

**Twentieth Session  
New York, 17 – 28 January 2000**

---

Item 11  
of the Provisional Agenda

**MEETING OF THE WORKING GROUP ON ROMANIZATION SYSTEMS**

**United Nations Romanization Systems for Geographical Names  
Preliminary Report on their Current Status**

United Nations Group of  
Experts on Geographical Names

Twentieth Session  
New York, 17-28 January 2000

Item No. 11 of the Provisional Agenda: Meeting of the Working Group on Romanization Systems

**UNITED NATIONS ROMANIZATION SYSTEMS FOR GEOGRAPHICAL NAMES**  
**Preliminary Report on Their Current Status**

Submitted by the UNGEGN Working Group on Romanization Systems

# UNITED NATIONS ROMANIZATION SYSTEMS FOR GEOGRAPHICAL NAMES

## Preliminary Report on Their Current Status

Compiled by the UNGEGN Working Group on Romanization Systems  
(Version 1.2, November 1999)

### Content

Introduction	1
List of languages, countries and romanization systems	3
Romanization systems	9

### *Languages covered by the systems recommended by the United Nations:*

Amharic	9	Kannada	25	Russian	38
Arabic	11	Khmer	27	Serbian	39
Assamese	14	Macedonian Cyrillic	29	Tamil	39
Bengali	15	Malayalam	30	Telugu	40
Bulgarian	17	Marathi	32	Thai	41
Chinese	18	Mongolian (in China)	32	Tibetan	43
Greek	19	Nepali	33	Uighur	45
Gujarati	20	Oriya	34	Urdu	46
Hebrew	22	Persian	35		
Hindi	23	Punjabi	37		

### *Other languages:*

Armenian	48	Kirghiz	56	Tajik	63
Belarusian	48	Korean	56	Tigrinya	64
Burmese	49	Laotian	58	Turkmen	64
Dzongkha	51	Maldivian	60	Ukrainian	65
Georgian	53	Mongolian (Cyrillic)	61	Uzbek	65
Japanese	53	Pashto	61		
Kazakh	55	Sinhalese	62		

## INTRODUCTION

The report will aim at clarifying 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: 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 that one of the general aims of the 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"*<sup>1</sup>. The Fourth United Nations Conference on the Standardization of Geographical Names (1982) also adopted the resolution 15 which called for guidelines governing the consideration of romanization systems. The resolution also stated that

- a) the new romanization systems for international use should be considered only on condition that the sponsoring nations implement such systems on their cartographical products;
- b) 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 put forward 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 on 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 to use as possible. E.g. if a plain 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 one should consult dictionaries and other sources to get 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 e.g. store electronically. That would mean the minimization of the use of diacritical marks, avoiding difficult and unusual character sequences, etc. A systematic representation of phonetic 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, incl. in 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

---

<sup>1</sup> Resolution 9 of the first UN Conference on the Standardization of Geographical Names (1967)

Organization for Standardization (ISO). For one such language there may exist two totally differing systems, as e.g. for Thai (UN system 1967, and ISO 11940:1998), or these two can be combined into one integral system, as e.g. for Greek (UN 1987, ELOT743) where there are two versions of romanization: one for general purposes, and the other for documenting, i.e. a reversible one. Reversibility is achieved by adding diacritical signs 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.

## 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 after their English name. The names of languages conform to those mentioned in the appropriate United Nations resolutions.

In column 4 the year of the adoption of the system by the United Nations is given. In brackets the first part refers to the United Nations Conference on the Standardization of Geographical Names, and the part after the slash to the number of the resolution.

In column 5 various other systems are mentioned that could have some international usage. An equation mark (=) will indicate that the systems in question more or less coincide. By 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 the 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 of 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 (1/17)	
				BGN/PCGN 1967

Language	Writing System	Country (Countries)	Romanization Systems	
			United Nations	Other
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
		Iraq	(see General)	
		Israel	(see General)	
		Jordan	(see General)	
				National: Royal Jordanian Geographic Centre System
		Kuwait	(see General)	
		Lebanon	(see General)	
		Libyan Arab Jamahiriya (the)	(see General)	
		Mauritania	(see General)	
		Morocco	(see (General)	
		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)	

Language	Writing System	Country (Countries)	Romanization Systems	
			United Nations	Other
		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
Belarusian (Byelorussian)	Cyrillic	Belarus	—	National 1998
				BGN/PCGN 1979
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
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) <sup>2</sup>	
Dzongkha	Tibetan	Bhutan	—	National 1994
Georgian	Georgian	Georgia	—	BGN/PCGN 1981
				ISO 9984:1996
Greek	Greek	Cyprus	1987 (V/19)	= ELOT 743, = ISO 843:1997
		Greece	1987 (V/19)	

