18 June 2002

English only

Eighth United Nations Conference on the
Standardization of Geographical Names
Berlin, 27 August-5 September 2002
Item 16 (a) of the provisional agenda*
Writing systems: romanization

Report on the current status of the United Nations romanization systems for geographical names

Submitted by the United Nations Group of Experts on Geographical Names

Report Contents

INTRODUCTION

The report is an updated version of a document presented to the 20th Session of the United Nations Group of Experts on Geographical Names (UNGEGN, New York, 17–28 January 2000, Working Paper No. 34 and 34/Corr.). The overall aim of the report is to clarify the present status and the degree of implementation of the United Nations romanization systems used for converting geographical names from other writing systems than Roman. Since 1967, the United Nations conferences on the standardization of geographical names have adopted resolutions recommending romanization systems that cover 28 languages and/or scripts: Amharic, Arabic, Assamese, Bengali, Bulgarian, Chinese, Greek, Gujarati, Hebrew, Hindi, Kannada, Khmer, Macedonian Cyrillic, Malayalam, Marathi, Mongolian (in China), Nepali, Oriya, Persian, Punjabi, Russian, Serbo-Croatian Cyrillic, Tamil, Telugu, Thai, Tibetan, Uighur, Urdu. Not all of them have been implemented on equal level, both nationally and internationally.

It has been agreed from the beginning that one of the general aims of international standardization is *"to arrive at an agreement on a single romanization system, based on scientific principles, from each non-Roman alphabet or script, for international application"*

(Resolution 9 of the first UN Conference on the Standardization of Geographical Names, 1967). The Fourth United Nations Conference on the Standardization of Geographical Names (1982) also adopted Resolution 15 which called for guidelines governing the consideration of romanization systems. The resolution stated that

1. the new romanization systems for international use should be considered only on condition that the sponsoring nations implement such systems on their cartographical products;
2. States should refrain from revising systems previously adopted for international use.

In the reports of the UNGEGN Working Group on Romanization Systems and elsewhere it has been emphasized that the process of adoption of a romanization system by the UN involves the following prerequisites.

1. Sufficient time is required for appropriate consultations and an expression of all views on technical matters between the sponsoring country and the WG members.
2. Specifically, the group considers whether the romanization system is based on sound scientific principles, the system's degree of reversibility, and the extent of its implementation on cartographic products (maps and charts) by the proposing country.

There are many aspects that one should consider when adopting new romanization systems for international use. These concern both the nature of the system itself, practical considerations and expectations for its future usage. The romanization systems would benefit from the following.

1. The system should be reversible, i.e., it should be possible to reconstruct the original non-Roman script form on the basis of romanization.
2. The system should be as simple and clear to use as possible. For example, if a character table is not adequate, the notes to the table should clarify all aspects of usage. It would not be to the advantage of the system if dictionaries and other sources would need to be consulted to obtain a correct romanization of names, or if the rules would allow for variations in the romanization of the same original script form.
3. The romanized name forms should be as easy as possible to write, read, memorize and also store electronically. That would mean minimal use of diacritical marks, avoiding difficult and unusual character sequences, etc. A systematic representation of phonological features is also to be recommended.

If the above conditions should be irreconcilable as is sometimes the case, it is the practical aspects that are the most decisive. For writing systems which would not allow for practicable and easy-to-use reversible romanization systems, it is often necessary to adopt two different romanization methods. The transcription method is the one recommended for wide usage, including the romanization of geographical names which is the area of interest for the United Nations. The second method, transliteration, could be applied for rendering the original script forms, e.g., in bibliography. Such systems fall into the sphere of activities of the International Organization for Standardization (ISO). For one such language there may exist two totally differing systems. For example, for Thai there is the UN system 1967 and ISO 11940:1998. Or these two can be combined into one integral system, as, for example, for Greek for which exists the UN system 1987 (ELOT 743) containing two versions of romanization: one for general purposes and the other for documentation, i.e., a fully reversible one. Reversibility is achieved by adding diacritical marks in case there could be ambiguities.

A new romanization system is typically submitted by the proposing country in form of a resolution to the next full United Nations Conference on the Standardization of Geographical Names. When the resolution is adopted, the subject romanization system becomes the United Nations standard.

The present report consists of two main parts. The first part lists languages with non-Roman writing systems and the romanization systems used to render geographical names from these languages.

The second part of the report presents an overview of romanization systems for each of the languages. A language section usually consists of an introduction that gives the origin of the UN system, describes the implementation of the system both nationally and internationally and briefly characterizes the system itself regarding the ease of applicability, reversibility, possible ambiguities and other problems.

The romanization system is presented with notes on its application if appropriate. While trying to preserve accurately the content of the original tables published in the UN materials, an attempt has been made to harmonize and consolidate their presentation. Where possible, all non-Roman characters are listed in the order which is traditional for the given language. Examples of romanization and details on pronunciation have been left out of the tables in this version as it is difficult to find sources for all the presented languages. Examples may be found in supplementary notes.

Briefly listed and compared with the main system are also other systems of romanization.

Updates

Compared with Version 1.3 (March 2000), the following major changes should be noted.
1. The general layout of the report was improved, some of the fonts (esp. vowel marks for the Perso-Arabic script) were replaced with more legible ones.
2. For reference purposes all characters in romanization tables were numbered.
3. New romanization systems were provided for Byelorussian, Georgian, Korean and Tigrinya.
4. Substantial new information was added in the sections for Armenian and Maldivian.
5. Turkmen and Uzbek sections were moved to Annex: Languages that have recently adopted Roman alphabets.

LIST OF LANGUAGES, COUNTRIES AND ROMANIZATION SYSTEMS

The following table contains languages with non-Roman writing systems as identified in the Report of the Working Group on Toponymic Data Exchange Formats and Standards (7th United Nations Conference on the Standardization of Geographical Names. New York, 13-22 January 1998. Document E/CONF.91/CRP.11). It also contains some other languages mentioned in the toponymic guidelines for individual countries.

Languages are arranged alphabetically according to their English name. The names of languages conform to those mentioned in the appropriate United Nations resolutions.

4

In column 3 countries are shown where at least one of the official languages is the language concerned. Official languages are identifed in the report of the Working Group of Country Names (7th United Nations Conference on the Standardization of Geographical Names. New York, 13-22 January 1998. Document E/CONF.91/L.65 and Add. 1).

In column 4 the year of the adoption of the system by the United Nations is given. The numerals that follow in parentheses refer to the United Nations Conference on the Standardization of Geographical Names (preceding the slash), and to the resolution number (following the slash).

In column 5 various other systems are mentioned that could have some international usage. The equal sign (=) will indicate that the systems in question more or less coincide. By a national system one would ordinarily mean those adopted by national cartographic services for rendering their geographical names. BGN/PCGN refers to systems jointly adopted by the United States Board on Geographic Names and the Permanent Committee on Geographical Names for British Official Use. Almost all of these systems have been published in a book "Romanization Systems and Roman-Script Spelling Conventions" by the Defense Mapping Agency (U.S.) in 1994. I.G.N. stands for *Institut Géographique National* in France. ISO is the abbreviation for the International Organization for Standardization, and its systems here are presented for the sake of integrity. ISO transliteration schemes are normally not used for rendering geographical names.

Language	Writing System	Country (Countries)	Romanization Systems	
			United Nations	Other
Amharic	Ethiopic	Ethiopia	1967 (I/17)	
				BGN/PCGN 1967
Arabic	Perso-Arabic	(General)	1972 (II/8)	= BGN/PCGN 1956
				I.G.N. System 1973
				ISO 233:1984 (transliteration) ISO 233-2:1993 (simplified transliteration)
		Algeria	(see General)	
		Bahrain	(see General)	
		Chad	(see General)	
		Comoros (the)	(see General)	
		Djibouti	(see General)	
		Egypt	(see General)	
				National: Survey of Egypt
		Eritrea	(see General)	
		Iraq	(see General)	
		Israel	(see General)	
		Jordan	(see General)	
				National: Royal Jordanian Geographic Centre System
		Kuwait	(see General)	

		Lebanon	(see *General*)	
				National: 1963
		Libyan Arab Jamahiriya (the)	(see *General*)	
		Mauritania	(see *General*)	
		Morocco	(see (*General*)	
				National: 1972
		Oman	(see *General*)	
		Qatar	(see *General*)	
		Saudi Arabia	(see *General*)	
		Somalia	(see *General*)	
		Sudan (the)	(see *General*)	
		Syrian Arab Republic (the)	(see *General*)	
		Tunisia	(see *General*)	
		United Arab Emirates (the)	(see *General*)	
		Yemen	(see *General*)	
Armenian	Armenian	Armenia	—	BGN/PCGN 1981
				ISO 9985:1996
Assamese	Bengali	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
Bengali	Bengali	Bangladesh	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
		India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
Bulgarian	Cyrillic	Bulgaria	1977 (III/10)	
				BGN/PCGN 1952
Burmese	Burmese	Myanmar	—	BGN/PCGN 1970
Byelorussian	Cyrillic	Belarus	—	National: 2000
				BGN/PCGN 1979
Chinese	Chinese	China	1977 (III/8)	= BGN/PCGN 1979; = ISO 7098:1991
				Modified Wade-Giles (1912)
		Singapore	—	Modified Wade-Giles (1912)
Dari	Perso-Arabic	Afghanistan	1967 (I/13)[1]	
Dzongkha	Tibetan	Bhutan	—	National: 1994
Georgian	Georgian	Georgia	—	National: 2002
				BGN/PCGN 1981
				ISO 9984:1996

Greek	Greek	Cyprus	1987 (V/19)	= ELOT 743, = ISO 843:1997
		Greece	1987 (V/19)	
Gujarati	Gujarati	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
Hebrew	Hebrew	Israel	1977 (III/13)	= BGN/PCGN 1962
				ISO 259:1984 (transliteration) ISO 259-2:1994 (simplified transliteration), ISO/DIS 259-3 (phonemic conversion)
Hindi	Devanagari	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
				ISO 15919:2001
Japanese	Sino-Japanese	Japan	—	National: *Kunrei-siki* (1954), = ISO 3602:1989
				National: Modified Hepburn, = BGN/PCGN
Kannada	Kannada	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
Kazakh	Cyrillic	Kazakhstan	—	BGN/PCGN 1979
Khmer	Khmer	Cambodia	1972 (II/10)	= BGN/PCGN 1972
				National: provisional 1996
Kirghiz	Cyrillic	Kyrgyzstan	—	BGN/PCGN 1979
Korean	Korean	(*General*)	—	McCune–Reischauer 1939, = BGN/PCGN
				Yale System 1967
				ISO/TR 11941:1996
		Democratic People's Republic of Korea (the)	—	National: 1992
		Republic of Korea (the)	—	National: 2000
Lao	Lao	Lao People's Democratic Republic (the)	—	National System (= BGN/PCGN 1966)
Macedonian Cyrillic	Cyrillic	The former Yugoslav Republic of Macedonia	1977 (III/11)	= BGN/PCGN 1981
Malayalam	Malayalam	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
Maldivian	Divehi (Thaana)	Maldives	—	National: 1987 (= BGN/PCGN 1988)
Marathi	Devanagari	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System

Mongolian	Cyrillic	Mongolia	—	BGN/PCGN 1964
	Mongolian	China	1977 (III/8)[2]	
		Mongolia	—	
Nepali	Devanagari	Nepal	1972 (II/11), 1977 (III/12)	
				BGN/PCGN 1964
Oriya	Oriya	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
Pashto	Perso-Arabic	Afghanistan	—	BGN/PCGN 1968
Persian	Perso-Arabic	Iran (Islamic Republic of)	1967 (I/13)	= BGN/PCGN 1958
Punjabi	Gurmukhi	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
Russian	Cyrillic	Belarus	1987 (V/18)	
		Kazakhstan	1987 (V/18)	
		Kyrgyzstan	1987 (V/18)	
		Russia	1987 (V/18)	
				BGN/PCGN 1947
				ISO 9:1995 (transliteration)
Serbo-Croatian	Cyrillic	Bosnia and Herzegovina	1977 (III/11)	
		Yugoslavia	1977 (III/11)	
Sinhalese	Sinhalese	Sri Lanka	—	
Tajik	Cyrillic	Tajikistan	—	BGN/PCGN 1994
Tamil	Tamil	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
		Singapore	1972 (II/11), 1977 (III/12)	
		Sri Lanka	1972 (II/11), 1977 (III/12)	
Telugu	Telugu	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
Thai	Thai	Thailand	1967 (I/14)	= BGN/PCGN 1970
				ISO 11940:1998 (transliteration)
Tibetan	Tibetan	China	1977 (III/8)[2]	
Tigrinya	Ethiopic	Eritrea	—	BGN/PCGN 1994
Uighur	Perso-Arabic	China	1977 (III/8)[2]	
Ukrainian	Cyrillic	Ukraine	—	National: 1996
				BGN/PCGN 1965

Urdu	Perso-Arabic	India	1972 (II/11), 1977 (III/12)	
				National: Hunterian System
		Pakistan	1972 (II/11), 1977 (III/12)	
				National: Hunterian System

Notes.

1. The system applies to Persian of which Dari is a variation.
2. Although the appropriate United Nations resolution does not mention these languages by name, it recognizes the Scheme for a Chinese Phonetic Alphabet (Pinyin) as China's official alphabet scheme and recommends that it be adopted as the international system for the romanization of Chinese geographical names. The Chinese Phonetic Alphabet (Pinyin) covers the direct transcription of Mongolian, Tibetan and Uighur scripts, as well as Chinese.

LANGUAGES/SCRIPTS COVERED BY SYSTEMS RECOMMENDED BY THE UNITED NATIONS

Amharic

The United Nations recommended romanization system was approved in 1967 (resolution I/17), based on the *Amharic to English Transliteration System* (2nd revised edition, August 1962) by the Imperial Ethiopian Mapping and Geography Institute. The table has been published in volume II of the conference report[1].

In all likelihood the UN-approved system is not being used in present-day Ethiopia. In international cartographic products mostly the BGN/PCGN 1967 system is used which differs from the former mainly by using diacritical marks differently. Often the same spellings are used without any diacritical marks.

Amharic uses the Ethiopic syllabic script whereby each syllable denotes a combination of a consonant and a vowel. It has 34 basic characters, each character having several different forms, usually called orders, according to the vowel with which the basic character is combined. The system is mainly reversible with the exception of some syllables of the 1st order which are pronounced in the same manner as the syllables of the 4th order, and some consonants.

Romanization

	1	2	3	4	5	6(A)	7	Combinations with W and Y[C]: (1)	(3)	(4)	(5)	(6)	(4)
1	U	U·	ጊ	ሃ	ዜ	U	ሆ			ኋ			
	ha[B]	hu	hi	ha	he	hị	ho			hwa			
2	ለ	ሉ	ሊ	ላ	ሌ	ል	ሎ			ሏ			
	le	lu	li	la	le	lị	lo			lwa			
3	ሐ	ሑ	ሒ	ሓ	ሔ	ሕ	ሖ						
	ha[B]	hu	hi	ha	he	hị	ho						
4	መ	ሙ	ሚ	ማ	ሜ	ም	ሞ			ሟ (ሟ)			ፙ
	me	mu	mi	ma	me	mị	mo			mwa			mya
5	ሠ	ሡ	ሢ	ሣ	ሤ	ሥ	ሦ						
	·se	su	si	sa	se	sị	so						
6	ረ	ሩ	ሪ	ራ	ሬ	ር	ሮ			ሯ			ፘ
	re	ru	ri	ra	re	rị	ro			rwa			rya
7	ሰ	ሱ	ሲ	ሳ	ሴ	ስ	ሶ			ሷ			
	se	su	si	sa	se	sị	so			swa			
8	ሸ	ሹ	ሺ	ሻ	ሼ	ሽ	ሾ			ሿ			
	she	shu	shi	sha	she	shị	sho			shwa			
9	ቀ	ቁ	ቂ	ቃ	ቄ	ቅ	ቆ	ቈ	ቊ	ቋ	(ቌ)	ቍ	
	k'ẹ	k'u	k'i	k'a	k'e	k'ị	k'o	k'we	k'wi	k'wa	k'we	k'wị	

---

[1] *Second United Nations Conference on the Standardization of Geographical Names*. London, 10-31 May 1972. Vol. II. *Technical papers*, pp. 165-168 (Table 2)

	1	2	3	4	5	6	7	8	9	10	11	12
10	be	bu	bi	ba	be	bi	bo			bwa		
11	te	tu	ti	ta	te	ti	to			twa		
12	che	chu	chi	cha	che	chi	cho			chwa		
13	ha[B]	hu	hi	ha	he	hi	ho	hwe	hwi	hwa	hwe	hwi
14	ne	nu	ni	na	ne	ni	no			nwa		
15	nye	nyu	nyi	nya	nye	nyi	nyo			nywa		
16	a[B]	u	i	a	e	i	o					
17	ke	ku	ki	ka	ke	ki	ko	kwe	kwi	kwa	kwe	kwi
18	he	hu	hi	ha	he	hi	ho					
19	we	wu	wi	wa	we	wi	wo					
20	a[B]	u	i	a	e	i	o					
21	ze	zu	zi	za	ze	zi	zo			zwa		
22	zhe	zhu	zhi	zha	zhe	zhi	zho			zhwa		
23	ye	yu	yi	ya	ye	yi	yo			ywa		
24	de	du	di	da	de	di	do			dwa		
25	je	ju	ji	ja	je	ji	jo			jwa		
26	ge	gu	gi	ga	ge	gi	go	gwe	gwi	gwa	gwe	gwi
27	t'e	t'u	t'i	t'a	t'e	t'i	t'o			t'wa		
28	ch'e	ch'u	ch'i	ch'a	ch'e	ch'i	ch'o			ch'wa		
29	p'e	p'u	p'i	p'a	p'e	p'i	p'o					
30	ts'e	ts'u	ts'i	ts'a	ts'e	ts'i	ts'o			ts'wa		
31	ts'e	ts'u	ts'i	ts'a	ts'e	ts'i	ts'o					
32	fe	fu	fi	fa	fe	fi	fo			fwa		fya

33	ፐ	ፑ	ፒ	ፓ	ፔ	ፕ	ፖ
	pe	pu	pi	pa	pe	pi	po
34 C	ቨ	ቩ	ቪ	ቫ	ቬ	ቭ	ቮ
	ve	vu	vi	va	ve	vi	vo

A The vowel of the sixth order (i) is eliminated in spelling except when the actual pronunciation requires it (e.g. not Menigisiti but Mengist).

B Sounds identical to fourth-order form.

C Row 34 and most combinations with W and Y are not present in the original table. These Amharic syllable characters have been taken from the BGN/PCGN 1967 table, preserving however the indication of vowels as it is done elsewhere in the table.

Other systems of romanization

The **BGN/PCGN 1967 System** is practically very close to the system above, except for the diacritical marks used with vowels:

	UN	BGN/PCGN
1st order	e	e
1st order	a	ā (ሀ hā, ሐ hā, ኀ hā, አ ā, ’ā, ዐ ‘ā)
3rd order	i	ī
5th order	e	ē
6th order	i	i

Other differences:

1. The vowel characters in row 16 should be romanized in the BGN/PCGN System ā, u, ī, a, ē, i, and o initially and ’ā, ’u, ’ī, ’a, ’ē, ’i, and ’o in all other positions.
2. The characters in row 20 should be romanized in the BGN/PCGN System ‘ā, ‘u, ‘ī, ‘a, ‘ē, ‘ or ‘i, and ‘o.
3. The syllables gwe, hwe, kwe, k'we of the UN system are treated as variations of the syllables go, ho, ko, k'o in the BGN/PCGN system.
4. For documentation purposes the following consonants may be romanized with diacritical marks in the BGN/PCGN system: ḥ (row 3), š (row 5), ḫ (row 13), ḥ (row 18), ts' (row 31).

Arabic

The United Nations recommended romanization system was approved 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].

Already in the UN resolution it was specifically pointed out that the system was recommended "for the romanization of the geographical names within those Arabic-speaking countries where this system is officially acknowledged". It cannot be definitely ascertained which of the Arabic-

---

[2] *Second United Nations Conference on the Standardization of Geographical Names.* London, 10-31 May 1972. Vol. II. *Technical papers*, p. 170.

speaking countries have adopted this system officially. Judging by the use of names in international cartographic products which rely mostly on national sources it appears that the UN system is more or less current in Iraq, Kuwait, the Libyan Arab Jamahiriya, Saudi Arabia, United Arab Emirates and Yemen, there and in some other countries the system is often used without diacritical marks. For the geographical names of the Syrian Arab Republic the international maps favour the UN system while the local usage seems to prefer a French-oriented romanization. Also in Egypt and Sudan there exist local romanization schemes or practices side by side with the UN system. 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.

Resolution 7 of the Seventh UN Conference on the Standardization of Geographical Names (1998) recommended that "the League of Arab States should, through its specialized structures, continue its efforts to organize a conference with a view to considering the difficulties encountered in applying the amended Beirut system of 1972 for the romanization of Arabic script, and submit, as soon as possible, a solution to the United Nations Group of Experts on Geographical Names".

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 proper pointing. One must also take into account dialectal and idiosyncratic deviations. The romanization is generally reversible though there are 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	2	3	4	5		1	2	3	4	5
1	ء				', - [A]	11	ر			ر	r	21	ف	ف	ف	ف	f
2	ا			ا	- [B]	12	ز			ز	z	22	ق	ق	ق	ق	q
3	ب	ب	ب	ب	b	13	س	س	س	س	s	23	ك	ك	ك	ك	k
4	ت	ت	ت	ت	t [C]	14	ش	ش	ش	ش	sh	24	ل	ل	ل	ل	l
5	ث	ث	ث	ث	th	15	ص	ص	ص	ص	ş	25	م	م	م	م	m
6	ج	ج	ج	ج	j	16	ض	ض	ض	ض	ḍ	26	ن	ن	ن	ن	n
7	ح	ح	ح	ح	ḥ	17	ط	ط	ط	ط	ṭ	27	ه	ه	ه	ه	h [C]
8	خ	خ	خ	خ	kh	18	ظ	ظ	ظ	ظ	ẓ	28	و			و	w
9	د			د	d	19	ع	ع	ع	ع	'	29	ي	ي	ي	ي	y
10	ذ			ذ	dh	20	غ	غ	غ	غ	gh						

[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	•	a	6	•	(A)	11	ٍ•	á
2	ٯ•	aw	7	١•	ā	12	•	a^n
3	ٯ•	ay	8	ٱ	ā	13	•	i^n
4	•	i	9	ٯ•	ī	14	•	u^n
5	•	u	10	ٯ•	ū	15	•	(B)

[A] Marks absence of the vowel.
[B] Marks doubling of the consonant.

Note. When the definite article *al* precedes a word beginning with one of the "sun letters" (*t, th, d, dh, r, z, s, sh, ṣ, ḍ, ṭ, ẓ, l, n*) the *l* of the definite article is assimilated with the first consonant of the word: الشارقة ash-Shāriqah.

Other systems of romanization

The **BGN/PCGN 1956 System** is identical to the UN system. The only difference lies in the treatment of articles. 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. البصره 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 first definite article in a name, e.g. Al Baṣrah, Ar Riyāḍ.

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

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

---

[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.

[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 (ī)	= â	ḥ	= ḥ	th	= t̲	
á	= ay	iⁿ	= í	uⁿ	= ú	
aⁿ	= á	j	= ǧ	y (ى)	= ẏ	
ḍ	= ḍ	kh	= ḫ	y (ي)	= y	
dh	= ḏ	ṣ	= ṣ	ẕ	= ẓ	
gh	= ġ	sh	= š			
h (ة)	= ẗ	ṭ	= ṭ			

The **Royal Jordanian Geographic Centre (RJGC) System**[4] is essentially the same as the amended Beirut system. The sub-macron is used instead of the cedilla.

