article:

- (a) The romanized form *al* is connected with the following word by a hyphen.

al-kitāb al-thānī	الكتاب الثاني
al-ittiḥād	الاتحاد
al-aṣl	الاصل
al-āthār	الآثار

- (b) When *al* is initial in the word, and when it follows an inseparable preposition or conjunction, it is always romanized *al* regardless of whether the preceding word, as romanized, ends in a vowel or a consonant.

ilá al-ān	الى الآن
Abū al-Wafā'	ابو الوفاء
Maktabat al-Nahḍah al-Miṣrīyah	مكتبة النهضة المصرية
bi-al-tamām wa-al-kamāl	بالتمام والكمال

Note the exceptional treatment of the preposition *l* followed by the article:

lil-Shirbīnī	للشربيني
--------------	----------

See also rule 23.

- (c) The *l* of the article is always romanized *l*, whether it is followed by a "sun letter" or not, i.e., regardless of whether or not it is assimilated in pronunciation to the initial consonant of the word to which it is attached.

al-ḥurūf al-abjadīyah	الحروف الابدعية
Abū al-Layth al-Samarqandī	ابو الليث السمرقندي

Orthography of Arabic in Romanization

18. Capitalization:

- (a) Rules for the capitalization of English are followed, except that the definite article *al* is given in lower case in all positions.
- (b) Diacritics are used with both upper and lower case letters.

al-Ījī	الإيجي
al-Ālūsī	الآلوسي

19. The macron or the acute accent, as appropriate, is used to indicate all long vowels, including those which in Arabic script are written defectively. The macron or the acute accent, as the case may be, is retained over final long vowels which are shortened in pronunciation before *hamzat al-waṣf*.

Ibrāhīm	ابراهيم ، ابرهيم
Dā'ūd	داؤود ، داؤد
Abū al-Ḥasan	ابو الحسن
ru'ūs	رؤس
dhālika	ذلك
'alā al-'ayn	على العين

20. The hyphen is used:

- To connect the definite article *al* with the word to which it is attached. See rule 17(a).
- Between an inseparable prefix and what follows. See rules 16(b) and 17(b) above.
- Between *bin* and the following element in personal names when they are written in Arabic as a single word. See rule 25.

21. The prime (') is used:

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

Ad'ham	ادهم
akramat'hā	اكرمته

- To mark the use of a letter in its final form when it occurs in the middle of a word.

Qal'ah'jī	قلعةجي
Shaykh'zādah	شيخزاده

22. As in the case of romanization from other languages, foreign words which occur in an Arabic context and are written in Arabic letters are romanized according to the rules for romanizing Arabic.

Jārmānūs (<i>not</i> Germanos <i>nor</i> Germanus)	جارمانوس
Lürd Ghrānīl (<i>not</i> Lord Granville)	لورد غرانفيل
Isāghūjī (<i>not</i> Isagoge)	ايساغوجي

For short vowels not indicated in the Arabic, the Arabic vowel nearest to the original pronunciation is supplied.

Gharsiyā Khayīn (<i>not</i> García Jaén)	غرسيا خين
---	-----------

Examples of Irregular Arabic Orthography

23. Note the romanization of الله, alone and in combination.
- | | |
|---------------------|----------------|
| Allāh | الله |
| billāh | بالله |
| lillāh | لله |
| bismillāh | بسم الله |
| al-Mustanṣir billāh | المستنصر بالله |
24. Note the romanization of the following personal names.
- | | |
|--------|-------------|
| Tāhā | طه |
| Yāsīn | يس ، يسن |
| 'Amr | عمرو |
| Bahjat | بهجت ، بهجة |
25. ابن and بن are both romanized *ibn* in all positions.
- | | |
|-------------------------------------|----------------------------|
| Aḥmad ibn Muḥammad ibn Abī al-Rabiʿ | احمد بن محمد بن ابي الربيع |
| Sharḥ Ibn 'Aqīl 'alā Alfīyat | شرح ابن عقيل على الفية |
| Ibn Mālīk | ابن مالك |

Exception is made in the case of modern names, typically North African, in which the element *bin* is pronounced *bin*.

Bin Khiddah	بن خده
Bin-'Abd Allāh	بنعبد الله

26. Note the anomalous spelling مائة, romanized *mi'ah*.

SPECIAL CHARACTERS AND CHARACTER MODIFIERS IN ROMANIZATION

Special Characters	Name	USMARC hexadecimal code
'	soft sign (prime)	A7
•	alif (hamzah)	AE
ء	ayn	B0
Character Modifiers	Name	USMARC hexadecimal code
acute	acute	E2
macron	macron	E5
dot below	dot below	F2