Arabic

The current United Nations recommended romanization system was approved in 2017 (resolution XI/3), based on the system adopted by Arabic experts at the conference held in Beirut in 2007, the Unified Arabic Transliteration System, taking into account the practical amendments and corrections carried out and agreed upon by the representatives of the Arabic-speaking countries at the Fourth Arab Conference on Geographical Names, held in Beirut in 2008, and some clarifications and amendments agreed in Riyadh in 2017[1].

Previously, the United Nations had approved a romanization system in 1972 (resolution II/8), based on the system adopted by Arabic experts at the conference held at Beirut in 1971 with the practical amendments carried out and agreed upon by the representatives of the Arabic-speaking countries at their conference. The table was published in volume II of the conference report[2].

In UN resolution XI/3 it is specifically stated that the system was recommended for the "romanization of the geographical names within those Arabic-speaking countries where this system is officially adopted". There is evidence of its partial implementation in Jordan, Oman and Saudi Arabia. The UNGEGN Working Group on Romanization Systems intends to continue monitoring the UN system's implementation across Arabic-speaking countries.

In some countries there exist local romanization schemes or practices. The geographical names of Algeria, Djibouti, Mauritania, Morocco and Tunisia are generally rendered in the traditional manner which conforms to the principles of the French orthography.

The previous UN-approved system is still found in considerable international usage.

Arabic is written from right to left. The Arabic script usually omits vowel points and diacritical marks from writing which makes it difficult to obtain uniform results in the romanization of Arabic. It is essential to identify correctly the words which appear in any particular name and to know the standard Arabic-script spelling including the relevant vowels. One must also take into account dialectal and idiosyncratic deviations. The romanization is generally reversible though there may be some ambiguous letter sequences (*dh, kh, sh, th*) which may also point to combinations of Arabic characters in addition to the respective single characters.

Romanization

In the romanization system below column 1 denotes an independent consonant character, column 2 the initial, column 3 the medial and column 4 the final form of a character. Column 5 gives the romanization equivalent.

	1	2	3	4	5
1	ء				', -

				A
2	ا			ٵ	_ [B]
3	ب	ب	بـ	ـب	b
4	ت	ت	تـ	ـت	t [C]
5	ث	ث	ثـ	ـث	th
6	ج	ج	جـ	ـج	j
7	ح	ح	حـ	ـح	ḥ
8	خ	خ	خـ	ـخ	kh
9	د			ـد	d
10	ذ			ـذ	dh
11	ر			ـر	r
12	ز			ـز	z
13	س	س	سـ	ـس	s
14	ش	ش	شـ	ـش	sh
15	ص	ص	صـ	ـص	ṣ
16	ض	ض	ضـ	ـض	ḍ
17	ط	ط	طـ	ـط	ṭ
18	ظ	ظ	ظـ	ـظ	ḍh
19	ع	ع	عـ	ـع	'
20	غ	غ	غـ	ـغ	gh
21	ف	ف	فـ	ـف	f
22	ق	ق	قـ	ـق	q
23	ك	ك	كـ	ـك	k
24	ل	ل	لـ	ـل	l
25	م	م	مـ	ـم	m
26	ن	ن	نـ	ـن	n
27	ه	ه	هـ	ـه	h [C]
28	و			ـو	w
29	ي	ي	يـ	ـي	y

- [A] Not romanized word-initially.
- [B] Not romanized, but see romanizations accompanying *alif* (ا) in the table for vowels.

- C In certain endings, an original *tā'* (ت) is written ة, i.e., like *hā'* (ه) with two dots, and is known as *tā' marbūṭah*. It is romanized *h*, except in the construct form of feminine nouns, where it is romanized *t*, instead.

Vowels, diphthongs and diacritical marks (- stands for any consonant)

1	ـَ	a
2	ـَوْ	aw
3	ـَيْ	ay
4	ـِ	i
5	ـُ	u
6	ـْ	(A)
7	ـَا	ā
8	آ	ā
9	ـِي	ī
10	ـُو	ū
11	ـَى	á
12	ـّ	(B)

- A Marks absence of the vowel.
- B Marks doubling of the consonant.

Notes.

1. When the definite article *al* precedes a word beginning with one of the "sun letters" (*t, th, d, dh, r, z, s, sh, s, ḍ, ṭ, dh, l, n*) the *l* of the definite article is assimilated with the first consonant of the word: الشارقة Ash Shāriqah.
2. The definite article is always written with a capital initial: الزيتون Az Zaytūn, البلد Al Balad, منية الضنية Minyat Aḏ Ḏinniyyah.
3. Nunation is unlikely to be found in geographical names and the last letter remains silent: جبلٌ = جبل Jabal (not Jabalun).
4. In order to disambiguate certain character sequences a middle dot (·) may be used: سهيلة S·haylah (cf. شيلة Shaylah), دهيب D·hayb (cf. ذيب Dhayb), أدهم Ad·ham (cf. أدم Adham).

Other systems of romanization

The previous **UN 1972 System** had the following differences:

1. the character (ظ) was romanized as *ẓ* instead of *dh*;
2. the cedilla (ؚ) was used instead of sub-macron (‾) in all characters with sub-macrons.