<sup>2</sup> - the recommendation applies to Persian of which Dari is a variation

Language	Writing System	Country (Countries)	Romanization Systems	
			United Nations	Other
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
Japanese	Sino-Japanese	Japan	—	National: <i>Kunrei-siki</i> (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
Kirghiz	Cyrillic	Kyrgyzstan	—	BGN/PCGN 1979
Korean	Korean	(General)	—	McCune Reischauer 1939, = BGN/PCGN
				ISO/TR 11941:1996
		Democratic People's Republic of Korea (the)	—	National 1992
		Republic of Korea (the)	—	National: Ministry of Education 1984
Laotian	Laotian	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)	



Language	Writing System	Country (Countries)	Romanization Systems	
			United Nations	Other
				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)	
		Mongolia	--	
Nepalese	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)
		Tajikistan	1987 (V/18)	
		Uzbekistan	1987 (V/18)	

<i>Language</i>	<i>Writing System</i>	<i>Country (Countries)</i>	<i>Romanization Systems</i>	
			<i>United Nations</i>	<i>Other</i>
Serbian <sup>3</sup>	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)	
Tigrinya	Ethiopic	Eritrea	—	BGN/PCGN 1994
Turkmen	Cyrillic	Turkmenistan	—	BGN/PCGN 1979
Uighur	Perso-Arabic	China	1977 (III/8)	
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
Uzbek	Cyrillic	Uzbekistan	—	BGN/PCGN 1979

<sup>3</sup> - referred to as Serbo-Croatian Cyrillic in Resolution 11 of the Third UN Conference on the Standardization of Geographical Names, 1977

## ROMANIZATION SYSTEMS

Structure of the description:

1. Origin of the system.
2. Implementation of the system (both nationally and internationally). Perspectives.
3. Brief characterization of the system: ease of applicability, reversibility, possible ambiguities and other problems.
4. The romanization system 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 to make concise their presentation. Where possible, all non-Roman characters are given 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 occur in supplementary notes.
5. Other systems of romanization, comparison with the UN recommended system.

### A. LANGUAGES COVERED BY THE SYSTEMS RECOMMENDED BY THE UNITED NATIONS

#### Amharic

1. 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<sup>4</sup>.

2. 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.

3. Amharic uses an Ethiopic syllabic script whereby each syllable denotes a combination of a consonant and a vowel. It has 34 basic symbols, each symbol having several different forms, usually called orders, according to the vowel with which the basic symbol 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.

4.

		Combinations with W and Y:											
		1	2	3	4	5	6 <sup>1</sup>	7 (1)	(3) <sup>3</sup>	(4) <sup>3</sup>	(5) <sup>3</sup>	(6) <sup>3</sup>	(4) <sup>3</sup>
1	ሀ	ሁ	ሂ	ሃ	ሄ	ሀ	ሀ	ሀ	ሀ	ሀ	ሀ	ሀ	ሀ
	ha <sup>2</sup>	hu	hi	ha	he	hi	ho			hwa			
2	ለ	ሉ	ሊ	ላ	ሌ	ለ	ለ	ለ	ለ	ለ	ለ	ለ	ለ
	le	lu	li	la	le	li	lo			lwa			

<sup>4</sup> *Second United Nations Conference on the Standardization of Geographical Names*. London, 10-31 May 1972. Vol. II. *Technical papers*, pp. 165-168 (Table 2)