UN	RJGC					
-ah (ة)	= -a	ḥ	= h̲	ṭ	= t̲	
ḍ	= ḏ	ṣ	= s̲	ẕ	= d̲h	

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

UN	SES	dh	= dh (z)	ṭ	= t̲
á	= a	ḥ	= h	th	= th (t)
ā	= â (a)	ī	= î	ū	= û
-ah (ة)	= -a	j	= g (j)	ẕ	= z̲ (d)
aw	= ô (au)	q	= q (k)	'	= '
ay	= ei (ai)	s	= s (c)		
ḍ	= ḍ	ṣ	= ṣ		

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).

A system that is used in **Lebanon** is mentioned in ISO 3166-2:1998 (Codes for the representation of names of countries and their subdivisions. Part 2: Country subdivision code): *Principles for romanization from Lebanese Arabic to Latin Characters* (National Ministry of Defence of the Lebanese Republic 1963). It is close to the I.G.N. 1973 System. In **Algeria** there is at present no official romanization system, the prospects of establishing such a system are being discussed at the National Institute of Cartography and Teledetection. In **Mauritania**, the romanized name

---

[4] *Activities in Jordan on the Standardization of Geographical Names.* UNGEGN, 18th Session. Geneva, 12-23 August 1996. WP. No. 86.

forms in the official maps edited since 1969 have been rendered in accordance with a simplified version of the I.G.N. system[5]. In **Morocco** the Ministry of Agriculture introduced in 1971 and 1972 an official romanization system for Arabic script which constituted a modification of the above-mentioned I.G.N. system. In **Tunisia** the Directorate of Topography and Cartography adopted in 1983 officially the amended Beirut system with minor modifications (e.g. adding a letter *g* to the table).

Assamese

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[6].

There is no evidence of the use of the system either in India or in international cartographic products.

Assamese (Asamīyā) uses an alphasyllabic script whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization table is unambiguous but the user would have to recognize many ligatures not given in the original table. The system is mostly reversible but there exist some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (ligatures vs. character sequences).

Romanization

I. Independent vowel characters

1 অ	a	3 ই	i	5 উ	u	7 ঋ	r̥	8 এ	e	10 ও	o
2 আ	ā	4 ঈ	ī	6 ঊ	ū			9 ঐ ·	ai	11 ঔ	au

II. Abbreviated vowel characters (• stands for any consonant character)

1 •	a	3 ি•	i	5 •ু	u[A]	7 •ৃ	r̥[C]	8 ে•	e	10 ে•া	o
2 •া	ā	4 •ী	ī	6 •ূ	ū[B]			9 ৈ•	ai	11 ে•ৗ	au

[A] Exceptions: গু gu; ৰু ru; শু shu; হু hu; ন্তু ntu; স্তু stu.

[B] Exception: ৰূ rū.

[C] Exception: হৃ hr̥.

---

[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] *Second United Nations Conference on the Standardization of Geographical Names*. London, 10-31 May 1972. Vol. II. *Technical papers*. United Nations. New York 1974, pp. 141-142; *Third United Nations Conference on the Standardization of Geographical Names*. Athens, 17 August - 7 September 1977. Vol. II, *Technical papers*, pp. 393 etc.

III. Other symbols (• stands for any consonant character)

1 •ৎ ṁ   2 • m̐   3 •ঃ ḥ   4 •্ (A)

A Pronunciation without a vowel; special form: ৎ t.

IV. Consonant characters

1 ক ka	6 চ cha	11 ট ṭa	16 ত ta	21 প pa	26 য ja[A]	30 শ sha
2 খ kha	7 ছ chha	12 ঠ ṭha	17 থ tha	22 ফ pha	27 র ra	31 ষ ṣha
3 গ ga	8 জ ja	13 ড ḍa[A]	18 দ da	23 ব ba	28 ল la	32 স sa
4 ঘ gha	9 ঝ jha	14 ঢ ḍha[A]	19 ধ dha	24 ভ bha	29 ব va	33 হ ha
5 ঙ ṅa	10 ঞ ña	15 ণ ṇa	20 ন na	25 ম ma		

A Dotted variants of the characters: ড় ṛa; ঢ় ṛha; য় ya.

V. Ligatures
See the section on the romanization of Bengali.

Other systems of romanization

For differences between the UN system and the **ISO transliteration standard ISO 15919: 2001** see the section on the romanization of Hindi.

Bengali

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[7].

There is no evidence of the use of the system either in Bangladesh, in India or in international cartographic products. The resolution IV/17 (1982) recommended association, inter alia, with Bangladesh, in carrying out further studies on the system.

Bengali (Bānglā) uses an alphasyllabic script whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization table is unambiguous but the user would have to recognize many ligatures not given in the original table. The system is mostly reversible but there exist some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (ligatures vs. character sequences).

---

[7] *Second United Nations Conference on the Standardization of Geographical Names.* London, 10-31 May 1972. Vol. II. *Technical papers.* United Nations. New York 1974, pp. 139-140; *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. II, *Technical papers,* pp. 393 etc.

Romanization

In the romanization system below character variations and the table of ligatures have been added.

I. Independent vowel characters

1 অ a	3 ই i	5 উ u	7 ঋ r̥	8 এ e	10 ও o
2 আ ā	4 ঈ ī	6 ঊ ū		9 ঐ ai	11 ঔ au

II. Abbreviated vowel characters (• stands for any consonant character)

1 • a	3 •ি i	5 •ু u[A]	7 •ৃ r̥[C]	8 ে• e	10 ে•া o
2 •া ā	4 •ী ī	6 •ূ ū[B]		9 ৈ• ai	11 ে•ৗ au

[A] Exceptions: গু gu; রু ru; শু shu; হু hu; ন্তু ntu; স্তু stu.
[B] Exception: রূ rū.
[C] Exception: হৃ hr̥.

III. Other symbols (• stands for any consonant character)

1 •ং ṁ	2 •ঁ m̐	3 •ঃ ḥ	4 •্ (A)

[A] Pronunciation without a vowel; special form: ৎ t.

IV. Consonant characters

1 ক ka	6 চ cha	11 ট ṭa	16 ত ta	21 প pa	26 য ja[A]	29 শ sha
2 খ kha	7 ছ chha	12 ঠ ṭha	17 থ tha	22 ফ pha	27 র ra	30 ষ ṣha
3 গ ga	8 জ ja	13 ড ḍa[A]	18 দ da	23 ব ba	28 ল la	31 স sa
4 ঘ gha	9 ঝ jha	14 ঢ ḍha[A]	19 ধ dha	24 ভ bha		32 হ ha
5 ঙ ṅa	10 ঞ ña	15 ণ ṇa	20 ন na	25 ম ma		

[A] Dotted variants of the characters: ড় ṛa; ঢ় ṛha; য় ya.

V. Ligatures

Adscript forms of some consonants
        Example
• r-: র্ক rka
ৄ -r: গ্র gra
•্য -y: ত্য tya

Other ligatures (the list is not complete)

ক্ক	ক্ট	ক্ত	ক্ন	ক্ম	ক্র	ক্ল	ক্ব	ক্ষ	ক্ষ্ণ	ক্ষ্ম	ক্ষ্ব
kka	kṭa	kta	kna	kma	kra	kla	kva	kṣha	kṣhna	kṣhma	kṣhva
ক্স	গ্গ	গ্দ	গ্ধ	গ্ন	গ্ম	গ্র	গ্ল	ঘ্র	ঙ্ক	ঙ্গ	চ্চ
ksa	gga	gda	gdha	gna	gma	gra	gla	ghra	ṅka	ṅga	chcha
চ্ছ	চ্ছ্ব	চ্ঞ	জ্জ	জ্ব	জ্ঝ	জ্ঞ	জ্ব	ঞ্চ	ঞ্ছ	ঞ্জ	ঞ্ঝ
chchha	chchhva	chña	jja	jjva	jjha	jña	jva	ñcha	ñchha	ñja	ñjha

ṭṭa	ḍḍa	ṇṭa	ṇṭha	ṇḍa	tta	ttva	ttha	tna	tma	tra	tla
tva	dda	ddva	ddha	ddhva	dna	dva	dbha	dma	dra	dla	dhra
ntha	nḍa	nka	nta	ntra	ntha	nda	ndra	ndha	nna	nma	nva
pna	pta	ppa	pra	pla	phra	bja	bda	bdha	bba	bra	bhra
mpa	mba	mbha	mbhra	mma	mra	mla	lka	lta	lḍa	lma	lla
shcha	shchha	shta	shna	shma	shra	shla	shva	ṣhka	ṣhṭa	ṣhṭra	ṣhṭha
ṣhña	ṣhpa	ṣhpha	ska	skra	skha	sta	sna	sma	sra	sva	hna
hma	hra	hla									

Other systems of romanization

For differences between the UN system and the **ISO transliteration standard ISO 15919: 2001** see the section on the romanization of Hindi.

Bulgarian

The United Nations recommended romanization system was approved in 1977 (resolution III/10), based on the system produced by the Council of Orthography and Transcription of Geographical Names, Sofia (1972). The table was published as an annex to the resolution[8]. The 1977 resolution adopted one of the two systems approved by the UN conference in 1972 (II/5).

The system is used in Bulgaria and increasingly in international cartographic products. In Bulgaria there are also some other systems in use, e.g. the Bulgarian Council of Ministers had promulgated on 2 April, 1999 in the Government Gazette the decree No. 61 on personal identity documents with an annex on the romanization of names. It has been stated by the Council of Orthography and Transcription of Geographical Names that at the moment there are no plans to submit any new system for the rendering of geographical names and thus replace the approved United Nations system.

Bulgarian uses the Cyrillic script which is alphabetic. The romanization table is unambiguous and can be applied automatically. The system is on the whole reversible, although one should know the spelling rules. For example, the romanized *j* will correspond to ь if used after a consonant, in other cases it represents the Cyrillic *ŭ*. As exception, when *j* is followed by *a* or *u*, the combinations *ja* and *ju* should be converted to the Cyrillic as я and ю, respectively. These romanizations could be ambiguous, if there should exist character sequences *ŭa* and *ŭy* which

---

[8] *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. I, *Report of the Conference*, p. 28

seems highly unlikely.

Romanization

1	А а	a	9	И и	i	17	Р р	r	25	Ш ш š
2	Б б	b	10	Й й	j	18	С с	s	26	Щ щ št
3	В в	v	11	К к	k	19	Т т	t	27	Ъ ъ ă
4	Г г	g	12	Л л	l	20	У у	u	28	Ь ь j
5	Д д	d	13	М м	m	21	Ф ф	f	29	Ю ю ju
6	Е е	e	14	Н н	n	22	Х х	h	30	Я я ja
7	Ж ж	ž	15	О о	o	23	Ц ц	c		
8	З з	z	16	П п	p	24	Ч ч	č		

Note. Cursive forms of some characters might be formed differently: *Аа Бб Вв Гг Дд Ее Жж Зз Ии Йй Кк Лл Мм Нн Оо Пп Рр Сс Тт Уу Фф Хх Цц Чч Шш Щщ Ъъ Ьь Юю Яя.*

Other systems of romanization

The system adopted by the **Bulgarian Council of Ministers** in April 1999 for the conversion of names in passports, provides for the romanization, as a single block, of the following characters differently from the UN system (the Cyrillic character is followed in parentheses by the romanization according to the UN system):

ж (ž)	zh		ц (c)	c		щ (št)	sht		ю (ju)	yu
й (j)	y		ч (č)	ch		ъ (ă)	a		я (ja)	ya
х (h)	h		ш (š)	sh		ь (j)	y			

The **BGN/PCGN 1952 System** provides for the romanization, as a single block, of the following characters differently from the UN system (the Cyrillic character is followed in parentheses by the romanization according to the UN system):

ж (ž)	zh		ц (c)	ts		щ (št)	sht		ю (ju)	yu
й (j)	y		ч (č)	ch		ъ (ă)	ŭ		я (ja)	ya
х (h)	kh		ш (š)	sh		ь (j)	'			

There are also notes concerning the romanization of characters abolished after the orthography reform of 1945.

Chinese

The United Nations conference approved the Scheme for a Chinese Phonetic Alphabet (Pinyin) in 1977 (III/8) for the romanization of Chinese geographical names.

The system is used in China and in international cartographic products almost universally.

Chinese uses a logographic script whereby each character represents a word or syllable. The relationship between the characters and their pronunciation is complex and therefore the phonetic notation of a Chinese syllable cannot be unambiguously converted to its written form.

Romanization

The approved readings of Chinese characters can be obtained from modern authoritative dictionaries, such as *Xinhua Zidian*.

Other systems of romanization

The **Modified Wade-Giles transcription** (1892) was previously one of the most well-known systems of rendering Chinese syllables, especially in English-language texts. The syllables of the Wade-Giles (WG) system may be transformed into Pinyin syllables as given in the table below[9] but bearing in mind that often it is not correct to change Wade-Giles into Pinyin without reference to current Chinese character sources.

WG	UN						
Syllable beginnings:		sz	= s	en	= an	u	= u, ou[F]
ch	= zh, j[A]	t	= d	ên	= en	ŭ	= i
ch'	= ch, q[A]	t'	= t	êng	= eng	ü	= u, ü[F]
hs	= x	ts	= z	êrh	= er	üan	= uan
j	= r	ts'	= c	i	= i, yi[B]	üeh	= ue, üe[F]
k	= g	tz	= z	ieh	= ie	uei	= ui
k'	= k	tz'	= c	ien	= ian	ün	= un
p	= b	Syllable endings:		ih	= i	ung	= ong
p'	= p	ê	= e	iung	= iong		
ss	= s	eh	= e	o	= o[C], uo[D], e[E]		

[A] Before *i* (except the ending *-ih*) and *ü*.

[B] Word-initially.

[C] After *p, p', m, f, w, y*.

[D] After *t, t', n, l, ch, ch', j, ts, ts', s*.

[E] After *k, k', h* and syllable-initially.

[F] After *n, l*.

The Pinyin alphabet has also been adapted to the romanization of names from three minority languages in China: **Mongolian, Tibetan** and **Uighur**.

Greek

The United Nations recommended system was approved in 1987 (resolution V/19), based on the ELOT 743 conversion system of the Greek Standardization Organization. The table was published as an annex to the resolution[10].

---

[9] Adapted from Klaus Kaden, *Die wichtigsten Transkriptionssysteme für die chinesische Sprache*. VEB Verlag Enzyklopädie, Leipzig 1975, S. 145.

[10] *Fifth United Nations Conference on the Standardization of Geographical Names*. Montreal, 18-31 August 1987. Vol. I. *Report of the Conference*, pp. 42-43.

The system is used in Greece and Cyprus, also increasingly in international cartographic products. Many systems of romanization adapted to transcribing names from older forms of Greek continue to be used in other countries. The 1987 resolution acknowledged that "a suitable transitional period will be required before the system can be fully implemented".

The romanization table is unambiguous and simple to use. There are two versions of romanization which complement each other: transcription and transliteration. The latter differs from the former only by adding a sub-macron to certain ambiguous romanization equivalents.

Romanization

The table below contains the transcription version of the romanization system. See note 1 for the transliteration equivalents.

1	A α	a	5	E ε	e	11	Λ λ	l	17	P ρ	r
	αυ	av[A], af[B]		ευ	ev[A], ef[B]	12	M μ	m	18	Σ σ ς	s
2	B β	v	6	Z ζ	z		μπ	b[C], mp[D]	19	T τ	t
3	Γ γ	g	7	H η	i	13	N ν	n	20	Y υ	y
	γγ	ng		ηυ	iv[A], if[B]	14	Ξ ξ	x	21	Φ φ	f
	γξ	nx	8	Θ θ	th	15	O o	o	22	X χ	ch
	γχ	nch	9	I ι	i		ου	ou	23	Ψ ψ	ps
4	Δ δ	d	10	K κ	k	16	Π π	p	24	Ω ω	o

[A] Before the consonants β, γ, δ, ζ, λ, μ, ν, ρ and all vowels.

[B] Before the consonants θ, κ, ξ, π, σ, τ, φ, χ, ψ and at the end of the word.

[C] At the beginning of the word and at the end of the word.

[D] In the middle of the word.

Notes

1. In order to ensure reversibility a sub-macron may be added to the following characters and character combinations: 1 αυ a*v*/a*f*, 3 γγ *ng*, γξ *nx*, γχ *nch*, 5 ευ e*v*/e*f*, 7 η *i*, ηυ i*v*/i*f*, 24 ω *o*.
2. Some examples of the ELOT 743 table contain also combinations in parentheses αι, άι, αϊ, γκ, ει, έι, εϊ, ντ, οι, όι, οϊ, υι which are converted according to the conversion rules of each character they consist of, resp. *ai, ái, aï, gk, ei, éi, eï, nt, oi, ói, oï, yi*.
3. The combinations αυ, ευ, ηυ, ου are converted according to the conversion rules of each character when the vowel before υ has an accent or when υ has the diaeresis sign.
4. When the character combinations αυ, ευ, ηυ are stressed in Greek, in the Roman script the accent is set on the vowel (*áv, áf, év, éf, ív, íf*).

Other systems of romanization

The **BGN/PCGN 1962 System** that continued to be used until recently, provided for the romanization, as a single block, of the following characters or character combinations differently from the UN system (the Greek character is followed in parentheses by the romanization according to the UN system):

αυ (av/af) av	δ (d) dh, d^D	μπ (b/mp) b^B, mb^C χ (ch) kh	
γ (g) g, y^A	ει (ei) i^E	ντ (nt) d^B, nd^C, nt^F	
γχ (nch) nkh	ευ (ev/ef) ev	υ (y) i	
γκ (gk) g^B, ng^C	ηυ (iv/if) iv	υι (yi) i	

^{A} Before αι, ε, ει, η, ι, οι, υ and υι.
^{B} Initially.
^{C} Medially.
^{D} Between v and ρ.
^{E} But εϊ = eï in both systems.
^{F} In the combination ντζ.

The transliteration **ISO 843:1997** which is also based on ELOT 743 presents the table of type 1 (transliteration of Greek characters into Latin characters) whereby each Greek character has only one Latin equivalent. The following consonants are romanized slightly differently from the UN-approved system:

H η   Ī ī or Ῑ ῑ
Ω ω   Ō ō or Ō̄ ō̄

Combinations of two or more characters are transliterated according to the provisions laid down for each independent character. The only exception to this rule for transliteration is the conversion of the Greek double vowels AY, αυ, EY, ευ, OY, ου which are mapped into Roman script as AU, au, EU, eu, OU, ou respectively.

The transliteration version of the system approved by the UN conference in 1987 is referred to in the standard as *reversible transcription* (type 2) and is given as informative Annex B.

Gujarati

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[11].

There is no evidence of the use of the system either in India or in international cartographic products.

Gujarati (Gujarātī) uses an alphasyllabic script whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization table is unambiguous but the user would have to recognize many ligatures not given in the original table. For two vowel letters (e/ĕ, o/ŏ) there is uncertainty as to the conditions in which either of the two variants is to be used. The system is mostly reversible but there exist

---

[11] *Second United Nations Conference on the Standardization of Geographical Names.* London, 10-31 May 1972. Vol. II. *Technical papers.* United Nations. New York 1974, pp. 134-135; *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. II, *Technical papers*, pp. 393 etc.

some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (ligatures vs. character sequences).

Romanization

In the romanization system below the table of ligatures has been added.

I. Independent vowel characters

1 અ	a	3 ઇ	i	5 ઉ	u	7 ઋ	r̥	8 એ	e, ĕ	10 ઓ	o, ŏ
2 આ	ā	4 ઈ	ī	6 ઊ	ū			9 ઐ	ai	11 ઔ	au

II. Abbreviated vowel characters (• stands for any consonant character)

1 •	a	3 • િ	i	5 •ુ	u	7 •ૃ	r̥	8 •ે	e, ĕ	10 •ો	o, ŏ
2 •ા	ā	4 •ી	ī[A]	6 •ૂ	ū[B]			9 •ૈ	ai	11 •ૌ	au

[A] Exception: જી jī.
[B] Exception: રૂ rū.

III. Other symbols
(These symbols were missing in the original table.)

1 •ં	ṁ	2 •ઁ	m̐	3 •ઃ	ḥ	4 •્	(A)

[A] Pronunciation without a vowel.

IV. Consonant characters

1 ક	ka	6 ચ	cha	11 ટ	ṭa	16 ત	ta	21 પ	pa	26 ય	ya	30 શ	sha
2 ખ	kha	7 છ	chha	12 ઠ	ṭha	17 થ	tha	22 ફ	fa	27 ર	ra	31 ષ	ṣha
3 ગ	ga	8 જ	ja	13 ડ	ḍa	18 દ	da	23 બ	ba	28 લ	la	32 સ	sa
4 ઘ	gha	9 ઝ	jha	14 ઢ	ḍha	19 ધ	dha	24 ભ	bha	29 વ	va	33 હ	ha
5 ઙ	ṅa	10 ઞ	ña	15 ણ	ṇa	20 ન	na	25 મ	ma			34 ળ	ḷa

V. Consonant ligatures

Adscript forms of some consonants
Examples

ર	r-	ર્ક rka, ર્મ rma
્	-r	ક્ર kra, બ્ર bra, દ્ર dra

Other ligatures (the list is not complete)

kka	kra	kṣha	khta	khra	gra	ghra	ṅka	ṅkha	ṅga	ṅgha	ṅma
chra	jja	jña	jra	jhra	ṭṭa	ṭṭha	ṭhṭha	ḍḍa	ḍḍha	ḍhḍha	tta
tna	tra	thra	dga	dda	ddha	dbha	dma	dya	dra	dva	dhra

nna	nra	pta	pra	phra	bra	bhra	mra	yra	vra	shna	shra
shva	shvra	ṣhṭa	ṣhṭha	sra	hṇa	hna	hma	hya	hla	hva	

Other systems of romanization

For differences between the UN system and the **ISO transliteration standard ISO 15919: 2001** see the section on the romanization of Hindi.

Hebrew

The United Nations recommended system was approved in 1977 (resolution III/13), based on the official system (1957) proposed by the Academy for the Hebrew Language and used by the Survey of Israel. The table was published as an annex to the resolution[12]. The 1977 resolution was the elaboration of an earlier resolution (II/9) adopted in 1972.

The system is used in Israel and in most international cartographic products. In recent years there have been discussions in Israel on whether to modify the official romanization system but no decision has yet been taken[13].

Hebrew is written from right to left. The correct romanization of the Hebrew script presumes the presence of fully pointed text, i.e. where all vowels are marked. These marks are usually omitted in everyday writing though a method of denoting vowels with certain consonant characters (*ketiv malé*) is gaining ground also in the rendering of geographical names. The romanization is not fully reversible, e.g. *t* denotes two different consonant characters and vowels (e.g. *o*) can have more than one way of writing in the Hebrew script.

Romanization

In the romanization table below five consonants have a special final form which is shown in addition to the main character after a comma. Superscript numbers refer to the notes at the end.

1	א	', (-)[2,9]	8	ז	z	15	מ, ם	m	22	ק	q
2	ב	b	9	ח	ḥ	16	נ, ן	n	23	ר	r
3	ב	v	10	ט	t	17	ס	s	24	שׁ	sh
4	ג (ג)[4]	g	11	י	y	18	ע	ʻ [8,9]	25	שׂ	s
5	ד (ד)[4]	d	12	כ, ך	k	19	פ	p	26	ת (ת)[4]	t
6	ה	h[7]	13	כ[3], ך ך	kh	20	פ[3], ף	f			
7	ו	w	14	ל	l	21	צ, ץ	ẓ			

---

[12] *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. I. *Report of the Conference*, pp. 29-30, 32-33.