The **BGN/PCGN 1956 System** is almost identical to the previous UN 1972 System. The only difference lies in the treatment of articles articles (and the -iyyah ending, represented -īyah in the BGN/PCGN system). The original transliteration table, published in vol. II of the report on the Second UN Conference on the Standardization of Geographical Names, contains examples (but not explicit rules) where the definite article is always written with a small initial and connected by hyphen to the main part of the name, e.g. البصرة al-Baṣrah, الرياض ar-Riyāḍ. The practice of the BGN and the PCGN, however, is not to use hyphens between articles and names and to capitalize the definite article where it appears at the beginning of a name, e.g. Al Baṣrah, Ar Riyāḍ.

The **I.G.N. System 1973** (sometimes also called Variant B of the Amended Beirut System[3]) has the following equivalents to the romanizations of the UN system:

UN	I.G.N.
a	= a, e, é, è[A]
ā	= â, ê[A]
á	= a
ḏ	= ḍ
dh	= ẓ
ẖ	= ḥ
i	= i, e[A]
ī	= î, ê[A]
j	= dj, j[A]
n	= n, ne[B]
q	= q, g[A], gu[C]
s	= s, ss[D]
ṣ	= ṣ, ç[E]
sh	= ch
ṯ	= ṭ
u	= ou, o[A]
ū	= oû, ô[A]
w	= ou
y	= i, ï[F], y[G]
ʼ	(*not romanized*)
ʻ	= ʼ, aa[H]

- [A] Used according to the local pronunciation.
- [B] *ne* is used word-finally after *a, e, i, o*.
- [C] *gu* is used before *e* and *i*.
- [D] *ss* is used between vowels.
- [E] *ç* is used in some conventional names.
- [F] *ï* is used as second element of a diphthong.

- ^G *y* is used word-initially or between vowels.
- ^H *aa* is specific to Lebanon.

The transliteration **ISO 233:1984** gives every character and diacritical mark a unique equivalent and e.g. long vowels in Arabic *ā, ī* and *ū* are consequently written *a', iy* and *uw* respectively in the ISO transliteration. Other main correspondences:

UN	ISO
ā (آ)	= â
á	= ay
ḏ	= ḍ
dh	= ḏ
dh	= ẓ
gh	= ġ
h (ة)	= ẗ
ḥ	= ḥ
j	= ǧ
kh	= ḫ
ṣ	= ṣ
sh	= š
ṯ	= ṭ
th	= ṯ
y (ى)	= ỳ
y (ي)	= y

The **Survey of Egypt System** (SES) of romanization has the following correspondences with the UN system:

UN	SES
á	= a
ā	= â (a)
-ah (ة)	= -a
aw	= ô (au)
ay	= ei (ai)
ḏ	= ḍ
dh	= dh (z)
dh	= ẓ (d)
ḥ	= ḥ
ī	= î

j	= g (j)
q	= q (k)
s	= s (c)
s̲	= ṣ
t̲	= ṭ
th	= th (t)
ū	= û
ʻ	= ʻ

The variants in parentheses are used depending on pronunciation and tradition. Not all the variations have been given above. The article is always written *el-* (El-Kafr el-Qadîm, Sharm el-Sheikh).

In **Algeria** there is at present no official romanization system, the possibility of establishing such a system is being discussed in the Permanent Commission for Toponymy (CPST) at the National Council of Geographical Information (CNIG)[4]. In **Mauritania**, the romanized name forms in official maps edited since 1969 have been rendered in accordance with a simplified version of the I.G.N. system[5]. In **Morocco** the official romanization system for Arabic script dates from June 17, 1932, although changes to this are being planned[6]. In **Tunisia** the Directorate of Topography and Cartography adopted officially the amended Beirut system with minor modifications (e.g. adding a letter *g* to the table) in 1983 but later reverted to traditional rendering. Information on some of the countries above might be out of date and needs further survey.

References

1. *Romanization system from Arabic letters to Latinized letters 2007*. Eleventh United Nations Conference on the Standardization of Geographical Names. New York, 8–17 August 2017. E/CONF.105/137/CRP.137 (+ Annex 1 and Annex 2).
2. *Second United Nations Conference on the Standardization of Geographical Names.* London, 10–31 May 1972. Vol. II. *Technical papers*, p. 170.
3. *Présentation de la variante B du système de translittération de l'arabe "Beyrouth amendé".* UNGEGN, 17th Session. New York, 13–24 June 1994. WP No. 61.
4. *Rapport de l'Algérie.* Huitième Conférence des Nations Unies sur la normalisation des noms géographiques. Berlin, 27 août – 5 septembre 2002. E/CONF.94/INF.37.
5. Report of the Working Group on a Single Romanization System for Each Non-Roman Writing System: Activities from 1 June 1972 to 16 August 1977. *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August – 7 September 1977. Vol. II. Technical papers, pp. 402–403.
6. *Rapport national sur la toponymie* (Maroc). Huitième Conférence des Nations Unies sur la normalisation des noms géographiques. Berlin, 27 août – 5 septembre 2002. E/CONF.94/INF.76.