[13] An overview is given by N. Kadmon in *Considerations for and against the revision of a romanization system – the case of Hebrew.* Seventh United Nations Conference on the Standardization of Geographical Names. New York, 13-22 January 1998. Document E/CONF.91/L.11.

Vowels (• stands for any consonant character)

1	•	a	5	•̤	e	9	•	i	13	�ו	o
2	•̱	a	6	•	e, é[5]	10	ٔ•	i	14	•̣	u
3	•	a, o	7	ٔ•	e	11	•̤	o	15	�ו	u
4	•	e	8	•	e[6]	12	•	o			

Notes

1. The addition of a dot (*dagesh ḥazaq*) within a consonant, except as shown in the list, doubles its value and is represented by doubling the respective letter (with the exception of the digraph *sh*).
2. The sign ' (*alef*) slightly separates two consecutive vowels or a consonant and a vowel; at the beginning of a word it is omitted in transliteration.
3. Does not occur at the beginning of a name.
4. At the beginning of a name or of a syllable after *sheva naḥ* these letters carry a dot (*dagesh qal*).
5. In some maps the *zere* (•̤), if accentuated, is transliterated é.
6. The *sheva* (•) is of two kinds: *sheva naḥ*, which is omitted in transliteration, and *sheva na'*, which occurs at the beginning of a word or syllable.
7. Final *h* (ה), unless vocalized, is omitted in transliteration.
8. The sign ' (*'ayin*) slightly separates two consecutive vowels or a consonant and a vowel; it is also retained at the beginning and end of a word.
9. In practical application א (') and ע (') are transliterated by the same symbol.

Hindi

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[14].

There is no evidence of the use of the system either in India or in international cartographic products. It was stated in 1987 that the appropriate resolution had not been implemented in India and the Hunterian system was still in use in large-scale mapping[15].

Hindi uses the alphasyllabic script Devanāgarī whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization table is unambiguous but the user would have to recognize many ligatures not given in the original table (only three are given). The system is mostly reversible but there may exist

---

[14] *Second United Nations Conference on the Standardization of Geographical Names.* London, 10-31 May 1972. Vol. II. *Technical papers.* United Nations. New York 1974, pp. 119-131; *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. II, *Technical papers*, pp. 393 etc.

[15] *Fifth United Nations Conference on the Standardization of Geographical Names.* Montreal, 18-31 August 1987. Vol. I. *Report of the Conference*, p. 30.

some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (ligatures vs. character sequences).

Romanization

In the romanization system below the table of ligatures has been added.

I. Independent vowel characters

1 अ a	3 इ i	5 उ u	7 ऋ ṛ	9 ऌ ḷ	10 ए e	12 ओ o
2 आ ā	4 ई ī	6 ऊ ū	8 ॠ ṝ		11 ऐ ai	13 औ au

II. Abbreviated vowel characters (• stands for any consonant character)

1 • a	3 ि i	5 ु u^B	7 ृ ṛ^B	8 े e	10 ो o
2 ा ā^A	4 ी ī	6 ू ū^B	9 ै ai	11 ौ au

^{A} With an additional mark: •ँ ằ.
^{B} Exceptions: रु ru; रू rū; हृ hṛ.

III. Other symbols

1 • ṁ	2 ँ m̐	3 •ः ḥ	4 ् $^{(A)}$

^{A} Pronunciation without a vowel: क् k.

IV. Consonant characters

1 क ka^A	6 च cha	11 ट ṭa	16 त ta	21 प pa	26 य ya	30 श sha
2 ख kha^A	7 छ chha	12 ठ ṭha	17 थ tha	22 फ pha^A	27 र ra	31 ष ṣha
3 ग ga^A	8 ज ja^A	13 ड ḍa^A	18 द da	23 ब ba	28 ल la	32 स sa
4 घ gha	9 झ (ऴ)jha	14 ढ ḍha^A	19 ध dha	24 भ bha	29 व va	33 ह ha
5 ङ ṅa	10 ञ ña	15 ण ṇa	20 न na	25 म ma		

^{A} Dotted variants of the characters: क़ qa; ख़ ḵha; ग़ g̱a; ज़ za; ड़ ṛa; ढ़ ṛha; फ़ fa.

V. Ligatures
Adscript form:

		Examples
्	r-	र्त rta, र्द्व rdva

Other ligatures (the list is not complete)

क्क	क्ख	क्त	क्त्व	क्य	क्र	क्ल	क्व	क्ष	क्ष्ण	क्ष्म	क्ष्य	क्ष्व
kka	kkha	kta	ktva	kya	kra	kla	kva	kṣha	kṣhṇa	kṣhma	kṣhya	kṣhva
क्स	ख्य	क्ष्श	ग्द	ग्ध	ग्न	ग्भ	ग्म	ग्य	ग्र	ग्ल	घ्न	घ्य
ksa	khya	ḵhsha	gda	gdha	gna	gbha	gma	gya	gra	gla	ghna	ghya
घ्र	ङ्क	ङ्ख	ङ्ग	ङ्घ	च्च	च्छ	च्छ्र	च्य	च्र	ज्ज	ज्झ	ज्ञ
ghra	ṅka	ṅkha	ṅga	ṅgha	chcha	chchha	chchhra	chya	chra	jja	jjha	jña
ज्य	ज्र	ज्व	ञ्च (ऄ्च)	ञ्छ	ञ्ज (ऄ्ज)	ट्ट	ट्ठ	ठ्ठ	ड्ड	ड्ढ	ढ्ढ	ण्ट
jya	jra	jva	ñcha	ñchha	ñja	ṭṭa	ṭṭha	ṭhṭha	ḍḍa	ḍḍha	ḍhḍha	ṇṭa

ण्ठ	ण्ड	ण्य	त्क	त्त	त्त्य	त्त्व	त्थ	त्न	त्प	त्म	त्म्य	त्य
ṇtha	ṇḍa	ṇya	tka	tta	ttya	ttva	ttha	tna	tpa	tma	tmya	tya
त्र	त्र्य	त्व	त्स	त्स्न	त्स्य	थ्य	थ्र	द्ग	द्द	द्ध	द्भ	द्म
tra	trya	tva	tsa	tsna	tsya	thya	thra	dga	dda	ddha	dbha	dma
द्य	द्र	द्व	ध्न	ध्म	ध्य	ध्र	ध्व	न्त	न्त्य	न्त्र	न्थ	न्द
dya	dra	dva	dhna	dhma	dhya	dhra	dhva	nta	ntya	ntra	ntha	nda
न्द्र	न्ध	न्ध्य	न्ध्र	न्न	न्म	न्य	न्र	न्व	न्स	न्ह	प्त	प्त्य
ndra	ndha	ndhya	ndhra	nna	nma	nya	nra	nva	nsa	nha	pta	ptya
प्न	प्य	प्र	प्ल	प्स	फ्य	फ्ल	ब्द	ब्न	ब्य	ब्र	भ्य	भ्र
pna	pya	pra	pla	psa	phya	phla	bda	bna	bya	bra	bhya	bhra
म्न	म्ब	म्भ	म्म	म्य	म्र	म्ल	म्व	म्ह	य्य	य्र	ल्क	ल्प
mna	mba	mbha	mma	mya	mra	mla	mva	mha	yya	yra	lka	lpa
ल्म	ल्य	ल्ल	ल्व	ल्स	व्य	व्र	श्च	श्च्य	श्छ	श्न	श्य	श्र
lma	lya	lla	lva	lsa	vya	vra	shcha	shchya	shchha	shna	shya	shra
श्ल	श्व (श्व)	ष्क	ष्क्र	ष्ट (ष्ट)	ष्ट्र	ष्ठ	ष्ण	ष्प	ष्प्र	ष्म	ष्य	ष्व
shla	shva	ṣhka	ṣhkra	ṣhṭa	ṣhṭra	ṣhṭha	ṣhṇa	ṣhpa	ṣhpra	ṣhma	ṣhya	ṣhva
स्क	स्ख	स्त	स्त्र	स्थ	स्न	स्प	स्फ	स्म	स्म्य	स्य	स्र	स्व
ska	skha	sta	stra	stha	sna	spa	spha	sma	smya	sya	sra	sva
ह्न	ह्म	ह्य	ह्र	ह्ल	ह्व							
hna	hma	hya	hra	hla	hva							

Notes

1. It is recommended that the vowel अ (*a*) should always be romanized except when it ends a name. If a name ends with a consonant, the consonant should carry a sub-macron. Such cases, however, will be very rare. For example, कानपुर *Kānapur* (not *Kānapura*), जगत् *Jagat̲*.
2. If each letter of a digraph or any two parts of a trigraph has a distinct independent sound then it should be indicated by a hyphen, thus *d-h*.

Other systems of romanization

The **Hunterian system** is the actually used national system of romanization in India.

UN	Hunterian							
ā	= ā, aᴬ	ṁ	= n, m	sh	= s, sh	v	= w, vᴮ
ḍ	= d	ṅ	= ng	ṣh	= sh		
ḍh	= dh	ṇ	= n	t	= t		
ī	= ī, iᴬ	ñ	= ny	ṭh	= th		

ᴬ Word-finally.
ᴮ Exceptionally, e.g. before *i*.

The transliteration standard for Indian languages **ISO 15919: 2001** provides different transliteration equivalents to the characters of the Indian languages. These are:

UN	ISO						
ă	= ĕ	e	= ē, eᴬ	i̱	= yᶜ	ḻ	= ḻ ᴮ
â	= ŏ	ĕ	= e	k̲h	= k̲h	m̐	= m̐, ~, ñᴱ
ch	= c	g̱	= ġ	ḷ	= ḷ	o	= ō, oᴬ
chh	= ch	ḥ	= ḥ, k̲ᴮ	ḹ	= ḹ	ŏ	= o

ṙ	= ṛ		sh	= ś		y	= ẏ [CF]
ṛ	= ṛ		ṣh	= ṣ		y̱	= y [F]
r̄	= r̄		ṭ	= ṭ [D]			

[A] Non-uniform vowels option: if the only scripts in view are without short *e*, *o*, then long *e*, *o* may be transliterated *e*, *o*, respectively.

[B] Used in Tamil.

[C] Used in Bengali.

[D] Used in Malayalam.

[E] Used in Telugu.

[F] Used in Oriya.

Kannada

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[16].

There is no evidence of the use of the system either in India or in international cartographic products.

Kannada uses an alphasyllabic script whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization table is unambiguous. The system is mostly reversible but there may exist some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (combinations with subscript consonants vs. character sequences).

Romanization

I. Independent vowel characters

1 ಅ	a	3 ಇ	i	5 ಉ	u	8 ಎ	ĕ	11 ಒ	ŏ		
2 ಆ	ā	4 ಈ	ī	6 ಊ	ū	9 ಏ	e	12 ಓ	o		
				7 ಋ	ṛ	10 ಐ	ai	13 ಔ	au		

[16] *Second United Nations Conference on the Standardization of Geographical Names.* London, 10-31 May 1972. Vol. II. *Technical papers.* United Nations. New York 1974, pp. 147-148; *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. II, *Technical papers,* pp. 393 etc.

II. Abbreviated vowel characters and other symbols (• stands for any consonant character)

1 •ᴶ	aᴬ	3 ᴐ	iᴮ	5 ಲ	uᶜ	8 •ᴾ	ĕ	11 •ᴶ಄	ŏ	14 ಂ	ṁ
2 ᴇᴐ	ā	4 ᴐᴇ	ī	6 •಄	ūᴰ	9 •ᴾᴇ	e	12 •ᴶᴐᴇ	o	15 ಃ	ḥ
				7 ಋ	ṛ	10 •ᴶ	ai	13 •ᴐ	au	16 ᴇ	(E)

ᴬ Some consonants do not have the diacritic, e.g. ಖ kha, ಜ ja, ಣ ṇa, ಬ ba, ಲ la.

ᴮ Variations of the diacritic, e.g. ಗಿ gi, ದಿ di.

ᶜ Variations of the diacritic: ಪು pu, ಫು phu, ವು vu.

ᴰ Variations of the diacritic: ಪೂ pū, ಫೂ phū, ವೂ vū.

ᴱ Pronunciation without a vowel: ಕ್ k.

III. Consonant characters

1 ಕ	ka	6 ಛ	cha	11 ಟ	ṭa	16 ತ	ta	21 ಪ	pa	26 ಯ	ya	30 ಶ	sha
2 ಖ	kha	7 ಛ	chha	12 ಠ	ṭha	17 ಥ	tha	22 ಫ	pha	27 ರ	ra	31 ಷ	ṣha
3 ಗ	ga	8 ಜ	ja	13 ಡ	ḍa	18 ದ	da	23 ಬ	ba	28 ಲ	la	32 ಸ	sa
4 ಘ	gha	9 ಝ	jha	14 ಢ	ḍha	19 ಧ	dha	24 ಭ	bha	29 ವ	va	33 ಹ	ha
5 ಙ	ṅa	10 ಞ	ña	15 ಣ	ṇa	20 ನ	na	25 ಮ	ma			34 ಳ	ḷa

IV. Subscript consonant characters

1 •ᴶ	kᴬ	6 •ᴶ	ch	11 •ᴶ	ṭᴮ	16 •ᴶ	tᶜ	21 •ᴶ	pᴰ	26 •ᴶ	y	30 •ᴶ	sh
2 •ᴶ	kh	7 •ᴶ	chh	12 •ᴶ	ṭh	17 •ᴶ	th	22 •ᴶ	ph	27 •ᴶ	r	31 •ᴶ	ṣh
3 •ᴶ	g	8 •ᴶ	j	13 •ᴶ	ḍ	18 •ᴶ	d	23 •ᴶ	b	28 •ᴶ	l	32 •ᴶ	s
4 •ᴶ	gh	9 •ᴶ	jh	14 •ᴶ	ḍh	19 •ᴶ	dh	24 •ᴶ	bh	29 •ᴶ	v	33 •ᴶ	h
5 •ᴶ	ṅ	10 •ᴶ	ñ	15 •ᴶ	ṇ	20 •ᴶ	n	25 •ᴶ	m			34 •ᴶ	ḷ

ᴬ Subscript combinations: •ᴶ kr, •ᴶ kṣh.

ᴮ Subscript combination: •ᴶ tr.

ᶜ Subscript combinations: •ᴶ tr, •ᴶ ty.

ᴰ Subscript combination: •ᴶ pr.

Combinations with r as the first component are written by adding ಼ after the second consonant: ಗ಼ rga.

Other systems of romanization

For differences between the UN system and the **ISO transliteration standard ISO 15919: 2001** see the section on the romanization of Hindi.

Khmer

The United Nations recommended system was approved in 1972 (resolution II/10), based on the system used by the BGN/PCGN (1972), this being a modified version of the Service Géographique Khmère (SGK) 1959 system. The table and notes on its usage were published in volume II of the conference report[17].

The system is used in many international cartographic products. In 1994 a Gazetteer of Cambodia was produced using this system with some proposed modifications. However, the standard ISO 3166-2:1998 (Codes for the representation of names of countries and their subdivisions. Part 2: Country subdivision code) gives the names of Cambodian provinces in a "provisional romanization by Geographic Department of the Office of the Council of Ministers [of Cambodia] 1996" which is a considerable departure from the UN adopted system.

Khmer uses an alphasyllabic script whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization system is complicated by many additional rules. In Khmer writing word division is not ordinarily indicated and Khmer diacritical marks are often omitted. The romanization is generally not reversible to its original script form.

Romanization

I. Consonant characters

1	kâ	6	châ	11	dâ	16	tâ	21	bâ[B]	26	yô	30	sâ
2	khâ	7	chhâ	12	thâ	17	thâ	22	phâ	27	rô	31	hâ
3	kô	8	chô	13	dô	18	tô	23	pô	28	lô	32	lâ
4	khô	9	chhô	14	thô	19	thô	24	phô	29	vô	33	'â[C]
5	ngô	10	nhô[A]	15	nâ	20	nô	25	mô				

[A] If used with a subscript character the lower element of the character is omitted: ញ *nhchô*.

[B] If used with a subscript character the romanization of the character will be *p* (see note 4).

[C] See note 5.

---

[17] *Second United Nations Conference on the Standardization of Geographical Names*. London, 10-31 May 1972. Vol. II. *Technical papers*. United Nations. New York 1974, pp. 163-164.

II. Subscript consonant characters (see also note 3)

Character numbers correspond to those in Table I.

1	k	6	ch	11	d	16	t	21	b	26	y	30	s
2	kh	7	chh	12	th	17	th	22	ph	27	r	31	h
3	k	8	ch	13	d	18	t	23	p	28	l	33	’
4	kh	9	chh	14	th	19	th	24	ph	29	v		
5	ng	10	nh[A]	15	n	20	n	25	m				

[A] With the same base character ⟡ is used (⟡), in all other cases only the lower element is used as subscript, e.g. ⟡ khnhâ.

III. Independent vowel characters

1	ĕ	3	ŏ,ŭ[A]	5	rœ	7	lœ̆	9	ê	11	aô[B]
2	ei	4	âu	6	rœ	8	lœ	10	ai	12	au

[A] See note 10.

[B] Variation: ⟡.

IV. Vocalic nuclei (• stands for any consonant character)

Where variants in romanization separated by a dash are given the first is to be used in the â-series and the second in the ô-series.

1	a—éa[A]	5	œ	9	aeu—eu	13	ê
2	ĕ—ĭ	6	ŏ—ŭ	10	œă	14	ai—ey
3	ei—i	7	o—u	11	iĕ	15	aô—oŭ
4	œ	8	uŏ	12	é	16	au—ŏu

[A] With the character ⟡ p a ligature ⟡ is used (see note 4).

V. Shortened syllables and vocalic nuclei with anuswara or visarga

Where variants in romanization separated by a dash are given the one before the dash is to be used in the â-series and the one(s) after the dash in the ô-series.

1	á—ó	4	âm—um	7	ăh—eăh	10	aôh—ŏăh
2	ă—oă,eă[A]	5	om—ŭm	8	ŏh—ŭh	11	ăng—eăng
3	ă—oă,eă[B]	6	ăm—ŏâm	9	éh		

[A] eă if followed by k, ng, h (see note 6).

[B] eă after a short vowel or if followed by k, ng, h (see note 6).

Notes

1. Khmer consonants are divided into 2 series, the *â*-series and the *ô*-series, as indicated in the consonant table. With vocalic nuclei these consonants may produce different romanizations, as shown in the abbreviated vowel table: ក *kâ*, ក្រ *krâ*, ក *kô*, ក្រ *krô*. A Khmer consonant in syllable-final position, not accompanied by a vowel marker or by ៈ, should generally be romanized without a vowel letter following: កក *kâk*, អង្គ *'ângk* (exception: ពង្រោ *pôngrô*, also written ពង្រោរ *pôngrôr* and ពង្រោះ *pôngrôh*).

2. The Khmer diacritical mark ̃ or ̈ written above an *â*-series consonant (except ប and ប៉; see note 4) changes it to the *ô*-series: ហ៊ាង *héang*. The diacritical mark ̈ written above an *ô*-series consonant changes it to the *â*-series: ញ៉ាង *nhâng*. When either of these marks would conflict with another symbol written above a character, the mark ៝ or ៎ may be written in its place: ហ៊ *hi*, ដំរី *dâmrei*. (These marks are frequently omitted in Khmer writing, particularly in words of Indic provenance.)

3. The second consonant of a Khmer graphic cluster is generally written below the base consonant in the special form called a "foot": ខ្នាង *khnâng*. There is no foot for the character ឡ *lâ*. The "feet" ៝ and ៎ usually represent the characters ដ and ថ respectively, rather than ក and ប: ក្ដី *kdei*, កន្ថាយ *kânthéay*, but កន្ត្រប *kântráb*.

A "foot" determines the series of the following vocalic nucleus unless it is a nasal (ង *ng*, ញ *nh*, ណ *n*, ន *n*, ម *m*) or យ *y*, រ *r*, ល *l*, វ *v*, ស *s*. In that case, the base consonant determines the vocalic series: ខ្ពង *khpông*, ល្អ *l'â*, ថ្ម *thmâ*, ស្វាយ *svay*. Syllable-final យ and ង sometimes appear as "feet": ស្វាយ or ស្វ្យ *svay*, ទាំង or ទ្យាំ *teăng*. This practice appears to be optional and such irregular Khmer spellings are not reflected as such in romanization.

4. The combination ប plus ា is written បា *ba*. The latter character is a graphic device designed to prevent confusion with ហា *hâ*. The characters ប and បា with the diacritical mark ̈ are romanized *p* in the *â*-series, rather than as *b* in the *ô*-series: ប៉ង *pâng*, ប៉ៅ *patau*. The diacritical mark ៝ or ៎ is substituted where a conflict with another symbol written above a character would occur: ប៊ី *pei*. The characters ប and បា when accompanied by a "foot" are also romanized as *p* in the *â*-series, although the Khmer diacritical mark is generally omitted: ប្លេង *plêng*, ប្អា *p'â*, ប្រាប *prăb*.

5. The *â*-series consonant អ is romanized by means of an apostrophe ('): ក្អែក *k'êk*, ចង្អៀត *châng'iĕt*, រអិល *rô'ĕl*, អ្វី *'vei*, អាង *'ang*. In word-initial position before a vowel, ' may be omitted: អាង *ang*.

6. The Khmer diacritical mark ' appears only in two combinations: '' (examples: ឋត់ *bát*, ខ្ចស់ *khpós*) and ា'. The diacritical mark ' appears only in the combination ''. In the *â*-series both ា' and '' are romanized *ă*: ឋាត់ and ឋត់ are both romanized *chăk*. In the *ô*-series both ា' and '' are romanized *eă* when followed by *k, ng* or *h*; otherwise, they are romanized *oă*: រោកត់ *rôpeăk*, ឋាត់ *moăt*, វ៉ែង *veăngk*, ភ៉ព *phoăpv*.

7. The combination '' is romanized *r* before a consonant: ឋម៌ *thôrm*. The combination '' is romanized as *r* before a consonant preceded by a shortened vocalic nucleus: ការ៌ *koărr*.

8. The symbol '' in syllable-initial position is ignored in romanization: ស៊ *sâ*, ស៊ស៊ *sâsâs*. In syllable-final position '' indicates that the consonant is vowelled, i.e. followed by *â* in the *â*-series, by *ô* in the *ô*-series: តំណ៌ *tâmnâ*, ពម៌ *pômô*.

9. The diacritical mark '' (which appears above characters and/or vowel markers which are not vocalized) is ignored in romanization: បុណ្យ *bŏny*, ពោធិ៍ *poŭthĭ*, ភូមិ៍ *phumĭ*.

10. The independent character ឥ is romanized either *ŏ* or *ŭ*. A reference source should be consulted where doubt arises.

Other systems of romanization

The **provisional romanization system by the Geographic Department** of the Office of the Council of Ministers (1996) is reflected in the list of Cambodian provinces (ISO 3166-2) and in a tourist map produced jointly by the said department and the I.G.N. in 1997. This diacriticless system renders the consonants essentially in the same way as described above but the rendering of vowels is somewhat different. Based on name examples the following correspondences can be noted with the UN system:

UN	Provisional							
a, ă	= aa	ĕ	= e	iĕ	= ie	rœ̆	= rue
-ăh	= -ah	é	= ee	o	= ou, o	u, ŭ	= u
-ăm	= -am	-éh	= -eh	ŏ, ô, ó	= o	uŏ	= uo
-ăng	= -ang	ê	= ae—eae[A]	oŭ	= ou		
â, á	= a	éa, eă	= ea	œ	= oeo[B]—ueu[A]		
aô	= ao	i, ĭ	= i	œ̆	= oe—ue[A]		

[A] The first romanization is applied in the *â*-series, the second in the *ô*-series.
[B] The romanization *oeo* is not verified.

Not all the vocalic nuclei occur in the name examples. It must be assumed that in most cases the diacritic is simply dropped in romanization.

Macedonian Cyrillic

The United Nations recommended system was approved in 1977 (III/11), based on the romanization of the Serbo-Croatian and Macedonian Cyrillic alphabets that had long been employed in official gazetteers and maps of Yugoslavia. The table was published as an annex to the resolution[18].

The system is used in The former Yugoslav Republic of Macedonia[19] and in international cartographic products.

The Cyrillic alphabet is used. The romanization table is unambiguous and can be applied automatically. The system is on the whole reversible with the exception of the characters ѓ and ќ. These are romanized as *g* and *k*, respectively, before е and и. But also г and к are occasionally used before е and и, so, for example, the name *Makedonija* is reconstituted as Македонија, not Маќедонија in the Cyrillic source script.

Romanization

1	А а	a	9	З з	z	17	Н н	n	25	У у	u
2	Б б	b	10	S ѕ	dz	18	Њ њ	nj	26	Ф ф	f
3	В в	v	11	И и	i	19	О о	o	27	Х х	h
4	Г г	g	12	Ј ј	j	20	П п	p	28	Ц ц	c
5	Д д	d	13	К к	k	21	Р р	r	29	Ч ч	č
6	Ѓ ѓ	đ, g^A	14	Л л	l	22	С с	s	30	Џ џ	dž
7	Е е	e	15	Љ љ	lj	23	Т т	t	31	Ш ш	š
8	Ж ж	ž	16	М м	m	24	Ќ ќ	ć, k^A	32	'	'

^{A} The second variant is used before е and и.

Note. Cursive forms of some characters might be formed differently: *Аа Бб Вв Гг Дд Ѓѓ Ее Жж Зз Ѕѕ Ии Јј Кк Лл Љљ Мм Нн Њњ Оо Пп Рр Сс Тт Ќќ Уу Фф Хх Цц Чч Џџ Шш '.*

---

[18] *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. I. *Report of the Conference*, pp. 29, 31.

[19] See e.g. *Report of The former Yugoslav Republic of Macedonia.* Seventh United Nations Conference on the Standardization of Geographical Names. New York, 13-22 January 1998. Document E/CONF.91/CRP.23.

Malayalam

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[20].

There is no evidence of the use of the system either in India or in international cartographic products.

Malayalam (Malayāḷam) uses an alphasyllabic script whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization table is unambiguous but now outdated because since the 1970's a new orthography has been introduced and the use of many character combinations and ligatures has been simplified. Nevertheless the user of the romanization table would have to recognize many ligatures not given in the original table. The system is mostly reversible but there exist some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (ligatures vs. character sequences), particularly in the case of older orthography.

Romanization

The romanization system below is based on the new orthography, with notes where appropriate on the older spelling. The table of ligatures has been added.

I. Independent vowel characters

1 അ a	3 ഇ i	5 ഉ u	7 ഋ ṛ	9 എ e	11 ഒ ŏ	13 ഔ au
2 ആ ā	4 ഈ ī	6 ഊ ū	8 ഏ ĕ	10 ഐ ai	12 ഓ o

II. Abbreviated vowel characters (• stands for any consonant character)

1 • a	3 •ി i	5 •ു uᴬ	7 •ൃ ṛᶜ	9 െ• e	11 ൊ• ŏ	13 •ൗ auᴰ
2 •ാ ä	4 •ീ ī	6 •ൂ ūᴮ	8 േ• ĕ	10 ൈ• ai	12 ൊ•ാ o

ᴬ Traditional script form: ಃ - ഖു khu, മു mu, etc. Exceptions in the older spelling: ku (kku, ṅku), gu, chhu, ju, ṇu, tu, nu (nnu), bhu, ru, shu, hu.

ᴮ Traditional script form: ಃ - khū, mū, etc. Exceptions in the older spelling: kū, gū, chhū, jū, ṇū, tū, nū, bhū, rū, shū, hū.

ᶜ Traditionally written in conjunction with the consonant character: kṛ.

ᴰ In the older spelling: •ൗ.

[20] *Second United Nations Conference on the Standardization of Geographical Names.* London, 10-31 May 1972. Vol. II. *Technical papers.* United Nations. New York 1974, pp. 149-150; *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. II, *Technical papers,* pp. 393 etc.

III. Other symbols

1 •ഠ ṁ   2 •ഃ ḥ   3 ് (A)

ᴬ Indicates absence of the inherent short [a] vowel, e.g. ക് k , സ് s. Special forms of consonants with that symbol: ൺ ṇ, ൻ n, ർ r, ൽ l, ൾ (ൔ) ḷ.

IV. Consonant characters

1 ക	ka	6 ച	cha	11 ട	ṭa	16 ത	ta	21 പ	pa	26 യ	ya	32 സ	sa
2 ഖ	kha	7 ഛ	chha	12 ഠ	ṭha	17 ഥ	tha	22 ഫ	pha	27 ര	ra	33 ഹ	ha
3 ഗ	ga	8 ജ	ja	13 ഡ	ḍa	18 ദ	da	23 ബ	ba	28 ല	la	34 ക്ഷ	kṣha
4 ഘ	gha	9 ഝ	jha	14 ഢ	ḍha	19 ധ	dha	24 ഭ	bha	29 വ	va	35 ള	ḷa
5 ങ	ṅa	10 ഞ	ña	15 ണ	ṇa	20 ന	na	25 മ	ma	30 ശ	sha	36 ഴ	ḻa
										31 ഷ	ṣha	37 �റ	ṟaᴬ

ᴬ Special ligature: ഺ ṯ.

V. Ligatures
Adscript forms of some consonants

		Example
.	r-	ഗ്ഗ rgga (in modern usage it is substituted by ർഗ്ഗ rgga)
ൄ	-r	ക്ര kra (older style: ക്ര)
•ൃ	-y	ത്യ tya
•ൟ	-v	സ്വ sva

Other ligatures (the list is not complete)

ക്ക	ക്ട	ക്ത	ക്ഷ	ക്ല	ഗ്ഗ	ഗ്ദ	ഗ്ദ്ധ*	ഗ്ന	ഗ്മ	ഗ്ല	ങ്ക
kka	kṭa	kta	kṣha	kla	gga	gda	gddha*	gna	gma	gla	ṅka
ങ്ങ	ച്ച	ച്ഛ	ജ്ജ	ജ്ഞ	ഞ്ച	ഞ്ഛ*	ഞ്ജ*	ഞ്ഞ	ട്ട	ഡ്ഡ	ണ്ട
ṅṅa	chcha	chchha	jja	jña	ñcha	ñchha*	ñja*	ñña	ṭṭa	ḍḍa	ṇṭa
ണ്ഠ*	ണ്ഡ	ണ്ണ	ണ്മ	ത്ത	ത്ഥ	ത്ന*	ത്ഭ	ത്മ	ത്സ	ദ്ദ	ദ്ധ
ṇṭha*	ṇḍa	ṇṇa	ṇma	tta	ttha	tna*	tbha	tma	tsa	dda	ddha
ന്ത	ന്ഥ	ന്ദ	ന്ധ	ന്ന	ന്മ	ന്റ	പ്ത*	പ്ന*	പ്പ	പ്ല	പ്സ*
nta	ntha	nda	ndha	nna	nma	nṟa	pta*	pna*	ppa	pla	psa*
ബ്ദ	ബ്ധ	ബ്ബ	ബ്ല	മ്ന	മ്പ	മ്മ	മ്ല	യ്ക	യ്ക്ക	യ്ത	യ്ത്ത
bda	bdha	bba	bla	mna*	mpa	mma	mla	yka*	ykka	yta*	ytta*
യ്പ	യ്മ	യ്യ	ല്ക	ല്ക്ക	ല്പ	ല്മ	ല്ല	വ്വ	ശ്ച	ശ്ല	ശ്ഷ
ypa*	yma*	yya	lka*	lkka*	lpa	lma*	lla	vva	shcha	shla	shsha
ഷ്ക	ഷ്ട	ഷ്ഠ	ഷ്ന	ഷ്പ	ഷ്ഫ	ഷ്മ	സ്ക	സ്ഇ	സ്ത	സ്ഥ	സ്ന
ṣhka*	ṣhṭa	ṣhṭha*	ṣhṇa*	ṣhpa*	ṣhpha*	ṣhma*	ska*	sïa	sta*	stha	sna*
സ്പ	സ്ഫ	സ്മ	സ്ല	സ്സ	ഹ്ന	ഹ്മ	ഹ്ല	ള്ള	ള്ക	ള്ച	ള്ട
spa*	spha*	sma*	sla	ssa	hna	hma	hla	ḷḷa	ḷka*	ḷcha*	ḷṭta*
ള്മ	ള്പ										
ḷma*	ḷpa*										

Many of the ligatures, especially those marked with the asterisk are thought to be obsolete since the introduction of the new spelling and have been replaced by a combination of two consonant characters, the first of which bears the ˘ diacritic, e.g. traditional പ്ന, modern പ്ന pna. However, the usage of ligatures in different sources varies greatly.

Other systems of romanization

For differences between the UN system and the **ISO transliteration standard ISO 15919: 2001** see the section on the romanization of Hindi.

The **State of Kerala** seems to use a regular pattern for rendering their geographical names. At least the following correspondences may be noted:

UN	State of Kerala
ḷ	= zh
sh	= s
ṭ	= th
ṭ	= t or d

Marathi

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[21].

There is no evidence of the use of the system either in India or in international cartographic products.

Marathi (Marāṭhī) uses the alphasyllabic script Devanāgarī. For further details see the Hindi section of the report.

Romanization

The romanization system is given in the Hindi section of the report. The main exceptions to the table are as follows.

1. Additional characters (absent in Hindi):
    ळ ḷa
    ऍ ă (abbreviated: कॅ kă)

2. Different shapes of some characters and ligatures:

Hindi	Marathi	Roman
ख	स	kha
ल	ऴ	la
क्त	क्त	kta
क्क	क्क	kka

3. Ligatures may be formed with r as the first component: र्य rya, र्ह rha.
4. Typical are ligatures ending in h: न्ह nh, म्ह mh, ल्ह lh, व्ह vh.

---

[21] *Second United Nations Conference on the Standardization of Geographical Names*. London, 10-31 May 1972. Vol. II. *Technical papers*. United Nations. New York 1974, p. 133; *Third United Nations Conference on the Standardization of Geographical Names*. Athens, 17 August - 7 September 1977. Vol. II, *Technical papers*, pp. 393 etc.

5. The following independent characters are sometimes used: इ i, ई ī, उ u, ऊ ū, ए e, ऐ ai.

Mongolian
(in China)

The United Nations resolution III/8 in 1977 recognized the Scheme for a Chinese Phonetic Alphabet (Pinyin) as China's official Roman alphabet scheme and recommended the alphabet as the international system for the romanization of Chinese geographical names. In China Mongolian geographical names are transcribed directly from the Mongolian script. The scheme was published in *Toponymic Guidelines for Map and Other Editors: China*, 1982[22].

The system is used in China and in international cartographic products.

Mongolian uses a vertical script. Transcription of Mongolian names is made from their proper pronunciation based on the written form of the Mongolian language and Qahar vernacular, with Zhenglan as its representative pronunciation. Due to the complex nature of the script the romanization scheme is not reversible.

Romanization

1	a	8	h	15	o	22	w
2	b	9	i	16	p	23	x
3	c	10	j	17	q	24	y
4	d	11	k	18	r	25	z
5	e	12	l	19	s	26	o [ô]
6	f	13	m	20	t	27	u [û]
7	g	14	n	21	u		

Notes
1. The long and short Mongolian vowels are not distinguished in the spelling for general use, but in recording the pronunciation of place-names, the long vowel is represented by duplication.
2. Where two Roman equivalents are given, the second (in brackets) is used for recording the pronunciation of place-names while the first form is for general use.
3. In the table only word-initial character variants are shown. Depending on the position in the word many variants of the characters are used as well as some ligatures. These features are not covered here.

---

[22] *Fourth United Nations Conference on the Standardization of Geographical Names.* Geneva, 24 August - 14 September 1982. Vol. II. *Technical papers*, pp. 121-125.

Nepali

The United Nations recommended system was approved in 1972 (II/11), based on a report prepared by D. N. Sharma. The note on the system was published in volume II of the conference report[23].

There is no evidence of the use of the system either in Nepal or in international cartographic products. The resolution IV/17 (1982) recommended association, inter alia, with Nepal in carrying out further studies on the system.

Nepali (Nepālī) uses the alphasyllabic script Devanāgarī. For further details see the Hindi section of the report.

Romanization

The romanization system is given in the Hindi section of the report. The following exceptions should be noted.

1. In normal spelling ṛ and ṛh are not marked, instead dotless ड ḍa, ढ ḍha are used to denote these sounds.
2. Ligatures may be formed with r as the first component: ऱ rha.

Other systems of romanization

The **BGN/PCGN 1964 System** has the following different Roman equivalents for the romanization of Nepali:

UN	BGN/PCGN				
ṁ	= ń	ṛ	= ri, ir[A]	ṣh	= ṣ
ṅ	= ng	ṝ	= rī, īr[A]	v	= w
ṛ	= ṛ	sh	= ś		

[A] The second variant is used in word-medial position.

Oriya

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[24].

---

[23] Second United Nations Conference on the Standardization of Geographical Names. London, 10-31 May 1972. Vol. II. Technical papers. United Nations. New York 1974, p. 133

[24] Second United Nations Conference on the Standardization of Geographical Names. London, 10-31 May 1972. Vol. II. Technical papers. United Nations. New York 1974, pp. 143-144; Third United Nations Conference on the Standardization of Geographical Names. Athens, 17 August - 7 September 1977. Vol. II, Technical papers, pp. 393 etc.

There is no evidence of the use of the system either in India or in international cartographic products.

Oriya uses an alphasyllabic script whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization table is unambiguous. The system is mostly reversible but there may exist some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (combinations with subscript consonants vs. character sequences).

Romanization

In the romanization system below the table of subscript and adscript consonants and the list of ligatures has been added.

I. Independent vowel characters

1 ଅ	a	3 ଇ	i	5 ଉ	u	7 ଋ	ṛ[A]	8 ଏ	e	10 ଓ	o[B]
2 ଆ	ā	4 ଈ	ī	6 ଊ	ū			9 ଐ	ai	11 ଔ	au

[A] Rarely used characters that are positioned between characters 7 and 8: ଌ ṝ, ୡ ḷ.

[B] Special form: ଓ va (not given in the original table).

II. Abbreviated vowel characters (• stands for any consonant character)

1 •	a	3 •ੰ	i[A]	5 •ੁ	u	7 •ੂ	ṛ	8 6•	e	10 6•l	o
2 •l	ā	4 •l	ī	6 •ੂ	ū			9 6•ੈ (•ੈ) ai		11 6• ℩(• ℩) au	

[A] Variations: ଥୁ thi, ଧୁ dhi.

III. Other symbols

1 •°	ṁ	2 •ˇ	m̐	3 •ঃ	ḥ	4 •ˎ	(A)

[A] Pronunciation without a vowel: କ୍ k.

IV. Consonant characters

1 କ	ka	6 ଚ	cha	11 ଟ	ṭa	16 ତ	ta	21 ପ	pa	26 ଯ	ya[B]	30 ଶ	sha
2 ଖ	kha	7 ଛ	chha	12 ଠ	ṭha	17 ଥ	tha	22 ଫ	pha	27 ର	ra	31 ଷ	ṣha
3 ଗ	ga	8 ଜ	ja	13 ଡ	ḍa[A]	18 ଦ	da	23 ବ	ba	28 ଲ	la	32 ସ	sa
4 ଘ	gha	9 ଝ	jha	14 ଢ	ḍha[A]	19 ଧ	dha	24 ଭ	bha	29 ଳ	ḷa	33 ହ	ha
5 ଙ	ṅa	10 ଞ	ña	15 ଣ	ṇa	20 ନ	na	25 ମ	ma			34 କ୍ଷ	kṣha

[A] Dotted variants of the characters: ଡ଼ ṙa, ଢ଼ ṙha.

[B] Special form: ଯ ya.

V. Subscript and adscript consonant characters

Character numbers correspond to those in Table IV.

1	k	6	ch	11	ṭ	16	t	21	p	26	y	
2	kh	7	chh	12	ṭh	17	th	22	ph	27	r	
3	g	8	j	13	ḍ	18	d	23	b	28	l	32 s
4	gh	9	jh	14	ḍh	19	dh	24	bh	29	ḷ	
				15	ṇ	20	n	25	m			

VI. Ligatures (the list is not complete)

kka	kta	kḷa	kṣha	kṣhṇa	ksa	ṅka	ṅkṣha	ṅkha	ṅga	ṅgha	chcha
chhchha	jja	jjha	jña	ñcha	ñchha	ñja	ñjha	ṭṭa	ḍḍa	ṇṭa	ṇṭha
ṇḍa	ṇḍha	ṇṇa	tka	tta	tpa	tra	tsa	dga	dda	ddha	dbha
nta	ntra	ntha	nda	ndha	nna	pta	psa	bda	bba	mpa	mpha
mba	mbha	mma	ḷka	ḷpa	ḷpha	ḷbha	shcha	shchha	shka	ṣhṭa	ṣhṭha
ṣhṇa	ṣhpa	ṣhpha	ska	skha	sta	stra	stha	spa	spha	hna	hla

Combinations with *r* as the first component are written ʿ , e.g. rka.

Other systems of romanization

For differences between the UN system and the **ISO transliteration standard ISO 15919: 2001** see the section on the romanization of Hindi.

Persian

The United Nations recommended system was approved in 1967 (I/13), based on the official system adopted by Iran and published in its English version as *Transliteration of Farsi Geographic Names to Latin Alphabet* (September 1966). The romanization table was also published as an annex to the Toponymic Guidelines for the Islamic Republic of Iran in 2000[25].

The system is used in the Islamic Republic of Iran and in international cartographic products.

Persian (Farsi) uses the Perso-Arabic script that is written from right to left. The Persian script usually omits vowel points and diacritical marks from writing which makes it difficult to obtain

---

uniform results in the romanization of Persian. The romanization is generally reversible though there are some ambiguous letter sequences (*kh, sh, th, zh*) which also may represent the romanized values of two Persian characters in addition to the respective single ones.

Romanization

For the initial, medial and final forms of the principal characters in the romanization table below see the Arabic section of the report.

1	ا	- [A]	9	خ	kh	17	ص	ş	25	ک	k
2	ب	b	10	د	d	18	ض	z̧	26	گ	g
3	پ	p	11	ذ	z̄	19	ط	ţ	27	ل	l
4	ت	t	12	ر	r	20	ظ	z̧	28	م	m
5	ث	s̱	13	ز	z	21	ع	‘	29	ن	n
6	ج	j	14	ژ	zh	22	غ	gh	30	و	v
7	چ	ch	15	س	s	23	ف	f	31	ه	h
8	ح	ḩ	16	ش	sh	24	ق	q	32	ی	y

[A] Not romanized, but see romanizations accompanying *alef* (ا) in the table for vowels.

Vowel characters and diacritical marks (• stands for any consonant)

1	•َ	a	5	َیٰ /ٰ•َ	á	9	ـیَ /ـیْ•َ	ey	13	•ّ	(C)
2	•ِ	e	6	آ	’ā, ā [A]	10	•ُو /ـوَ•ُ	ow	14	•ْ, ـیْ•	-e [D], -ye [E]
3	•ُ	o	7	•ِی /ـی	ī	11	•ٔ	(B)			
4	ـا /ٰا•	ā	8	•ُو /ـو	ū	12	•ٔ	’			

[A] Word-initially.
[B] Not romanized; marks absence of the vowel.
[C] Doubling of the consonant letter.
[D] After a consonant (excl. -ah).
[E] After a vowel (see also note 2).

Notes

1. The adjectival ending of Arabic origin ـیّه in Persian is romanized *-īyeh*. In romanizing the definite article the same rules of assimilation of consonants are applied as in Arabic, e.g. زین الدین Zeyn od Dīn.
2. The relational suffix (*ezāfeh*) *-e* is usually not expressed in Persian writing after a consonant. After final ا or و it is written with ی, e.g. پای آب Pā-ye Āb. After final ی and ه it is expressed by writing *hamzeh* over the character: دهانهٔ ممبر Dahāneh-ye Mambar.
3. To point Persian vowels two systems are in use that are separated by a slash in the table. The first system is a Persian one while the other adheres to the Arabic tradition. In normal spelling vowel points are not used.

Other systems of romanization

The simplified transliteration of Persian in the standard **ISO 233-3:1999** has the following

equivalents to the UN system:

UN	ISO						
ā	= â	ī	= y?	sh	= š	ẕ	= ż
ch	= č*	kh	= ḵ	ṭ	= ṭ	ⱬ	= ẓ*
gh	= ġ	s̄	= s̱	ū	= v?	zh	= ž*
ḥ	= ḥ	ṣ	= ṣ	z̄	= ẕ		

Note. The ISO standard obviously contains some typographic errors which are corrected and indicated by an asterisk in the table above. Consequently, not all the correspondences here can be considered as verified.

Punjabi

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[26].

There is no evidence of the use of the system either in India or in international cartographic products.

Punjabi (Panjābī) uses an alphasyllabic script (Gurmukhi) whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization table is unambiguous. The system is mostly reversible but there exist some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (combinations with subscript consonants vs. character sequences).

Romanization

1 ੳ	(u)[A]	6 ਕ	ka	11 ਚ	cha	16 ਟ	ṭa	21 ਤ	ta	26 ਪ	pa	31 ਯ	ya
2 ਅ	a[A]	7 ਖ	kha[B]	12 ਛ	chha	17 ਠ	ṭha	22 ਥ	tha	27 ਫ	pha[B]	32 ਰ	ra
3 ੲ	(i)[A]	8 ਗ	ga[B]	13 ਜ	ja[B]	18 ਡ	ḍa	23 ਦ	da	28 ਬ	ba	33 ਲ	la
4 ਸ	s[B]	9 ਘ	gha	14 ਝ	jha	19 ਢ	ḍha	24 ਧ	dha	29 ਭ	bha	34 ਵ	va
5 ਹ	h	10 ਙ	ṅa	15 ਞ	ña	20 ਣ	ṇa	25 ਨ	na	30 ਮ	ma	35 ੜ	ṛa

[A] These characters are used in combination with abbreviated vowel characters: ਉ u, ਊ ū, ਆ a, ਆ ā, ਐ ai, ਔ au, ਇ i, ਈ ī, ਏ e; exceptional variation: ੳ o.

[B] Dotted variants of the characters: ਸ਼ sha, ਖ਼ ḵha, ਗ਼ ġa, ਜ਼ za, ਫ਼ fa.

---

[26] *Second United Nations Conference on the Standardization of Geographical Names.* London, 10-31 May 1972. Vol. II. *Technical papers.* United Nations. New York 1974, pp. 136-138; *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. II, *Technical papers,* pp. 393 etc.

Abbreviated vowel characters and other symbols (• stands for any consonant character)

1	•	a	3	f•	i	5	•͜	u	7	•̇	e	9	•̇	o	11	•̇	ṁ^A	13	•̇	(C)

1	•	a	3	f•	i	5	•͜	u	7	•̇	e	9	•̇	o	11	•̇	ṁ^A	13	•̇	(C)
2	•ᵀ	ā	4	•ᵌ	ī	6	•͜	ū	8	•̇	ai	10	•̇	au	12	•̇	ṁ^B			

A Used if it is preceded by short vowels (a, i, u) and ū, excluding the independent vowel character
ੳ (u, ū), e.g. ਸੰਗ samg, ਸਿੰਗ simg, ਬੁੰਦਾ bumdā, ਬੂੰਦ būmd, ਇੰਜਨ imjan.

B Used in all other occasions, e.g. ਝੋਂਕਾ jhomkā.

C Marks doubling of the following consonant: ਨਿੱਕਾ nikkā, ਲੱਭਣਾ labhbhṇā, ਕੁੱਤਾ kuttā, ਹਿੱਸਾ hissā, ਲੱਮਾ lammā.

Subscript consonant characters

1	•̣	-h	2	•̣	-v	3	•̣	-r	4	•̣	-y

Note. Absence of the inherent vowel (-a) is not marked in the spelling in any way except for the combinations with subscript characters and those which are doubled by ˘ (adhaka).

Other systems of romanization

For differences between the UN system and the **ISO transliteration standard ISO 15919: 2001** see the section on the romanization of Hindi.

Russian

The United Nations recommended system was approved in 1987 (V/18), based on the official system of the Main Administration of Geodesy and Cartography of the former Soviet Union, also known as the GOST 1983 system (GOST 16876-71). The table was published as an annex to the resolution[27].

The system is used in the Russian Federation and increasingly in international cartographic products.

Russian uses the Cyrillic alphabet which is alphabetic. The romanization table is unambiguous and can be applied automatically. The system is reversible, although very rarely there can be ambiguities. For example, a geographical name Ojusardah of the Sakha Republic (Yakutia) is reconstituted as Ойусардах in the Cyrillic source script, not Оюсардах as would seem logical.

---

[27] *Fifth United Nations Conference on the Standardization of Geographical Names.* Montreal, 18-31 August 1987. Vol. I. *Report of the Conference,* pp. 40-41.

Romanization

1	A a	a	10	И и	i	19	C c	s	28	Ъ ъ	"
2	Б б	b	11	Й й	j	20	Т т	t	29	Ы ы	y
3	В в	v	12	К к	k	21	У у	u	30	Ь ь	'
4	Г г	g	13	Л л	l	22	Ф ф	f	31	Э э	è
5	Д д	d	14	М м	m	23	Х х	h	32	Ю ю	ju
6	Е е	e	15	Н н	n	24	Ц ц	c	33	Я я	ja
7	Ё ё	ë	16	О о	o	25	Ч ч	č			
8	Ж ж	ž	17	П п	p	26	Ш ш	š			
9	З з	z	18	Р р	r	27	Щ щ	šč			

Note. Cursive forms of some characters might be formed differently: *Аа Бб Вв Гг Дд Ее Ёё Жж Зз Ии Йй Кк Лл Мм Нн Оо Пп Рр Сс Тт Уу Фф Хх Цц Чч Шш Щщ Ъъ Ыы Ьь Ээ Юю Яя.*

Other systems of romanization

The **BGN/PCGN 1947 System** provides for the romanization, as a single block, of the following characters differently from the UN system (the Cyrillic character is followed in parentheses by the romanization according to the UN system):

е (e)	e, ye[A]	й (j)	y	ч (č)	ch	э (è)	e
ё (ë)	ë, yë[A]	х (h)	kh	ш (š)	sh	ю (ju)	yu
ж (ž)	zh	ц (c)	ts	щ (šč)	shch	я (ja)	ya

[A] Initially, after the vowel characters and after й, ъ, and ь.

The transliteration of Cyrillic characters contained in the standard **ISO 9:1995** provides for the romanization, as a single block, of the following characters differently from the UN system:
щ (šč)   ŝ   ю (ju)   û   я (ja)   â

Serbo-Croatian Cyrillic

The United Nations recommended system was approved in 1977 (III/11), based on the Roman (Croatian) equivalent of the Serbo-Croatian Cyrillic alphabet that had long been employed in official gazetteers and maps of Yugoslavia. The table was published as an annex to the resolution[28].

The system is used in Yugoslavia, in Bosnia and Herzegovina, and in international cartographic products.

The Cyrillic script used in Serbo-Croatian (here more precisely: Serbian) is alphabetic. The romanization table is unambiguous and can be applied automatically. The system is reversible.

---

[28] *Third United Nations Conference on the Standardization of Geographical Names*. Athens, 17 August - 7 September 1977. Vol. I. *Report of the Conference*, pp. 29, 31.

Romanization

1	A a	a	9	З з	z	17	Њ њ	nj	25	Ф ф	f
2	Б б	b	10	И и	i	18	O o	o	26	X x	h
3	В в	v	11	J j	j	19	П п	p	27	Ц ц	c
4	Г г	g	12	К к	k	20	P p	r	28	Ч ч	č
5	Д д	d	13	Л л	l	21	C c	s	29	Џ џ	dž
6	Ђ ђ	đ (Đ)	14	Љ љ	lj	22	T т	t	30	Ш ш	š
7	E e	e	15	М м	m	23	Ћ ћ	ć			
8	Ж ж	ž	16	Н н	n	24	У y	u			

Note. Cursive forms of some characters might be formed differently: *Aa Бб Вв Гг Дд Ђђ Ее Жж Зз Ии Јј Кк Лл Љљ Мм Нн Њњ Оо Пп Рр Сс Тт Ћћ Уу Фф Хх Цц Чч Џџ Шш.*

Tamil

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[29].

There is no evidence of the use of the system either in India, Sri Lanka or in international cartographic products.

Tamil uses an alphasyllabic script whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization table is unambiguous. The system is on the whole reversible.

Romanization

I. Independent vowel characters

1	அ	a	3	இ	i	5	உ	u	7	எ	ĕ	9	ஐ	ai	11	ஒ	o
2	ஆ	ā	4	ஈ	ī	6	ஊ	ū	8	ஏ	e	10	ஓ	ŏ	12	ஔ	au

II. Abbreviated vowel characters and other symbols (• stands for any consonant character)

1	•	a	3	•ி	i[B]	5	•ு	u[D]	7	ெ•	ĕ	9	ை•	ai[F]	11	ொ•	o	13	∴	ḥ[G]
2	•ா	ā[A]	4	•ீ	ī[C]	6	•ூ	ū[E]	8	ே•	e	10	ோ•	ŏ	12	ௌ•	au	14	•	(H)

[A] Exceptions in the older and Classical literature: ணா ṇā, ரா rā, ணா ṇā.

[B] Variation: டி ti.

---

[29] *Second United Nations Conference on the Standardization of Geographical Names.* London, 10-31 May 1972. Vol. II. *Technical papers.* United Nations. New York 1974, pp. 153-154; *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. II, *Technical papers,* pp. 393 etc.

^C Variation: ட tī.

^D Variations: கு ku, நு ṅu, சு chu, ஞு ñu, டு ṭu, ணு ṇu, து tu, நு nu, பு pu, மு mu, யு yu, ரு ru, லு lu, etc.

^E Variations: கூ kü, ஙூ ṅü, சூ chü, ஞூ ñü, டூ ṭü, ணூ ṇü, தூ tü, நூ nü, பூ pü, மூ mü, etc.

^F Exceptions in the older and Classical literature: லை lai, ளை ḷai, ணை ṇai, னை ṉai.

^G Absent in the original table.

^H Pronunciation without a vowel: க் k.

III. Consonant characters

1 க ka	5 ட ṭa	9 ப pa	11 ய ya	15 ழ ḷa	19 ஜ ja^A	23 ஹ ha^A
2 ங ṅa	6 ண ṇa	10 ம ma	12 ர ra	16 ள ḷa	20 ஶ sha^{AB}	24 க்ஷ kṣha
3 ச cha	7 த ta		13 ல la	17 ற ṟa	21 ஷ ṣha	25 ற்ற ṟṟa
4 ஞ ña	8 ந na		14 வ va	18 ன ṉa	22 ஸ sa^A	26 ன்ற ṉṟa

^A Characters borrowed from the Grantha script.

^B Special syllable character: ஸ்ரீ shrī.

Other systems of romanization

For differences between the UN system and the **ISO transliteration standard ISO 15919: 2001** see the section on the romanization of Hindi.

Telugu

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[30].

There is no evidence of the use of the system either in India or in international cartographic products.

Telugu uses an alphasyllabic script whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants. The romanization table is unambiguous. The system is mostly reversible but there may exist some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (combinations with subscript consonants vs. character sequences).

Romanization

In the romanization system below the table for subscript consonant characters has been added.

---

[30] *Second United Nations Conference on the Standardization of Geographical Names.* London, 10-31 May 1972. Vol. II. *Technical papers.* United Nations. New York 1974, pp. 145-146; *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. II, *Technical papers*, pp. 393 etc.

I. Independent vowel characters

1 అ a	4 ఈ ī	7 ఋ ṛ	9 ఎ ĕ	12 ఒ ŏ
2 ఆ ā	5 ఉ u	8 ౠ ṝ	10 ఏ e	13 ఓ o
3 ఇ i	6 ఊ ū		11 ఐ ai	14 ఔ au

II. Abbreviated vowel characters and other symbols (• stands for any consonant character)

1 ా, • a	4 ీ ī	7 ృ ṛ	8 •ె ĕ	11 •ొ ŏ[C]	14 ం ṁ	17 ్ (F)
2 ా ā	5 ు u[B]		9 •ే e	12 •ో o[D]	15 ఁ m̆[E]
3 •ి i[A]	6 ూ ū		10 •ై ai	13 •ౌ au	16 ః ḥ

[A] Variations: చి chi, ని ni, యి yi, లి li, etc.

[B] Variations: జు ju, పు pu, వు vu, etc.

[C] Variations: మొ mŏ, యొ yŏ.

[D] Variations: మో mo, యో yo.

[E] Signified historic nasal sound, now obsolete.

[F] End-of-syllable mark (i.e., a consonant without a vowel): రంగపూర్ Raṁgapūr.

Note. Due to technical reasons, the vowel diacritics above that should appear on top of a character, are shifted to the right. To exemplify the position of the diacritics, here is the row with the consonant k:

క	కా	కి	కీ	కు	కూ	కృ	కె	కే	కై	కొ	కో	కౌ
ka	kā	ki	kī	ku	kū	kṛ	kĕ	ke	kai	kŏ	ko	kau

III. Consonant characters

1 క ka	6 చ cha	11 ట ṭa	16 త ta	21 ప pa	26 య ya	31 ష ṣha
2 ఖ kha	7 ఛ chha	12 ఠ ṭha	17 థ tha	22 ఫ pha	27 ర ra[A]	32 స sa
3 గ ga	8 జ ja	13 డ ḍa	18 ద da	23 బ ba	28 ల la	33 హ ha
4 ఘ gha	9 ఝ jha	14 ఢ ḍha	19 ధ dha	24 భ bha	29 వ va	34 ళ ḷa
5 ఙ ṅa	10 ఞ ña	15 ణ ṇa	20 న na	25 మ ma	30 శ sha	

[A] Variant: ఱ.

IV. Subscript consonant characters

1 ్క k	3 ్న n	5 ్య y	7 •్ల l
2 ్త t	4 ్మ, ్మ m	6 ్ర r	8 ్వ v

Most other subscript consonant characters are recognizable modifications of the main character, e.g. చ్చ chcha, ట్ట ṭṭa, ప్ప ppa.

Other systems of romanization

For differences between the UN system and the **ISO transliteration standard ISO 15919: 2001** see the section on the romanization of Hindi.

Thai

The United Nations recommended system was approved in 1967 (resolution I/14), based on the modified general system of the Royal Institute of Thailand. The table and notes on its application were published in vol. II of the conference reports[31].

The system is used in Thailand and in international cartographic products. In recent years, however, a revised version of the system had been prepared and presented as *Principles of Romanization for Thai Script by Transcription Method* to the UNGEGN Working Group on Romanization Systems in October 1997 (hereinafter referred to as the 1997 revised version). The revision was about to be presented to the Seventh United Nations Conference on the Standardization of Geographical Names (New York, January 1998) but this was postponed because of the absence of the Thai delegation at the conference.

Thai uses an alphasyllabic script. In Thai writing word division is not ordinarily indicated which may complicate the romanization. The romanization system is not reversible to its original script form.

Romanization

Thai characters are romanized in two ways depending on whether they denote syllable-initials or syllable-finals. The romanization of syllable-finals is indicated in parentheses.

1 ก k (k)	8 ฉ ch (t)	14 ฎ d (t)	20 ด d (t)	26 บ b (p)	34 ย y	42 ศ s (t)
2 ข kh (k)	9 ช ch (t)	15 ฏ t (t)	21 ต t (t)	27 ป p (p)	35 ร r (r)	43 ษ s (t)
3 ฃ kh (k)	10 ฌ ch (t)	16 ฐ th (t)	22 ถ th (t)	28 ผ ph (p)	36 ฤ ru[BC]	44 ส s (t)
4 ค kh (k)	11 ซ s (t)	17 ฑ th,d(t)	23 ท th (t)[A]	29 ฝ f (p)	37 ฤๅ ru[B]	45 ห h[D]
5 ฅ kh (k)	12 ฎ ch (t)	18 ฒ th (t)	24 ธ th (t)	30 พ ph (p)	38 ล l (n)	46 ฬ l (n)
6 ฆ kh (k)	13 ญ y (n)	19 ณ n (n)	25 น n (n)	31 ฟ f (p)	39 ฦ lu[B]	47 อ o[E]
7 ง ng(ng)				32 ภ ph (p)	40 ฦๅ lu[B]	48 ฮ h
				33 ม m (m)	41 ว w	

[A] Combination ทร is romanized *s* if so pronounced.

[B] These Thai characters are traditionally considered as vowel characters; for their revised romanizations see below.

[C] Also romanized *ri* and *ree*.

[D] See note 7.

[E] See note 8.

---

[31] *United Nations Conference on the Standardization of Geographical Names.* Geneva, 4–22 September 1967. Vol. II. *Proceedings of the Conference and technical papers*, pp. 122-123.

Vocalic nuclei (• stands for any consonant character)

#		Rom.	#		Rom.	#		Rom.	#		Rom.
1	•ะ	a	14	เ˘••	e	27	เ•ียะ	ia	40	•าว	ao
2	˘••	a	15	เ•	e	28	เ•ีย	ia	41	•ุย	ui
3	•รร•	a	16	แ•ะ	ae	29	เ•ือะ	ua*	42	โ•ย	oi
4	•า	a	17	แ˘••	ae	30	เ•ือ	ua*	43	•อย	oi
5	•รร	an	18	แ••	ae	31	•˘วะ	ua	44	เ•ย	oei
6	•ำ	am	19	โ•ะ	o	32	•˘ว	ua	45	เ•ือย	uai*
7	•̂	i	20	••	o	33	•ว•	ua	46	•วย	uai
8	•̃	i	21	โ•	o	34	ใ•	ai	47	•̂ว	iu*
9	•̃	u*	22	เ•าะ	o	35	ไ•	ai	48	เ•ว	eo
10	•̃	u*	23	•อ	o	36	•˘ย	ai	49	เ•ว	eo
11	•ุ	u	24	เ•อะ	oe	37	ไ•ย	ai	50	แ•ว	aeo
12	•ู	u	25	เ˘••	oe	38	•าย	ai	51	แ•ว	aeo
13	เ•ะ	e	26	เ•อ	oe	39	เ•า	ao	52	เ•ยว	ieo*

* Romanization equivalents marked with an asterisk have been revised in the 1997 version.

Notes[32]

1. Geographical names are transliterated individually according to their pronunciation, words should not be divided into meaningless syllables, e.g. *Chorakhe* (not *Cho Ra Khe*), *Samphao* (not *Sam Phao*).

2. Names comprising many syllables, of which the last and the first character of the first and second syllable respectively are vowels should be separated by a hyphen, e.g. *Saing* is written *Sa-ing*.

3. A word or a group of words followed by a *maiyamok* (ๆ) is written twice according to the reading rules, e.g. ทำบ่อย ๆ *tham boi boi*, ไฟไหม้ ๆ *fai mai fai mai*.

4. A word or a group of words followed by a *paiyannoi* (ฯ), a shortened form of a well known word or expression, is to be written in full according to the reading rules, e.g. กรุงเทพฯ *Krung Thep Maha Nakhon*.

---

[32] Notes 1-2 have been taken from the original paper (1967), 3-4 from the revised version (1997), and 5-10 have been adapted from the BGN/PCGN 1970 romanization system, to cover the most essential features of the Thai script.

5. A Thai consonant character with no accompanying vowel mark or diacritical mark may be romanized with the vowel letter *a* or *o*: ปฐม *pathom*, พลบ *phlop*.

6. When a single consonant character represents both the final consonant of one syllable and the initial consonant of the next syllable, both consonants should be reflected in romanization: ศักราช *sakkarat*, ราชสีห์ *ratchasi*.

7. ห at the beginning of a syllable and serving as a tone mark should not be romanized: หวอ *wo*, หลืบ *lup*; otherwise, ห should be romanized *h*: หวย *huai*, โวหาร *wohan*.

8. อ at the beginning of a syllable generally represents a glottal stop followed by a vowel: อรุณ *arun*; less frequently it serves as a tone mark: อยาก *yak*. In both instances, อ should not be romanized. See also the vocalic nuclei.

9. The silencing mark • written over the final character of a word, together with the vowel and/or consonant character(s) to which the mark applies, should not be romanized: กรรมสิทธิ์ *kammasit*, พระอินทร์ *Phra-in*.

10. The vowel shortening mark ั and the tone marks ่, ้, ๊, and ๋ should not be romanized: ก็ *ko*, ยี่ห้อ *yiho*, กระต๊อบ *kratop*.

Other systems of romanization

The **1997 revised version**[33] of the official Thai romanization system gives different Roman equivalents to the following Thai characters:

| ฤ | rue,ri,roe | ฦ | lue | ื | ue | เือะ | uea | เือย | ueai | เียว | iao |
| ฤๅ | rue | ฦๅ | lue | ื | ue | เือ | uea | ็อ | io |

The transliteration of Thai in the standard **ISO 11940:1998**[34] gives each Thai character a unique equivalent in the Roman script (a single letter or a combination of a letter plus a "modifier letter"). The Thai characters are transliterated strictly in the order they are written, diacritical marks accompanying a consonant are transliterated starting by the uppermost mark and finishing with the subscript mark. As an example, here are some transliteration equivalents: ก k, ข k̄h, ข

---

[33] See *Thailand. Principles of Romanization for Thai Script by Transcription Method.* United Nations Group of Experts on Geographical Names Newsletter. Number 22. January 1999, pp. 35-44a.

[34] This is a modification of a document presented in 1992: *Romanized Transliteration of Thai.* Seventh United Nations Conference on the Standardization of Geographical Names. Vol. II. *Technical papers.* New York, 25 August - 3 September 1992, pp. 369-379. Some consonants have been given revised equivalents in the ISO standard.

Ḵh, ᠊ᠠ kh, ᠊ᠠ k'h, ᠊ᠠ ḵh, ᠊ᠠ ng; ꞌ a, ᷅ a̧, ᠊ ā, ᠊ å̊, ᠊ e, ᠊ æ, etc.

Tibetan

The United Nations resolution III/8 in 1977 recognized the Scheme for a Chinese Phonetic Alphabet (Pinyin) as China's official Roman alphabet scheme and recommended the alphabet as the international system for the romanization of Chinese geographical names. In China Tibetan geographical names are transcribed directly from the Tibetan script. The scheme was published in *Toponymic Guidelines for Map and Other Editors: China*, 1982[35].

The system is used in China and in international cartographic products.

Tibetan uses an alphasyllabic script. The romanization scheme is based on the pronunciation of Tibetan names in the Tibetan broadcasts by the Central People's Broadcasting Station (the Lhasa pronunciation). Due to the historic nature of the script the rules of correspondence between the script and the pronunciation are complicated and the romanization is not reversible to its original written Tibetan form.

Romanization

Consonant characters

1	ग	ga	5	ज	ja	9	द	da	13	ब	ba	17	za
2	ख	ka	6	छ	qa	10	त	ta	14	प	pa	18	ca
3		ka,ga[A]	7		qa,ja[A]	11		ta,da[A]	15		pa,ba[A]	19	ca,za[A]
4		nga	8		nya	12		na	16		ma	20	wa

21		xa
22		sa
23		a
24		ya
25		ra

26		la
27		xa
28		sa
29		ha
30		a

[A] If the character is not accompanied by a prefixed character or a superscripted consonant, it is read as aspirated and the first romanization equivalent is used; otherwise it is read as nonaspirated and the second equivalent is used.

Consonant combinations at the beginning of syllables (the list is not complete)

1		gya	3		kya, gya[A]	5		cha	7		sha
2		kya	4		zha	6		cha, zha[A]	8		lha

[A] If the combination is not accompanied by a prefixed character or a superscripted consonant, it is read as aspirated and the first romanization equivalent is used; otherwise it is read as nonaspirated and the second equivalent is used.

---

[35] *Fourth United Nations Conference on the Standardization of Geographical Names*. Geneva, 24 August - 14 September 1982. Vol. II. *Technical papers*, pp. 121-125.

Vowel characters (ᵀ stands for any consonant character)

1 ᵀ a   2 ᵀ i   3 ᵀ u   4 ᵀ e   5 ᵀ o

Other symbols

˙ syllable separation character, e.g. ᩒᩔ Nag Qu.

Syllable endings

1	a	21	iu	41	um	61	êm
2	au	22	ig	42	um	62	êr
3	ag	23	ig	43	ur	63	ên
4	ag	24	i	44	ü	64	o
5	ang	25	i	45	ü	65	og
6	ang	26	ing	46	ü	66	og
7	ab	27	ing	47	ü	67	ong
8	ab	28	ib	48	ün	68	ong
9	am	29	ib	49	ê	69	ob
10	am	30	im	50	ê	70	ob
11	ar	31	im	51	ê	71	om
12	ai [ä]	32	ir	52	êg	72	om
13	ai [ä]	33	in	53	êg	73	or
14	ai [ä]	34	u	54	ê	74	oi [ö]
15	ai [ä]	35	ug	55	ê	75	oi [ö]
16	ain [än]	36	ug	56	êng	76	oi [ö]
17	i	37	ung	57	êng	77	oi [ö]
18	i	38	ung	58	êb	78	oin [ön]
19	i	39	ub	59	êb		
20	iu	40	ub	60	êm		

Notes

1. Prefixed consonant characters ᵀ (k), ᵀ (t), ᵀ (p), ᵀ (m), ᵀ (a), and superscripted consonant characters ᵀ (r), ᵀ (l), ᵀ (s) remain mostly unpronounced and are not reflected in the romanization.
2. Subscript characters ᵀ (w), ᵀ (y), ᵀ (r), ᵀ (l) are pronounced only in certain combinations (see also the consonant combinations at the beginning of syllables).
3. A number of homonyms exist in the Tibetan language and they are not all listed here.
4. The characters ᵀ and ᵀ are treated as without initials when transcribed with Pinyin.
5. When the prefixed characters ᵀ and ᵀ or the superscripted consonant ᵀ give the preceding

syllable an additional nasal or make the vowel in it nasalized, it is spelt according to the actual pronunciation. The nasalized vowels (a), (o), (u) are represented by the finals *an, on, un* in the Pinyin system.

6. Where two Roman equivalents are given, the second (in brackets) is used for recording the pronunciation of place-names while the first form is for general use.

Uighur

The United Nations resolution III/8 in 1977 recognized the Scheme for a Chinese Phonetic Alphabet (Pinyin) as China's official Roman alphabet scheme and recommended the alphabet as the international system for the romanization of Chinese geographical names. In China Uighur geographical names are transcribed directly from the Uighur script. The scheme was published in *Toponymic Guidelines for Map and Other Editors: China*, 1982[36].

The system is used in China and in international cartographic products.

Uighur (Uygur) uses the Perso-Arabic script which is written from right to left. The romanization is based on the Scheme for a new Uygur script. The letters of that scheme which had no counterparts in the Pinyin system, such as ŋ, ḥ, ḳ, ə, ө, ӽ were replaced by *g, h, k, a* (or *e*), *o* and *y* (or *j*). The scheme is unambiguous but in the form intended for general use does not differentiate some of the Uighur characters.

Romanization

1	ئا, ا[A]	a	9	د	d	17	ق	k [k̂]	25	ئو, و[A]	o
2	ئە, ە[A]	a, e[B] [ä]	10	ر	r	18	ك	k	26	ئۇ, ۇ[A]	u
3	ب	b	11	ز	z	19	گ	g	27	ئۆ, ۆ[A]	o [ö]
4	پ	p	12	ژ	y, j[C]	20	ڭ	ng	28	ئۈ, ۈ[A]	ü
5	ت	t	13	س	s	21	ل	l	29	ھ	h [ḥ]
6	ج	j	14	ش	x	22	م	m	30	ئې, ې[A]	e [ê]
7	چ	q	15	غ	g [ĝ]	23	ن	n	31	ئى,ى(ـ)[A]	i
8	خ	h	16	ف	f	24	ۋ	w, v	32	ي	y

[A] Form used word-initially.
[B] *e* is used where confusion of a word's meaning arises.
[C] *y* is used at the beginning of a syllable, and *j* at the end.

---

[36] *Fourth United Nations Conference on the Standardization of Geographical Names*. Geneva, 24 August - 14 September 1982. Vol. II. *Technical papers*, pp. 121-125.

Note. Where several Roman equivalents are given, the one in brackets is used for recording the pronunciation of place-names while the other forms are for general use.

Urdu

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports[37].

There is no evidence of the use of the system either in Pakistan, India or in international cartographic products. Instead, in Pakistan the Hunterian system is officially used[38]. The resolutions III/12 (1977) and IV/17 (1982) recommended association, inter alia, with Pakistan, in carrying out further studies on the system.

Urdu (Urdū) uses the Perso-Arabic script which is written from right to left. In the script vowel points are usually omitted which makes it difficult to obtain uniform romanizations. Some of the Arabic consonants are undifferentiated in romanization which means that the system is not fully reversible.

Romanization

For the initial, medial and final forms of the principal characters in the romanization table below see the Arabic section of the report.

1	ا	a, i, u	11	د	d	21	ض	z̤	31	م	m
2	ب	b	12	ڈ	ḍ	22	ط	t̤	32	ن، ں	n
3	پ	p	13	ذ	z̤	23	ظ	z̤	33	و	v, w[B]
4	ت	t	14	ر	r	24	ع	ʻ[A]	34	ہ (ھ)	h, ḥ[C]
5	ٹ	ṭ	15	ڑ	ṛ	25	غ	gh	35	ھ	h[D]
6	ث	s	16	ز	z	26	ف	f	36	ی	y
7	ج	j	17	ژ	y	27	ق	q	37	ے	e, ai[E]
8	چ	ch	18	س	s	28	ک	k			
9	ح	ḥ	19	ش	sh	29	گ	g			
10	خ	kh	20	ص	s	30	ل	l			

[A] If preceded by short *a*, it is romanized ʻā, e.g. معمول M'āmul.
[B] When و is imperceptible, e.g. in a few words of Persian origin when preceded by خ (kh).
[C] Word-finally after a short vowel.
[D] Marks aspiration of consonants.

---

[37] *Second United Nations Conference on the Standardization of Geographical Names.* London, 10-31 May 1972. Vol. II. *Technical papers.* United Nations. New York 1974, pp. 155-161; *Third United Nations Conference on the Standardization of Geographical Names.* Athens, 17 August - 7 September 1977. Vol. II, *Technical papers,* pp. 393 etc.

[38] See e.g. *Geographical Names Romanization in Pakistan.* UNGEGN, 18th Session. Geneva, 12-23 August 1996. Working Papers No. 85 and No. 85 Add. 1.

^E The character ے is used only word-finally.
Note. The Arabic ه has two distinct shapes in Urdu: ہ (ﮧ) and ھ (ﮬ). The latter denotes aspiration after consonants b, ch, d, ḍ, j, k, p, ṙ, t, ṭ.

Vowels

1	اَ	a	6	اِ	i	11	ـَے, ـَیْ	ai^A	16	اّ	(C)
2	آ	ā	7	ـِیْ	ī	12	وْ	o			
3	ـیٰ	ā	8	اُ	u	13	ـَوْ	au			
4	ـیٰ	ā	9	ـُوْ	ū	14	ا	(B)			
5	آ	ā	10	ـِے, ـِیْ	e^A	15	اـ	- (hyphen)			

^A The character ے is used word-finally.
^B Not romanized; marks absence of the vowel.
^C Doubling of the consonant letter.

Other systems of romanization

The **Hunterian system** of romanization marks the following sounds differently:

UN	Hunterian							
-ah (ه)	= a	g̲h̲	= gh	ṇ	= n	v	= w, v^B	
ā	= ā, a^A	ḥ	= h	ṙ	= r	ẏ	= zh	
ḍ	= d	ī	= ī, i^A	ṭ	= t			
ḍh	= dh	k̲h̲	= kh	ṭh	= th			

^A Word-finally.
^B Exceptionally, e.g. before i.

OTHER LANGUAGES/SCRIPTS

Armenian

No romanization systems for Armenian have been put forward at the United Nations conferences on the standardization of geographical names or at sessions of the United Nations Group of Experts on Geographical Names. No information has been received on whether there exist any national systems of romanization in Armenia.

Systems of romanization

The **BGN/PCGN 1981 System** gives the following romanization:

1	Ա ա	a	11	Ի ի	i	21	Յ յ	y	31	Տ տ	t
2	Բ բ	b	12	Լ լ	l	22	Ն ն	n	32	Ր ր	r
3	Գ գ	g	13	Խ խ	kh	23	Շ շ	sh	33	Ց ց	ts'
4	Դ դ	d	14	Ծ ծ	ts	24	Ո ո	o, vo[B]	34	Ււ ււ	u
5	Ե ե	e, ye[A]	15	Կ կ	k	25	Չ չ	ch'	35	Փ փ	p'
6	Զ զ	z	16	Հ հ	h	26	Պ պ	p	36	Ք ք	k'
7	Է է	e	17	Ձ ձ	dz	27	Ջ ջ	j	37	Եւ և	ev, yev[C]
8	Ը ը	y	18	Ղ ղ	gh	28	Ռ ռ	rr	38	Օ օ	o
9	Թ թ	t'	19	Ճ ճ	ch	29	Ս ս	s	39	Ֆ ֆ	f
10	Ժ ժ	zh	20	Մ մ	m	30	Վ վ	v			

[A] *ye* is romanized initially and after the vowel characters ա, ե, է, ը, ի, ո, ււ, and o.

[B] *vo* is romanized initially except in the word ով which should be romanized *ov*.

[C] *yev* is romanized initially, in isolation, and after the vowel characters ա, ե, է, ը, ի, ո, ււ, and o.

Note. In newer sources the character 37 is used in the lower case form only; in the upper case it is written as a combination of characters 5 and 30 (եվ), cf. ԵՐԵՎԱՆ *YEREVAN*, Երևան *Yerevan*. Also there is some indication that in 1997 the alphabet was modified so that the character ււ *w* was restored as an independent character. According to the same source the end of the alphabet would read as follows (starting from the point of difference with the above table):

34	ււ	w	36	Ք ք	k'	38	Օ օ	o	40	Ււ ււ	u
35	Փ փ	p'	37	և	ev/yev	39	Ֆ ֆ	f			

The transliteration standard **ISO 9985:1996** gives, as a single block, different Roman equivalents to the following Armenian letters (the BGN/PCGN equivalents are in parentheses):

ե (e/ye)	e	ծ (ts)	ç	ո (o/vo)	o	ււ (u)	ow
է (e)	ē	ձ (dz)	j	չ (ch')	č	o (o)	ò
ը (y)	ë	ղ (gh)	ġ	ջ (j)	ǰ		
ժ (zh)	ž	ճ (ch)	č̣	ռ (rr)	ṙ		
խ (kh)	x	շ (sh)	š	ց (ts')	c'		

Burmese

No romanization systems for Burmese have been put forward at the United Nations conferences on the standardization of geographical names or at sessions of the United Nations Group of Experts on Geographical Names. No information has been received on whether there exist any official systems of romanization in Myanmar. In 1989 the government of Myanmar changed the spellings of the names of several cities in Myanmar to bring them closer to Burmese pronunciation but there appears to be no new romanization schemes behind the changes.

The only known romanization system of Burmese is that of the BGN and the PCGN (1970) which is based on the 1907 version of the *Tables for the Transliteration of Burmese into English*, published in 1908 by the Office of the Superintendent, Government Printing, Rangoon, Burma.

Burmese uses an alphasyllabic script. The romanization is not reversible to its original script form.

System of romanization

The **BGN/PCGN 1970 System** is given below.

Consonant characters

1 က	ka[A]	6 ဝ	sa[A]	11 ဋ	ta[A]	16 တ	ta[A]	21 ပ	pa[A]	26 ယ	ya	31 ဟ	ha
2 ခ	ka[A]	7 ဆ	sa[A]	12 ဌ	ta[A]	17 ထ	ta[A]	22 ဖ	pa[AC]	27 ရ	ya	32 ဠ	la
3 ဂ	ga	8 ဇ	za	13 ဍ	da	18 ဒ	da	23 ဗ	ba	28 လ	la	33 အ	a[E]
4 ဃ	ga	9 ဈ	za	14 ဎ	da	19 ဓ	da	24 ဘ	ba	29 ဝ	wa		
5 င	nga	10 ည	nya[B]	15 ဏ	na	20 န	na	25 မ	ma	30 သ	tha[D]		

[A] See note 1.
[B] Syllable-final variant character: ဉ.
[C] ဖ is romanized *sha*.
[D] Syllable-final variant character: ဿ.
[E] See note 3.

Independent vowel characters (see also note 4)

1 ဣ	i	3 ဥ	u	5 ဦး	u	7 ဧ	e
2 ဤ	i	4 ဦ	u	6 ဩ	e	8 ဪ	aw

Abbreviated vowel characters (သ stands for any consonant character)

1 သံ	an	3 သိ	i	5 သု, သုံ	u	7 သဲ	è	9 သော်	aw	11 ကော် aw	13 ကေ e
2 သာ, သါ	a	4 သီ	i	6 သူ, သူ	u	8 ကော aw	10 ကော် aw	12 သို	o		

Other symbols (သ stands for any consonant character)

1 ျ	-y[A]	2 ြ	-y[B]	3 ွ	-w	4 ှ	h-[C]	5 စ်	in-[D]	6 စ်	(E)

[A] Exceptional romanization: ဈ *cha* (see also note 1).

ᴮ Exceptional romanization: ᥅ *cha* (see also note 1).

ᶜ Adds *h* before the consonant, e.g. ᥕ *hma*; exceptional romanization: ᥐ *sha*.

ᴰ Adds *-in* to the end of the preceding syllable: �’ဘ�__:ᗑ *Bin-gala Aw*.

ᴱ End-of-syllable mark (marks the absence of the inherent vowel).

Symbols 1, 2 and 3 may also occur in combinations: ᗑᗑ *-yw*, ᗑᗑ *-yw*.

Syllable endings (သ stands for any initial consonant character)

#		rom.	#		rom.	#		rom.	#		rom.
1	သက်	et	17		eik	33		ein	49		un
2		eik	18		ôk	34		ôn	50		wun
3		ôk	19		ut	35		un	51		è
4		aik	20		wut	36		wun	52		an
5		auk	21		it	37		an	53		aing
6		et	22		at	38		ein	54		at
7		in	23		wut	39		ôn	55		eik
8		ein	24		wut	40		aing	56		in
9		aing	25		at	41		at	57		an
10		aung	26		eik	42		eik	58		ein
11		it	27		ôk	43		ôk	59		ôn
12		ôk	28		at	44		ut	60		un
13		it	29		at	45		wut	61		wun
14		eik	30		eik	46		an	62		at
15		i,in,eᴬ	31		at	47		ein			
16		at	32		an	48		ôn			

ᴬ The choice of romanization depends on the actual pronunciation.

Notes

1. Certain consonant sounds change their value if used after any vowel letter or nasal consonant except as cited in note 5. The changes, as expressed in Roman letters, are the following:

    ch → gy
    k → g
    p → b
    s → z
    t → d

2. In romanization, *-a* is always accompanying the Burmese consonant character unless this has got a dependent vowel character or an end-of-syllable mark: _____ *madama*.

3. သ is romanized at the beginning of a word as *a* if followed by a consonant character that does not bear a vowel character or an end-of-syllable mark; at the beginning of a medial or final syllable it should be rendered by a hyphen: ____ *ma-u*, _____ *Thibin-aing*.

4. The independent vowel characters should be romanized with a hyphen at the beginning of medial and final syllables: ____ *kye-e*.

5. In consonant clusters the first component is romanized according to the main table, not according to the table of syllable endings: ___ *missa*

6. When the letter *n* at the end of a syllable is followed by *g* or *y* at the beginning of the next syllable, hyphen should be added in between them; similarly hyphen should be used between *t* and *h* if they are in different syllables: _____ *in-gut* (but ____ *shwengan*).

7. Tone marks . and : should not be romanized.

Byelorussian

No romanization systems for Byelorussian (Belarusian) have been put forward at the United Nations conferences on the standardization of geographical names or at sessions of the United Nations Group of Experts on Geographical Names. In Belarus there have been several projects of romanization, the most notable until recently being a scheme of 1992-1993, based on GOST 1983 (for Russian) and used by the national cartographic authority on maps of Belarus.

A new national system of romanization for Byelorussian was approved by the Government Committee on Land Resources, Geodesy and Cartography, Republic of Belarus, on 23 November, 2000[39] modifying an earlier system of 20 March, 1998. While the system is still based on GOST 1983, it takes more precisely into account the peculiarities of the Byelorussian orthography. The system is reversible though there may exist some ambiguous consonant combinations.

Systems of romanization

The **national system of romanization (2000)** is as follows.

1	A a	a	10	I i	i	19	C c	s	28	'	*(not*
2	Б б	b	11	Й й	j	20	Т т	t			*romanized)*
3	В в	v	12	К к	k	21	У у	u	29	Ы ы	y
4	Г г	h	13	Л л	l	22	Ў ў	ú	30	Ь ь	'
5	Д д	d	14	М м	m	23	Ф ф	f	31	Э э	e
6	E e	je^A, ie^B	15	Н н	n	24	X x	ch	32	Ю ю	ju^A, iu^B
7	Ё ё	jo^A, io^B	16	O o	o	25	Ц ц	c	33	Я я	ja^A, ia^B
8	Ж ж	ž	17	П п	p	26	Ч ч	č			
9	З з	z	18	P p	r	27	Ш ш	š			

^{A} Initially, after vowels, apostrophe, ь and ў.

^{B} After consonants.

Note. Cursive forms of some characters might be formed differently: *Aa Бб Вв Гг Дд Ее Ёё Жж Зз Іі Йй Кк Лл Мм Нн Оо Пп Рр Сс Тт Уу Ўў Фф Хх Цц Чч Шш ' Ыы Ьь Ээ Юю Яя.*

The **BGN/PCGN 1979 System** gives, as a single block, different Roman equivalents to the following Byelorussian characters (the national equivalents are in parentheses):

e (je/ie)	ye	й (j)	y	ц (c)	ts	'	"
ё (jo/io)	yo	ў (ú)	w	ч (č)	ch	ю (ju/iu)	yu
ж (ž)	zh	x (ch)	kh	ш (š)	sh	я (ja/ia)	ya

---

[39] Published in *Nacional'nyj Reestr Pravovyh Aktov Respubliki Belarus'* (National Registry of Legal Acts, Republic of Belarus) No. 3, 11 January 2001

Dzongkha

No romanization systems for Dzongkha have been put forward at the United Nations conferences on the standardization of geographical names or at sessions of the United Nations Group of Experts on Geographical Names.

On 16 September 1991, a phonological romanization for Dzongkha was introduced as an official standard by the Chairman of the Dzongkha Development Commission, Minister of Finance. Roman Dzongkha, as it was called, was not intended to replace traditional Bhutanese writing, but to accurately and adequately represent the phonology of the living language and to serve as a standard for representing Dzongkha names and words in the international media. The system was further elaborated and published as a handbook *Guide to Roman Dzongkha* in 1994.

Dzongkha uses the Tibetan script (*'Ucen*). Relationship between the script and the pronunciation is complex and the phonological transcription cannot be reverted to its original script form.

System of romanization

The **official romanization (1994)** is as follows[40].

Consonant characters

1	ka	9	ta	17	tsa	25	ra
2	kha	10	tha	18	tsha	26	la ('la)
3	g'a (ga)	11	d'a (da)	19	dza	27	sha
4	nga ('nga)	12	na ('na)	20	wa	28	sa
5	ca	13	pa	21	zh'a (zha)	29	ha
6	cha	14	pha	22	z'a (za)	30	'a
7	j'a (ja)	15	b'a (ba)	23	a		
8	nya ('nya)	16	ma ('ma)	24	ya		

Some characters have two romanizations; in that case, as general, the romanization in parentheses is to be used if the character is accompanied by a prefixed character or a superscripted consonant. Apostrophe by plosives signifies devoiced or soft pronunciation; by nasals, liquids and vowels it indicates a high register tone with abrupt glottal onset.

Syllable-initial consonant combinations (the list is not complete)
Where there are multiple romanizations, the romanization in parentheses is to be generally used if the character is accompanied by a prefixed character or a superscripted consonant.

---

[40] Adapted from *Dzongkha* by George van Driem with the collaboration of Karma Tshering of Gaselô. Research School CNWS, School of Asian, African, and Amerindian Studies. Leiden, The Netherlands 1998.

1	ཀ	k, c^A	6	བྱ	j' (j), bj' (bj)	11	ཁྲ	tr	16	ད	dr
2	ཁ	kh, ch^A	7	ཉ	ny	12	ཐྲ	thr	17	ཧྲ	hra
3	ག	g' (g), j' (j)^A	8	ཀྲ	tr	13	དྲ	dr' (dr)	18	དྭ	wa
4	ཅ	c, pc	9	ཁྲ	thr	14	པྲ	tr	19	ཌ	d
5	ཆ	ch, pch	10	གྲ	dr' (dr)	15	ཕྲ	thr	20	ལྷ	lh

^A Palatal variants *j, j', c, ch* are generally used before *a (â), o (ô)* and *u (û)*.

Vowel characters (ཀ stands for any consonant character)

1	ཀ	a	2	ཀི	i	3	ཀུ	u	4	ཀེ	e	5	ཀོ	o

Other symbols
' syllable separation character

Syllable endings (the list is not complete)

1		a	20		î, i, ik	39		um	58		em
2		â, a, ak	21		î, i	40		û	59		ê
3		â, a, ak	22		î, i	41		ü	60		en
4		a, ang	23		i, ing	42		ü	61		o, ô
5		a, ang	24		i, ing	43		ü	62		ô, ok
6		ap	25		ip	44		ü	63		ô, ok
7		ap	26		ip	45		ü	64		o, ong
8		am	27		im	46		e, i	65		o, ong
9		am	28		im	47		ê	66		op
10		â	29		î	48		iu	67		op
11		ai, ê	30		in	49		ê, ek	68		om
12		ä	31		u, û	50		ê, ek	69		om
13		ê, ä	32		û, u, uk	51		ê	70		ô
14		ä, â	33		û, u, uk	52		ê	71		ö
15		en	34		u, ung	53		e, eng	72		ö
16		i	35		u, ung	54		e, eng	73		ö, ô
17		î, i	36		up	55		ep	74		ön
18		iu	37		up	56		ep			
19		î, i, ik	38		um	57		em			

The choice of variants depends on pronunciation. There are often exceptions to the equivalents above. The table was compiled on the basis of the rules expressed in the handbook of Dzongkha

and with regard to the examples given. In addition there exist more complex syllables with consonants ᓓ *p*, ᓐ *m* or ᓇ *w* at the end, e.g. ᐅᓈᑉ *ap*, ᐄᓈᑉ *ip*, ᐆᓈᑉ *op*, ᐅᓵᑉ *âp*, ᐄᓵᑉ *îp*, ᐁᓵᑉ *êp*, ᐅᓵᒻ *âm*, ᐅᓈᐤ *û*, etc. These are not covered here.

Georgian

No romanization systems for Georgian have been put forward at the United Nations conferences on the standardization of geographical names or at sessions of the United Nations Group of Experts on Geographical Names.

According to preliminary information, in early 2002 the State Department of Geodesy and Cartography of Georgia had adopted a romanization system that employs digraphs and the apostrophe for the expression of specific Georgian sounds. This system had been utilized already earlier on e.g. driving licenses (1998).

Systems of romanization

The **national system of romanization (2002)** is as follows:

№	Georgian	Roman		№	Georgian	Roman		№	Georgian	Roman		№	Georgian	Roman
1	ა	a		10	კ	k'		19	ტ	t'		28	ძ	dz
2	ბ	b		11	ლ	l		20	უ	u		29	წ	ts'
3	გ	g		12	მ	m		21	ფ	p		30	ჭ	ch'
4	დ	d		13	ნ	n		22	ქ	k		31	ხ	kh
5	ე	e		14	ო	o		23	ღ	gh		32	ჯ	j
6	ვ	v		15	პ	p'		24	ყ	q'		33	ჰ	h
7	ზ	z		16	ჟ	zh		25	შ	sh				
8	თ	t		17	რ	r		26	ჩ	ch				
9	ი	i		18	ს	s		27	ც	ts				

The **BGN/PCGN 1981 System** is almost identical to the national system above but differs radically in the use of the apostrophe. While in the national system the apostrophe signifies an abruptive sound in contrast to the respective aspirated sound, in the BGN/PCGN system the apostrophe marks the aspirated sound vs. the unmarked abruptive sound. The differences are as follows (the national equivalents are in parentheses):

თ (t)	t'		ტ (t')	t		ყ (q')	q		წ (ts')	ts
კ (k')	k		ფ (p)	p'		ჩ (ch)	ch'		ჭ (ch')	ch
პ (p')	p		ქ (k)	k'		ც (ts)	ts'			

The transliteration standard **ISO 9984:1996** gives, as a single block, different Roman equivalents to the following Georgian characters (the national equivalents are in parentheses):

თ (t)	t'		ფ (p)	p'		ჩ (ch)	č'		ხ (kh)	x
კ (k')	k		ქ (k)	k'		ც (ts)	c'		ჯ (j)	ǰ
პ (p')	p		ღ (gh)	ḡ		ძ (dz)	j			
ჟ (zh)	ž		ყ (q')	q		წ (ts')	c			
ტ (t')	t		შ (sh)	š		ჭ (ch')	č			

Japanese

No romanization system for Japanese has been approved at the United Nations conferences on the standardization of geographical names. Systems for the romanization of Japanese have been presented at several sessions of the United Nations Group of Experts on Geographical Names (UNGEGN).

In Japan there are two current systems for the romanization of Japanese. The official romanization, also known under the name *Kunrei-siki*, was adopted on 9 December, 1954 (Cabinet Notification No. 1, table 1). Another system, the **Modified Hepburn System** (*Hebon-shiki*), continues to be used in Japan in certain spheres where it is estimated that a change would seriously affect customs and considerations of international relations, e.g. in passports. In international cartographic products the Modified Hepburn System remains the most used system.

Japanese is written in two forms: in *kanji*, which are Sino-Japanese characters, and in *kana*, which are syllabic symbols. There are two styles of *kana* – *katakana*, the squared form, and *hiragana*, the cursive form. *Katakana* symbols are used primarily for the phonetic transcription of foreign words and as the equivalent of italics. *Hiragana* symbols are used to represent grammatical particles and suffixes.

Systems of romanization

Below is a list of Japanese syllables with Roman equivalents. In the table line A gives katakana, line B hiragana and line C romanization in Kunrei-siki. In line D romanization in the Modified Hepburn System in italics is given, if different from the Kunrei-siki. Superscript numbers added to Japanese syllable characters or romanized letters refer to the notes at the end.

Combined forms:

		1	2	3	4	5	6	7	8	9	10	11
1	A	ア	イ	ウ	エ	オ						オゥ (オオ[5])
	B	あ	い	う	え	お						おぅ (おお[5])
	C	a	i	u	e	o						ô
	D											ō
2	A	カ	キ	ク	ケ	コ	キャ	キュ	キュゥ	キョ	キョゥ	コゥ
	B	か	き	く	け	こ	きゃ	きゅ	きゅぅ	きょ	きょぅ	こぅ
	C	ka	ki	ku	ke	ko	kya	kyu	kyû	kyo	kyô	kô
	D				*ke (ga[4])*				*kyū*		*kyō*	*kō*
3	A	サ	シ	ス	セ	ソ	シャ	シュ	シュゥ	ショ	ショゥ	ソゥ
	B	さ	し	す	せ	そ	しゃ	しゅ	しゅぅ	しょ	しょぅ	そぅ
	C	sa	si	su	se	so	sya	syu	syû	syo	syô	sô
	D		*shi*				*sha*	*shu*	*shū*	*sho*	*shō*	*sō*
4	A	タ	チ	ツ	テ	ト	チャ	チュ	チュゥ	チョ	チョゥ	トゥ
	B	た	ち	つ	て	と	ちゃ	ちゅ	ちゅぅ	ちょ	ちょぅ	とぅ
	C	ta	ti	tu[1]	te	to	tya	tyu	tyû	tyo	tyô	tô
	D		*chi*	*tsu*			*cha*	*chu*	*chū*	*cho*	*chō*	*tō*
5	A	ナ	ニ	ヌ	ネ	ノ	ニャ	ニュ	ニュゥ	ニョ	ニョゥ	ノゥ
	B	な	に	ぬ	ね	の	にゃ	にゅ	にゅぅ	にょ	にょぅ	のぅ
	C	na	ni	nu	ne	no	nya	nyu	nyû	nyo	nyô	nô
	D								*nyū*		*nyō*	*nō*

#												
6	A	ハ	ヒ	フ	ヘ	ホ	ヒャ	ヒュ	ヒュウ	ヒョ	ヒョウ	ホゥ
	B	は	ひ	ふ	へ	ほ	ひゃ	ひゅ	ひゅう	ひょ	ひょう	ほぅ
	C	ha	hi	hu	he	ho	hya	hyu	hyû	hyo	hyô	hô
	D	*ha (wa[4])*		*fu*	*he (e[4])*				*hyū*		*hyō*	*hō (ō[4])*
7	A	マ	ミ	ム	メ	モ	ミャ	ミュ	ミュウ	ミョ	ミョウ	モゥ
	B	ま	み	む	め	も	みゃ	みゅ	みゅう	みょ	みょう	もぅ
	C	ma	mi	mu	me	mo	mya	myu	myû	myo	myô	mô
	D								*myū*		*myō*	*mō*
8	A	ヤ		ユ		ヨ						ヨゥ
	B	や		ゆ		よ						よぅ
	C	ya		yu		yo						yô
	D											*yō*
9	A	ラ	リ	ル	レ	ロ	リャ	リュ	リュウ	リョ	リョウ	ロゥ
	B	ら	り	る	れ	ろ	りゃ	りゅ	りゅう	りょ	りょう	ろぅ
	C	ra	ri	ru	re	ro	rya	ryu	ryû	ryo	ryô	rô
	D								*ryū*		*ryō*	*rō*
10	A	ワ				ヲ						
	B	わ				を						
	C	wa				o						
	D											
11	A	ン										
	B	ん										
	C	n[2]										
	D	*n, n', m*										
12	A	ガ	ギ	グ	ゲ	ゴ	ギャ	ギュ	ギュウ	ギョ	ギョウ	ゴゥ
	B	が	ぎ	ぐ	げ	ご	ぎゃ	ぎゅ	ぎゅう	ぎょ	ぎょう	ごぅ
	C	ga	gi	gu	ge	go	gya	gyu	gyû	gyo	gyô	gô
	D								*gyū*		*gyō*	*gō*
13	A	ザ	ジ	ズ	ゼ	ゾ	ジャ	ジュ	ジュウ	ジョ	ジョウ	ゾゥ
	B	ざ	じ	ず	ぜ	ぞ	じゃ	じゅ	じゅう	じょ	じょう	ぞぅ
	C	za	zi	zu	ze	zo	zya	zyu	zyû	zyo	zyô	zô
	D		*ji*				*ja*	*ju*	*jū*	*jo*	*jō*	*zō*
14	A	ダ	ヂ	ヅ	デ	ド						ドゥ
	B	だ	ぢ	づ	で	ど						どぅ
	C	da	di	du	de	do						dô
	D		*ji*	*zu*								*dō*
15	A	バ	ビ	ブ	ベ	ボ	ビャ	ビュ	ビュウ	ビョ	ビョウ	ボゥ
	B	ば	び	ぶ	べ	ぼ	びゃ	びゅ	びゅう	びょ	びょう	ぼぅ
	C	ba	bi	bu	be	bo	bya	byu	byû	byo	byô	bô
	D								*byū*		*byō*	*bō*
16	A	パ	ピ	プ	ペ	ポ	ピャ	ピュ	ピュウ	ピョ	ピョウ	ポゥ
	B	ぱ	ぴ	ぷ	ぺ	ぽ	ぴゃ	ぴゅ	ぴゅう	ぴょ	ぴょう	ぽぅ
	C	pa	pi	pu	pe	po	pya	pyu	pyû	pyo	pyô	pô
	D								*pyū*		*pyō*	*pō*

Notes

1. A small-script *tu/tsu* form (ッ / っ) is inserted between kana symbols to indicate a double consonant (*kk, ss, tt, pp* in Kunrei-siki; *kk, ss, ssh, tt, tts, tch, pp* in Hepburn).

2. ン/ ん is romanized always *n* in Kunrei-siki; when it is necessary to separate the sound *n* from

the vowel or *y* to follow, the apostrophe is added after the *n*. In modified Hepburn the character is romanized *n'* before *y* or a vowel letter, *n* in all other cases; earlier also *m* was used before *b, p,* or *m*.

3. Long vowels are expressed in Kunrei-siki by placing a circumflex (^) over a vowel; in Hepburn long vowels are indicated by a macron (‾).
4. The romanization in parentheses (in modified Hepburn) is used only in those cases where the kana symbol is known to be pronounced in the manner indicated.
5. The combination in parentheses is used to denote the word meaning 'big, great'.

Kazakh

No romanization systems for Kazakh have been put forward at the United Nations conferences on the standardization of geographical names or at sessions of the United Nations Group of Experts on Geographical Names. It is not known whether there exist any national systems of romanization in Kazakhstan. According to the information received in March 2001 from the Kazakh Linguistics Institute in Almaty, a decree was issued by the President of Kazakhstan in 1998 on the necessity to prepare a new Kazakh alphabet based on the Roman script. Subsequently a committee was formed to further this aim. There have been several draft alphabets, the aim is to work on this task systematically, taking into consideration also the experience of other countries in the transition to the Roman alphabet.

Systems of romanization

The **BGN/PCGN 1979 System** gives the following romanization:

1	А а	a	12	И и	ī	23	Р р	r	34	Ш ш	sh
2	Ә ә	ä	13	Й й	y	24	С с	s	35	Щ щ	shch
3	Б б	b	14	К к	k	25	Т т	t	36	Ъ ъ	"
4	В в	v	15	Қ қ	q	26	У у	ū	37	Ы ы	y
5	Г г	g	16	Л л	l	27	Ұ ұ	u	38	І і	i
6	Ғ ғ	gh	17	М м	m	28	Ү ү	ü	39	Ь ь	'
7	Д д	d	18	Н н	n	29	Ф ф	f	40	Э э	è
8	Е е	e	19	Ң ң	ng	30	Х х	kh	41	Ю ю	yu
9	Ё ё	yo	20	О о	o	31	һ һ	h	42	Я я	ya
10	Ж ж	zh	21	Ө ө	ö	32	Ц ц	ts			
11	З з	z	22	П п	p	33	Ч ч	ch			

The transliteration standard **ISO 9:1995** gives, as a single block, different Roman equivalents to the following Kazakh characters (the BGN/PCGN equivalents are in parentheses):

ә (ä)	ä̌	қ (q)	ķ	x (kh)	h	i (i)	ì
ғ (gh)	ġ	ң (ng)	ņ	һ (h)	ḥ	э (è)	è
ё (yo)	ë	ө (ö)	ô	ц (ts)	c	ю (yu)	û
ж (zh)	ž	у (ū)	u	ч (ch)	č	я (ya)	â
и (ī)	i	ұ (u)	ů	ш (sh)	š		
й (y)	j	ү (ü)	ù	щ (shch)	ŝ		

Kirghiz

No romanization systems for Kirghiz (Kyrgyz) have been put forward at the United Nations conferences on the standardization of geographical names or at sessions of the United Nations Group of Experts on Geographical Names. No information has been received on whether there exist any national systems of romanization in Kyrgyzstan.

Systems of romanization

The **BGN/PCGN 1979 System** gives the following romanization:

1	А а	a	10	И и	i	19	П п	p	28	Ч ч	ch
2	Б б	b	11	Й й	y	20	Р р	r	29	Ш ш	sh
3	В в	v	12	К к	k	21	С с	s	30	Щ щ	shch
4	Г г	g	13	Л л	l	22	Т т	t	31	Ъ ъ	"
5	Д д	d	14	М м	m	23	У у	u	32	Ы ы	y
6	Е е	e	15	Н н	n	24	Ү ү	ü	33	Ь ь	'
7	Ё ё	yo	16	Ң ң	ng	25	Ф ф	f	34	Э э	e
8	Ж ж	j	17	О о	o	26	Х х	kh	35	Ю ю	yu
9	З з	z	18	Ө ө	ö	27	Ц ц	ts	36	Я я	ya

The transliteration standard **ISO 9:1995** gives, as a single block, different Roman equivalents to the following Kirghiz characters (the BGN/PCGN equivalents are in parentheses):

ё (yo)	ё	ө (ö)	ô	ч (ch)	č	ю (yu)	û
ж (j)	ž	ү (ü)	ù	ш (sh)	š	я (ya)	â
й (y)	j	х (kh)	h	щ (shch)	ŝ		
ң (ng)	ņ	ц (ts)	c	э (e)	è		

Korean

No romanization system for Korean has been approved at the United Nations conferences on the standardization of geographical names, although systems for the romanization of Korean have been presented at several sessions of the United Nations Group of Experts on Geographical Names (UNGEGN). Experts at the sessions have repeatedly expressed their wish that the Democratic People's Republic of Korea and the Republic of Korea should continue their efforts in aspiring to agree on a single international system for the romanization of Korean geographical names.

In the Democratic People's Republic of Korea there is a national system adopted in 1992 and presented to the 17th session of UNGEGN in 1994[41].

In the Republic of Korea the Ministry of Culture and Tourism adopted in July 2000 a new system

---

[41] *Guidelines for Romanization of Korean* (Democratic People's Republic of Korea). UNGEGN, 17th Session. New York, 13-24 June 1994, WP No. 57.

of romanization for Korean[42] which superseded another system approved in 1984. It is provided that the full transition to the new system will be completed before the end of 2005 but it is already compulsory since March 2002 to follow this system in new publications.

Apart from the national systems the most widely used international system is that of McCune-Reischauer (1939) which has also been adopted by the BGN and the PCGN[43].

A transliteration system that was provisionally agreed upon by the ISO experts of the Democratic People's Republic of Korea and the Republic of Korea, is given in ISO TR 11941:1996 (a provisional technical report, not enforced as a standard)[44].

In linguistics also the so-called Yale system of romanization is widely accepted[45].

Korean uses an alphabetic script in which the characters are grouped graphically together into complex syllable blocks. So, the geographical name P'anmunjŏm is written 판문점, not ㅍㅏㄴㅁㅜㄴㅈㅓㅁ.

Systems of romanization

The following table gives a comparison of all the main romanization systems. The character sequence has been taken from ISO TR 11941, there are variations to this in national usage. Romanization equivalents in the columns are as follows: **1** – ISO TR 11941, **2** – national system of the Democratic People's Republic of Korea (1992), **3** – national system of the Republic of Korea (2000), **4** – McCune-Reischauer system (1939), **5** – Yale system of romanization.

		1	2	3	4	5
1	ㄱ	k/g	k	g, k^F	k	k
2	ㅋ	kh/k	kh	k	k'	kh
3	ㄲ	kk/gg	kk	kk	kk	kk
4	ㄷ	t/d	t	d, t^F	t	t
5	ㅌ	th/t	th	t	t'	th
6	ㄸ	tt/dd	tt	tt	tt	tt
7	ㅂ	p/b	p	b, p^F	p	p
8	ㅍ	ph/p	ph	p	p'	ph
9	ㅃ	pp/bb	pp	pp	pp	pp
10	ㅈ	c/j	j	j	ch	c
11	ㅊ	ch/c	ch	ch	ch'	ch
12	ㅉ	cc/jj	jj	jj	tch	cc

---

[42] First introduced in *Recent Changes in Geographical Names of Korean Cities and other Developments* (Republic of Korea). UNGEGN, 20th Session. New York, 17–28 January 2000. WP No. 48. The full text has been published at various websites in the Republic of Korea, e.g. http://www.korea.net/kois/government/focus/romanization/4.html.

[43] *Romanization Systems and Roman-script Spelling Conventions*. Prepared by the U.S. Board on Geographic Names Foreign Names Committee Staff. Published by the Defense Mapping Agency. 1994, pp. 57-61.

[44] The summary of it was published at the Internet site http://www.phon.ucl.ac.uk/home/kskim/11941.html.

[45] See e.g. *The World's Writing Systems*. Edited by Peter T. Daniels and William Bright. New York, Oxford: Oxford University Press 1996, pp. 218-227.

13	ᄉ	s	s	s	s	s
14	ᄊ	ss	ss	ss	ss	ss
15	ᄒ	h	h	h	h	h
16	ᄋ	'A, -B, ngC	-, ngE	-, ngE	-, ngE	-, ngE
17	ᄂ	n	n	n	n	n
18	ᄅ	l, rD	r	r, lG	r, n, l	l
19	ᄆ	m	m	m	m	m
20	ᅡ	a	a	a	a	a
21	ᅥ	eo	ŏ	eo	ŏ	e
22	ᅩ	o	o	o	o	o
23	ᅮ	u	u	u	u	wu
24	ᅳ	eu	ŭ	eu	ŭ	u
25	ᅵ	i	i	i	i	i
26	ᅢ	ae	ae	ae	ae	ay
27	ᅦ	e	e	e	e	ey
28	ᅬ	oe	oe	oe	oe	oy
29	ᅣ	ya	ya	ya	ya	ya
30	ᅧ	yeo	yŏ	yeo	yŏ	ye
31	ᅭ	yo	yo	yo	yo	yo
32	ᅲ	yu	yu	yu	yu	ywu
33	ᅤ	yae	yae	yae	yae	yay
34	ᅨ	ye	ye	ye	ye	yey
35	ᅪ	wa	wa	wa	wa	wa
36	ᅯ	weo	wŏ	wo	wŏ	we
37	ᅱ	wi	wi	wi	wi	wuy
38	ᅫ	wae	wae	wae	wae	way
39	ᅰ	we	we	we	we	wey
40	ᅴ	yi	ŭi	ui	ŭi	uy

A See notes 2, 3 and 4.

B Word-initially.

C At the end of a syllable.

D At the beginning of a syllable.

E ᄋ is not romanized syllable-initially, syllable-finally it is romanized *ng*.

F The first romanization equivalent is used when occurring before a vowel, the second equivalent is used when followed by another consonant or if it forms the final sound of a word.

G ᄅ is romanized as *r* when followed by a vowel, and as *l* when followed by a consonant or when appearing at the end of a word. ᄅᄅ is romanized as *ll*.

Notes to ISO TR 11941: 1996

1. Where there are pairs of romanizations, the first corresponds to Method I and the second to Method II.

2. To avoid ambiguity the apostrophe is used to denote the beginning of the next syllable in a polysyllabic word.

3. By Method I the apostrophe is needed in the following cases:

    a. if the consonant ᄋ is at the initial position of a non-initial syllable of a polysyllabic word: 꽃이 *kkoch'i*, 강에 *kang'e*, 앉아라 *anc'ara*;

b. if the five double consonants appear at the initial positions of the non-initial syllables of polysyllabic words: 아까 *a'kka*, 흰떡 *hyin'tteok*;

c. if the four aspirated consonants appear at the initial position of the non-initial syllables of polysyllabic words: 유퀘하다 *yu'khwaehata*;

4. By Method II the apostrophe is needed in the following cases:

    a. if the consonant ⁰ is at the initial position of a non-initial syllable of a polysyllabic word: 꽃이 *ggoc'i*, 강에 *gang'e*, 앉아라 *anj'ara*;

    b. if the five double consonants appear at the initial positions of the non-initial syllables of polysyllabic words: 아까 *a'gga*, 흰떡 *hyin'ddeog*.

There are complicated rules based on pronunciation that determine the conversion of Korean syllables into Roman and the romanizations given in the table reflect only the most typical values. Most changes concern consonants which are often assimilated when used in combinations. The rules are given differently in various romanization systems and for reasons of economy these are not reproduced in this report.

Lao

No romanization system for Lao (Laotian) has been approved at the United Nations conferences on the standardization of geographical names. In 1966 the BGN and the PCGN jointly adopted a system of romanization that was based on the Lao Commission Nationale de Toponymie (CNT) system, to which a complete vowel table and supplementary notes had been added.

System of romanization

The **BGN/PCGN 1966 System** gives the following romanization[46]:

Consonant characters

1 ກ	k (k)	5 ຈ	ch	9 ດ	d (t)	14 ບ	b (p)	21 ຍ	y	25 ຫ	h^D
2 ຂ	kh	6 ສ	s	10 ຕ	t	15 ປ	p	22 ຣ	r	26 ອ	(E)
3 ຄ	kh	7 ຊ	x	11 ຖ	th	16 ຜ	ph	23 ລ	l^B	27 ຮ	h
4 ງ	ng (ng)	8 ຢ	gn (y)^A	12 ທ	th	17 ຝ	f	24 ວ	v, o^C		
				13 ນ	n (n)	18 ພ	ph				
						19 ຟ	f				
						20 ມ	m				

^{A} Character variation ຽ is used at the end of a syllable (ຂາຽ *khay*), after ຫ (ຫຽາບ *gnap*) and between consonants (ສຽງ *siang*).

^{B} Subscript variation: ຼ, e.g. ຫຼາມ *lam*.

---

[46] For full text and explanations see the original table in *Romanization Systems and Roman-script Spelling Conventions*. Prepared by the U.S. Board on Geographic Names Foreign Names Committee Staff. Published by the Defense Mapping Agency. 1994, pp. 63-67.

^C See note 4.
^D See note 2.
^E See note 3.

Lao characters are romanized in two ways depending on whether they denote syllable-initials or syllable-finals. The romanization of syllable-finals is indicated in parentheses. Some other consonant characters which are not accompanied by a vowel character form a full syllable when used within a word: ຂ *kkha*, ຄ *kkha*, ຈ *tcha*, ສ *tsa*, ຊ *tsa*, ຕ *tta*, ຖ *ttha*, ທ *ttha*, ປ *ppa*, ພ *ppha*, ຟ *pfa*.

Vocalic nuclei (• stands for any consonant character)

1	•ະ	a	12	ເ•	é	23	ເ•ຶະ	eu	34	•ົວະ	oua
2	•ັ	a	13	ແ•ະ	è	24	ເ•ຶ•	eu	35	•ົວ•	oua
3	•າ	a	14	ແ•ັ	è	25	ເ•ື	eu	36	•ົວ	oua
4	•ິ	i	15	ແ•	è	26	ເ•ືອ	eu	37	•ວ•	oua
5	•ີ	i	16	ໂ•ະ	ô	27	ເ•ັຍະ	ia	38	ໄ•	ai
6	•ຶ	u	17	•ົ•	ô	28	•ັຍ•	ia	39	ໃ•	ai
7	•ື	u	18	ໂ•	ô	29	ເ•ັຍ	ia	40	ເ•າ	ao
8	•ຸ	ou	19	ເ•າະ	o	30	•ຍ•	ia	41	•ຳ	am
9	•ູ	ou	20	•ັອ•	o	31	ເ•ືອະ	ua			
10	ເ•ະ	é	21	•ໍ	o	32	ເ•ືອ•	ua			
11	ເ•ັ	é	22	•ອ•	o	33	ເ•ືອ	ua			

Reduplicating mark

ໆ   (e.g. ແຊວໆ *xèoxèo*)

Notes

1. Tone marks ່, ້, ໊, ໋ should not be romanized.

2. At the beginning of a syllable, the character ຫ unaccompanied by a vowel or tone mark and occurring immediately before ຍ *gn*, ນ *n*, ມ *m*, ຣ *r*, ລ *l*, or ວ *v* should generally not be romanized: ຫລວງພະບາງ *Louangphrabang*. Often ligatures ຫນ, ຫມ, ໜ are used in such cases.

3. At the beginning of a word ອ should not be romanized: ອິກ *ik*. At the beginning of a syllable within a word the character should be romanized by a hyphen: ຊະເອມ *xa-ém*, ບ້ານແກ້ງອີ *Ban Kèng-I*.

4. The character ວ at the beginning of a syllable should be romanized *v*: ທັນວາ *thanva*. As the second character of a combination, ວ should be romanized *o*. At the end of a syllable the character is romanized *o* generally (ນາວ *nao*, ແກວ *kèo*, ດຽວ *diao*) but after *i* it is romanized *ou* (ຄິວ *khiou*, ຕິວ *tiou*).

5. In consonant combinations where the second character is ຣ *r*, ລ *l* or ວ *v*, it should be

romanized character by character, even if the second character may represent no pronounced sound: ບ້ານວັດພຣະໄຊ *Ban Vatphraxai*, �ຫຽຣນ *Khlitsatian*, ສວາຽ *soay*, ແຂວງ *khoèng*.

6. A vowel which is not represented in Lao writing occasionally occurs in the pronunciation of certain words. The presence of that vowel should be indicated in romanization by insertion of the letter *a*: ສວັນນະເຂດ *Savannakhét* (also written ສະຫວັນນະເຂດ), ຂວາວ *khavao* (cf. ຂວານ *khoan*).

7. In some words of foreign origin the variant ຣ̌ *r* occurs: ເບີຣ̌ *beur*.

Maldivian

No romanization system for Maldivian (Divehi) has been approved at the United Nations conferences on the standardization of geographical names.

In 1987 the Maldivian government adopted a system for the romanization of Maldivian. It was also approved by the BGN and the PCGN in 1988.

Maldivian is written from right to left. Vowels are marked as diacritical marks accompanying a consonant character.

System of romanization

The **official romanization (1987)** is as follows.
Consonants

1	✎	h	7	✎	k	13	✎	th[c]	19	✎	z
2	✎	sh	8	✎	(B)	14	✎	l	20	✎	t
3	✎	n[A]	9	✎	v	15	✎	g	21	✎	y
4	✎	r	10	✎	m	16	✎	gn	22	✎	p
5	✎	b	11	✎	f	17	✎	s	23	✎	j
6	✎	lh	12	✎	dh	18	✎	d	24	✎	ch

[A] Romanized *n'* if used without any vowel or auxiliary sign: ✎ *kan'du.*

[B] Not romanized but see note 1.

[C] Character ✎ is romanized *iy*, e.g. ✎ *aiypulhu.*

Consonants with diacritical marks (used mainly in words of Arabic origin, corresponding Arabic characters and their romanizations are shown in parentheses)

25	th'	(ث th)	29	x	(ز zh)	33	t'	(ط ṭ)	37	q	(ق q)
26	h'	(ح ḥ)	30	sh'	(ش sh)	34	z'	(ظ ẓ)	38	w	(و w)
27	kh	(خ kh)	31	s'	(ص ṣ)	35	'	(ع ')			
28	dh'	(ذ dh)	32	l'	(ض ḍ)	36	gh	(غ gh)			

Note. Romanization values for consonant characters with diacritical marks, and for variations described in notes A and C to the consonant table have been taken from the ISO draft transliteration for Thaana (Anthony P. Stone, June 26, 2000, http://ourworld.compuserve.com/homepages/stone_catend/tha02.pdf).

Vowels (• stands for any consonant character)

1	•	a	4	•	ey	7	•	o	10	•	oo
2	•	aa	5	•	i	8	•	oa			
3	•	e	6	•	ee	9	•	u			

Diacritical mark

• (not romanized; marks absence of the vowel)

Notes

1. The character ʾ is not romanized. If it bears a vowel character, that vowel character alone is romanized, e.g. *irugai*.

2. Characters ʾ and ʿ word-finally are romanized h, e.g. *rah*. In other cases, when ʾ and ʿ are used, the romanized value of the following consonant character is doubled, e.g. *batteh*.

Mongolian
(Cyrillic)

No romanization system for Mongolian Cyrillic has been approved at the United Nations conferences on the standardization of geographical names. No information has been received on whether there exist any national systems of romanization in Mongolia. The language is nowadays also written in its traditional Mongolian script which would need a different romanization approach.

System of romanization

The **BGN/PCGN 1964 System** gives the following romanization:

1	A a	a	10	И и	i	19	P p	r	28	Ш ш	sh
2	Б б	b^A	11	Й й	y^C	20	C c	s	29	Щ щ	shch^D
3	B в	v^B	12	K к	k^D	21	T т	t	30	Ъ ъ	ʼ ^{E}
4	Г г	g	13	Л л	l	22	У у	u	31	Ы ы	ï^F
5	Д д	d	14	M м	m	23	Y ү	ü	32	Ь ь	ĭ
6	E e	yö	15	H н	n	24	Ф ф	f^D	33	Э э	e
7	Ё ё	yo	16	O o	o	25	X x	h	34	Ю ю	yu, yü^G
8	Ж ж	j	17	Ө ө	ö	26	Ц ц	ts	35	Я я	ya
9	З з	dz	18	П п	p^D	27	Ч ч	ch			

^{A} The character б occurs initially and after л, м and б.

^{B} The character в occurs medially in Mongolian words and initially in foreign words.

^{C} The character й occurs only as the second element of diphthongs.

^{D} The characters к, п, ф and щ occur only in foreign words.

^{E} The character ъ occurs only in Russian words.

^{F} The character ы occurs only in suffixes to words containing the vowel character a, o or y.

^{G} In monosyllables, the character ю is romanized *yu* or *yü* depending on pronunciation; in polysyllables, it is romanized *yu* when followed by *a*, *o*, or *u*, but *yü* when followed by *i*, *e*, *ö*, or *ü*.

Pashto

No romanization system for Pashto (Afghan) has been approved at the United Nations conferences on the standardization of geographical names.

Pashto uses a Perso-Arabic script that is written from right to left. As in Arabic and Persian, vowels are marked only sporadically to indicate pronunciation, otherwise they are omitted.

System of romanization

The **BGN/PCGN 1968 System** gives the following romanization:

1	ا	-^{A}	11	ح	ḥ	21	س	s	31	ق	q
2	ب	b	12	خ	kh	22	ش	sh	32	ک	k
3	پ	p	13	د	d	23	ښ	s̲h̲	33	گ (ګ)	g
4	ت	t	14	ډ	ḍ	24	ص	ṣ	34	ل	l
5	ټ	ṯ	15	ذ	z̄	25	ض	z̤	35	م	m
6	ث	s̄	16	ر	r	26	ط	ṯ	36	ن	n
7	ج	j	17	ړ	ṛ	27	ظ	z̧	37	ڼ	ṇ
8	چ	ch	18	ز	z	28	ع	ʻ	38	و	w
9	څ	ts	19	ژ	zh	29	غ	gh	39	ه	h
10	ځ	dz	20	ږ	z̲h̲	30	ف	f	40	ی (ي)	y

^{A} Not romanized, but see romanizations accompanying *alip* (ا) in the table for vowels.

Vowel characters and diacritical marks (• stands for any consonant)

1	٠	a	6	ٚ٠	ay	11	ٚ٠	ê	16	ٚوى	ūy
2	٠ا	ā	7	ٚ٠	ey	12	ـَ	êy	17	ٱ	ā
3	٠	i	8	ى	ey	13	٠	u	18	•ٕ•	'
4	ٚى	ī	9	ﭗ	ē	14	٠و	ū	19	٠	(A)
5	ٚى	ī	10	ٚو	ew	15	٠ۆ	ow			

^A Doubling of consonant letter.

Sinhalese

No information has been received on whether there exist any national systems of romanization for Sinhalese (Sinhala) in Sri Lanka.

In 1972, D. N. Sharma prepared a romanization table for Sinhalese (Sinhala) which was published in vol. II of the conference report together with the other languages of the Indian group[47].

Sinhalese uses an alphasyllabic script whereby each character represents a syllable rather than one sound. Vowels and diphthongs are marked in two ways: as independent characters (used syllable-initially) and in an abbreviated form, to denote vowels after consonants.

System of romanization

Character variants and a list of ligatures have been added to the table prepared by D. N. Sharma.

I. Independent vowel characters

1	අ	a	5	ඉ	i	9	ඍ	ṛ	13	ඓ	ai^A
2	ආ	ā	6	ඊ	ī	10	ඎ	r̄	14	ඔ	ŏ
3	ඇ	ǣ^A	7	උ	u	11	ඏ	ĕ	15	ඕ	o
4	ඈ	ǣ^A	8	ඌ	ū	12	එ	e	16	ඖ	au

^A The character ඇ was romanized as ăi and the character ඈ as ai in Sharma's table, but there was no romanization for the character ඓ which, though rarely used, represents the ai proper of the Indian languages; here the romanization of these characters is given in accordance with the pronunciation.

---

[47] *Second United Nations Conference on the Standardization of Geographical Names*. London, 10-31 May 1972. Vol. II. Technical papers, pp. 151-152.

II. Abbreviated vowel characters and other symbols (• stands for any consonant character)

1 •	a	5 ⌐	i	9 •a	ṛ	13 ☺☺• ai	17 ○	ṁ
2 •)	ā	6 ⌐	ī	10 •aa	r̄	14 ☺•) ŏ	18 8	ḥ
3 •ᴈ	ăᴬ	7 ⌐	uᴮ	11 ☺•	ĕ	15 ☺•ʃ o	19 •ʳ	(D)
4 •ᴈ	æᴬ	8 ⌐	ūᶜ	12 ☺ʃ	e	16 ☺•ꓵ au		

ᴬ Exceptions: Ớᴈ ră, Ớᴈ ræ.

ᴮ Exceptions: ꜗ ku, ꝙ gu, ꝗ tu, ꝙ bhu, Ớᴈ ru, ꝙ shu.

ᶜ Exceptions: ꜗ kū, ꝙ gü, ꝗ tü, ꝙ bhū, Ớᴈ rü, ꝙ shū.

ᴰ Marks absence of the inherent vowel. Special forms: ꝺ kh, ꝺ ṅ, ꝺ ch, ꝺ ṭ, ꝺ ḍ, ꝺ dh, ꝺ b, ꝺ m, ꝺ v.

III. Consonant characters

1 ꝺ	ka	6 ꝺ	chaᴬ	11 ꝺ	ṭa	16 ꝺ	ta	21 ꝺ	pa	26 ꝺ	ya	31 ꝺ	shaᴬ
2 ꝺ	khaᴬ	7 ꝺ	chhaᴬ	12 ꝺ	ṭhaᴬ	17 ꝺ	tha	22 ꝺ	phaᴬ	27 ꝺ	ra	32 ꝺ	ṣhaᴬ
3 ꝺ	ga	8 ꝺ	ja	13 ꝺ	ḍa	18 ꝺ	da	23 ꝺ	ba	28 ꝺ	la	33 ꝺ	sa
4 ꝺ	ghaᴬ	9 ꝺ	jhaᴬ	14 ꝺ	ḍhaᴬ	19 ꝺ	dha	24 ꝺ	bhaᴬ	29 ꝺ	va	34 ꝺ	ha
5 ꝺ	ṅaᴬ	10 ꝺ	ñaᴬ	15 ꝺ	ṇa	20 ꝺ	na	25 ꝺ	ma	30 (ꝺ fa)		35 ꝺ	ḷa

ᴬ These consonants are used only in Sanskrit and Pali loanwords.

IV. Consonant ligatures

Examples

•ᴈ r-   ꝺ rka, ꝺ ṛta, ꝺ rma, ꝺ rva, ꝺ ṛsha.

⌐ -r   ꝺ kra, ꝺ gra, ꝺ ghra, ꝺ jra, ꝺ ṭra, ꝺ ḍra, ꝺ tra, ꝺ dra, ꝺ dhra, ꝺ pra, ꝺ bra, ꝺ bhra, ꝺ shra.

•ꝺ -y   ꝺꝺ kya, ꝺꝺ tya, ꝺꝺ vya.

Prenasalized sounds: ꝺ ṅga, ꝺ ñja, ꝺ ṇḍa, ꝺ nda, ꝺ mba.

Other ligatures: ꝺ kṣha, ꝺ jña, ꝺ dva, ꝺ nda.

Tajik

No romanization systems for Tajik have been put forward at the United Nations conferences on the standardization of geographical names or at sessions of the United Nations Group of Experts on Geographical Names. No information has been received on whether there exist any national systems of romanization in Tajikistan.

Systems of romanization

The **BGN/PCGN 1994 System** gives the following romanization:

1	A a	a	8	Ё ё	yo	15	Қ қ	q	22	C c	s
2	Б б	b	9	Ж ж	zh	16	Л л	l	23	T т	t
3	В в	v	10	З з	z	17	М м	m	24	У у	u
4	Г г	g	11	И и	i	18	Н н	n	25	Ӯ ӯ	ŭ
5	Ғ ғ	gh	12	Ӣ ӣ	í	19	О о	o	26	Ф ф	f
6	Д д	d	13	Й й	y	20	П п	p	27	Х х	kh
7	Е е	e	14	К к	k	21	Р р	r	28	Ҳ ҳ	h

| 29 | Ч ч | ch | 31 | Ш ш | sh | 33 | Э э | è | 35 | Я я | ya |
| 30 | Ҷ ҷ | j | 32 | Ъ ъ | ' | 34 | Ю ю | yu | | | |

The above table reflects the new Tajik alphabetical sequence as decreed by the Tajik government in 1998. Some of the characters were then abolished but should they appear in names they are romanized as follows:

| 36 | Ц ц | ts | 37 | Щ щ | shch | 38 | Ь ь | " | 39 | Ы ы | y |

The transliteration standard **ISO 9:1995** gives, as a single block, different Roman equivalents to the following Tajik characters (the BGN/PCGN equivalents are in parentheses):

ғ (gh)	ġ	й (y)	j	ӯ (ū)	ū	ъ (')	"
ё (yo)	ë	қ (q)	ķ	ч (ch)	č	э (è)	è
ж (zh)	ž	х (kh)	h	ҷ (j)	ç	ю (yu)	û
ӣ (í)	ī	ҳ (h)	ḩ	ш (sh)	š	я (ya)	â

Tigrinya

No romanization system for Tigrinya has been approved at the United Nations conferences on the standardization of geographical names. No information has been received on whether there exist any national systems of romanization in Eritrea.

Tigrinya, like Amharic, uses the Ethiopic syllabic script whereby each syllable denotes a combination of a consonant and a vowel. It has 35 basic characters, each character having several different forms, usually called orders, according to the vowel with which the basic symbol is combined.

System of romanization

A **BGN/PCGN 1994 Agreement** employs the BGN/PCGN 1967 System for Amharic also in the romanization of Tigrinya. Note, however, that there is one additional basic character in Tigrinya (row 10 below), and some combinations with W are romanized differently (see Note E).[48]

Combinations with W[C]:

	1(D)	2	3	4	5	6(A)	7	(1,E)	(3)	(4)	(5)	(6)
1	ሀ	ሁ	ሂ	ሃ	ሄ	ህ	ሆ					
	hā	hu	hī	ha	hē	h, hi	ho					
2	ለ	ሉ	ሊ	ላ	ሌ	ል	ሎ			ሏ		
	le	lu	lī	la	lē	l, li	lo			lwa		
3[B]	ሐ	ሑ	ሒ	ሓ	ሔ	ሕ	ሖ					
	hā	hu	hī	ha	hē	h, hi	ho					
4	መ	ሙ	ሚ	ማ	ሜ	ም	ሞ			ሟ		
	me	mu	mī	ma	mē	m, mi	mo			mwa		

---

[48] The syllabary for Tigrinya is adapted from *Tigrinya Grammar*, edited by John Mason, published by The Red Sea Press, Inc., Lawrenceville, New Jersey, 1996.

№												
5B	se	su	sī	sa	sē	s, si	so					
6	re	ru	rī	ra	rē	r, ri	ro			rwa		
7	se	su	sī	sa	sē	s, si	so			swa		
8	she	shu	shī	sha	shē	sh, shi	sho			shwa		
9	k'e	k'u	k'ī	k'a	k'ē	k', k'i	k'o	k'weᴱ	k'wī	k'wa	k'wē	k'wi
10B	k'e	k'u	k'ī	k'a	k'ē	k', k'i	k'o	k'weᴱ	k'wī	k'wa	k'wē	k'wī
11	be	bu	bī	ba	bē	b, bi	bo			bwa		
12	te	tu	tī	ta	tē	t, ti	to			twa		
13	che	chu	chī	cha	chē	ch, chi	cho			chwa		
14B	hā	hu	hī	ha	hē	h, hi	ho	hweᴱ	hwī	hwa	hwē	hwi
15	ne	nu	nī	na	nē	n, ni	no			nwa		
16	nye	nyu	nyī	nya	nyē	ny, nyi	nyo			nywa		
17C	ā, 'ā	u, 'u	ī, 'ī	a, 'a	ē, 'ē	i, 'i, (-)	o			'wa		
18	ke	ku	kī	ka	kē	k, ki	ko	kweᴱ	kwī	kwa	kwē	kwi
19B	he	hu	hī	ha	hē	h, hi	ho	hweᴱ	hwī	hwa	hwē	hwi
20	we	wu	wī	wa	wē	w, wi	wo					
21	'ā	'u	'ī	'a	'ē	', 'i	'o					
22	ze	zu	zī	za	zē	z, zi	zo			zwa		
23	zhe	zhu	zhī	zha	zhē	zh, zhi	zho					
24	ye	yu	yī	ya	yē	y, yi	yo					
25	de	du	dī	da	dē	d, di	do			dwa		
26	je	ju	jī	ja	jē	j, ji	jo					
27	ge	gu	gī	ga	gē	g, gi	go	gweᴱ	gwī	gwa	gwē	gwi
28	t'e	t'u	t'ī	t'a	t'ē	t', t'i	t'o			t'wa		

29	ch'e	ch'u	ch'ī	ch'a	ch'ē	ch',ch'i	ch'o	ch'wa
30	p'e	p'u	p'ī	p'a	p'ē	p', p'i	p'o	
31	ts'e	ts'u	ts'ī	ts'a	ts'ē	ts',ts'i	ts'o	ts'wa
32ᴮ	ts'e	ts'u	ts'ī	ts'a	ts'ē	ts',ts'i	ts'o	
33	fe	fu	fī	fa	fē	f, fi	fo	fwa
34	pe	pu	pī	pa	pē	p, pi	po	
35	ve	vu	vī	va	vē	v, vi	vo	

ᴬ The sixth-order characters should be romanized with or without the letter *i*, depending on pronunciation.

ᴮ For documentation purposes the following consonants may be romanized with diacritical marks: ḥ (row 3), š (row 5), k' (row 10, specific to Tigrinya), ḫ (row 14), ḥ (row 19), ṯs̱' (row 32).

ᶜ The vowel characters in row 17 should be romanized *ā, u, ī, a, ē, i,* and *o* initially and *'ā, 'u, 'ī, 'a, 'ē, 'i,* and *'o* in all other positions.

ᴰ According to *Tigrinya Grammar* (1996) the alphabet has recently been written in a more consistent way, so that e.g. in the first order the exceptional pronunciation of some vowels (*hā, 'ā,* etc.) has been abolished and the characters are given the same vowel *e* as for others. This means that the character ኣ is being replaced by the 4th-order syllable in many words, e.g. the old spelling ኣሎ *ālo*, the preferred new spelling ኣሎ *alo*.

ᴱ Romanization equivalents of combinations with W in column 1 are presented according to the pronunciation indicated in *Tigrinya Grammar* (1996). They are different from those of the BGN/PCGN Romanization System for Amharic.

Ukrainian

At sessions of the United Nations Group of Experts on Geographical Names there have been presented several documents on the romanization of Ukrainian geographical names. The first scheme for romanizing Ukrainian names was adopted in Ukraine in 1993[49].

The current national system of romanization was approved by the Minister of Justice (also Chairman of the Committee on Law Terminology) in 1996[50]. However, for final adoption it is envisaged to approve it as a national standard. The system which is based on the English orthography, is for the most part reversible, although certain character combinations may cause

---

[49] *Provisional Rules of Reproducing Letters of the Ukrainian Alphabet with Latin (English) Characters*. UNGEGN, 17th Session. New York, 13-24 June 1995. WP No. 73.

[50] *Romanization of Ukrainian Geographical Names*. UNGEGN, 18th Session. Geneva, 12-23 August 1996. WP No. 78.

ANNEX:
LANGUAGES THAT HAVE RECENTLY ADOPTED ROMAN ALPHABETS

Azerbaijani

Since 1991 a new Roman alphabet has been officially decreed for the Azerbaijani language that was formerly written in the Cyrillic alphabet.

The Cyrillic characters have the following equivalents in the new **official Roman alphabet**.

1	A a	a	10	З з	z	19	O o	o	28	X x	x
2	Б б	b	11	И и	i	20	Ѳ ѳ	ö	29	h h	h
3	B в	v	12	Ы ы	ı	21	П п	p	30	Ч ч	ç
4	Г г	q	13	J j	y	22	P p	r	31	Ч ч	c
5	F ғ	ğ	14	K к	k	23	C c	s	32	Ш ш	ş
6	Д д	d	15	K к	g	24	T т	t	33	'	'
7	E e	e	16	Л л	l	25	У у	u			
8	Ə ə	ə	17	M м	m	26	Y y	ü			
9	Ж ж	j	18	H н	n	27	Ф ф	f			

Turkmen

A new Roman alphabet was officially decreed in 1993 for the Turkmen language that was formerly written in the Cyrillic alphabet. Some of the Roman letters were later changed as confirmed by the Turkmen Standard TDS 565, 1999.

The Cyrillic characters have the following equivalents in the new **official Roman alphabet**.

1	A a	a	11	И и	i	21	P p	r	31	Щ щ	şç[D]
2	Б б	b	12	Й й	ý[B]	22	C c	s	32	Ъ ъ	(omitted)
3	B в	w	13	K к	k	23	T т	t	33	Ы ы	y
4	Г г	g	14	Л л	l	24	У у	u	34	Ь ь	(omitted)
5	Д д	d	15	M м	m	25	Y y	ü	35	Э э	e
6	E e	e, ýe[AB]	16	H н	n	26	Ф ф	f	36	Ə ə	ä
7	Ё ё	ýo[B]	17	Ң ң	ň[B]	27	X x	h	37	Ю ю	ýu[B]
8	Ж ж	ž[B]	18	O o	o	28	Ц ц	(t)s[C]	38	Я я	ýa[B]
9	Җ җ	j	19	Ѳ ѳ	ö	29	Ч ч	ç			
10	З з	z	20	П п	p	30	Ш ш	ş[B]			

[A] The letter sequence *ýe* is used initially, after vowel characters (a, e, ë, и, o, ѳ, у, ү, ы, э, ю, я) and й.

[B] Earlier variations of the letters that may still be encountered are as follows (1999 = 1993): Ňň = Ññ, Şş = $¢ Ýý = ¥ÿ, Žž = £ɾ.

[C] The Cyrillic ц was used in loanwords only, it has been replaced mainly by s, e.g. Франция *Fransiýa*.

[D] This letter sequence is rarely encountered.

Uzbek

In May 1995 a new Roman alphabet was officially decreed for the Uzbek language, amending the alphabet adopted earlier, in 1993. The Roman alphabet is now in parallel use with the Cyrillic alphabet. Full implementation of it has been postponed to the year 2005[51].

The Cyrillic characters have the following equivalents in the new **official Roman alphabet**.

1	A a	a	10	И и	i	19	C c	s	28	Ь ь	'
2	Б б	b	11	Й й	y	20	T т	t	29	Э э	e
3	B в	v	12	K к	k	21	У у	u	30	Ю ю	yu
4	Г г	g	13	Л л	l	22	Ф ф	f	31	Я я	ya
5	Д д	d	14	M м	m	23	X x	x	32	Ў ў	oʻ [B]
6	E e	e, ye[A]	15	H н	n	24	Ц ц	ts	33	Қ қ	q
7	Ё ё	yo	16	O o	o	25	Ч ч	ch	34	F ғ	gʻ [B]
8	Ж ж	j	17	П п	p	26	Ш ш	sh	35	Ҳ ҳ	h
9	З з	z	18	P p	r	27	Ъ ъ	'			

[A] The letter sequence *ye* is used initially, after the vowel characters (а, е, ё, и, о, у, э, ю, я, ў), й and ъ.

[B] These letters are decreed to have a different shape when handwritten:

Printed:	gʻ	oʻ
Handwritten:	ḡ (*or* ğ)	ō (*or* ŏ).

---

[51] *Report by the Convenor of the Working Group on Romanization Systems*. UNGEGN, 18th Session. Geneva, 12-23 August 1996. WP No. 21.