3	ሐ	ሐ	ሐ	ሐ	ሐ	ሐ	ሐ					
4	ha <sup>2</sup>	hu	hi	ha	he	hi	ho					
5	መ	ሙ	ሚ	ማ	ሚ	ም	ሞ			ግ (ግግ)	ግ	
6	me	mu	mi	ma	me	mi	mo			mwa	mya	
7	ሠ	ሠ	ሢ	ሣ	ሢ	ሥ	ሦ					
8	se	su	si	sa	se	si	so					
9	ረ	ሩ	ሪ	ራ	ሪ	ሮ	ሾ			ሯ	ሯ	
10	re	ru	ri	ra	re	ri	ro			rwa	rya	
11	ሰ	ሱ	ሲ	ሳ	ሴ	ሰ	ሶ			ሷ		
12	se	su	si	sa	se	si	so			swa		
13	ሸ	ሹ	ሺ	ሻ	ሼ	ሽ	ሾ			shwa		
14	she	shu	shi	sha	she	shi	sho			shwa		
15	ቀ	ቁ	ቂ	ቃ	ቄ	ቅ	ቆ	ቇ	ቈ	቉	ቊ	
16	k'e	k'u	k'i	k'a	k'e	k'i	k'o	k'we	k'wi	k'wa	k'we	k'wi
17	በ	ቡ	ቢ	ባ	ቤ	ብ	ቦ			ጠ		
18	be	bu	bi	ba	be	bi	bo			bwa		
19	ተ	ቲ	ቲ	ታ	ቲ	ቲ	ቲ			ጡ		
20	te	tu	ti	ta	te	ti	to			twa		
21	ቸ	ቹ	ቺ	ቻ	ቼ	ች	ቾ			ጢ		
22	che	chu	chi	cha	che	chi	cho			chwa		
23	ኀ	ኁ	ኂ	ኃ	ኄ	ኅ	ኆ	ኇ	ኈ	኉	ኊ	
24	ha <sup>2</sup>	hu	hi	ha	he	hi	ho	hwe	hwi	hwa	hwe	hwi
25	ኀ	ኁ	ኂ	ኃ	ኄ	ኅ	ኆ			ኇ(ኀ)		
26	ne	nu	ni	na	ne	ni	no			nwa		
27	ኘ	ኙ	ኚ	ኛ	ኜ	ኝ	ኞ			ኟ		
28	nye	nyu	nyi	nya	nye	nyi	nyo			nywa		
29	አ	ሁ	አ	አ	ኤ	አ	አ			ኦ		
30	a <sup>2</sup>	u	i	a	e	i	o					
31	ከ	ከ	ከ	ካ	ኬ	ከ	ኮ	ኰ	኱	ኲ	ኳ	
32	ke	ku	ki	ka	ke	ki	ko	kwe	kwi	kwa	kwe	kwi
33	ኸ	ኸ	ኸ	ኹ	ኺ	ኸ	ኹ			ኻ		
34	he	hu	hi	ha	he	hi	ho					
35	ወ	ወ	ወ	ወ	ወ	ወ	ወ			ወ		
36	we	wu	wi	wa	we	wi	wo					
37	ዐ	ዐ	ዐ	ዐ	ዐ	ዐ	ዐ			ዐ		
38	a <sup>2</sup>	u	i	a	e	i	o					
39	ዘ	ዘ	ዘ	ዘ	ዘ	ዘ	ዘ			ዘ		
40	ze	zu	zi	za	ze	zi	zo			zwa		
41	ኸ	ኸ	ኸ	ኹ	ኺ	ኸ	ኹ			ኻ		
42	zhe	zhu	zhi	zha	zhe	zhi	zho			zhwa		
43	የ	የ	የ	የ	የ	የ	የ			የ		
44	ye	yu	yi	ya	ye	yi	yo			ywa		
45	ደ	ደ	ደ	ደ	ደ	ደ	ደ			ደ		
46	de	du	di	da	de	di	do			dwa		
47	ጆ	ጆ	ጆ	ጆ	ጆ	ጆ	ጆ			ጆ		
48	je	ju	ji	ja	je	ji	jo			jwa		
49	ገ	ገ	ገ	ገ	ገ	ገ	ገ	ገ	ገ	ገ	ገ	
50	ge	gu	gi	ga	ge	gi	go	gwe	gwi	gwa	gwe	gwi

27	ጠ	ጡ	ጢ	ጣ	ጤ	ጥ	ጦ	ጧ	
	t'e	t'u	t'i	t'a	t'e	t'i	t'o	t'wa	
28	ጪ	ጫ	ጬ	ጭ	ጮ	ጯ	ጰ	ጱ	
	ch'e	ch'u	ch'i	ch'a	ch'e	ch'i	ch'o	ch'wa	
29	ጸ	ጹ	ጺ	ጻ	ጼ	ጽ	ጾ		
	p'e	p'u	p'i	p'a	p'e	p'i	p'o		
30	ጺ	ጻ	ጼ	ጽ	ጾ	ጿ	ጺ	ጻ	
	ts'e	ts'u	ts'i	ts'a	ts'e	ts'i	ts'o	ts'wa	
31	ፀ	ፁ	ፂ	ፃ	ፄ	ፅ	ፆ		
	ts'e	ts'u	ts'i	ts'a	ts'e	ts'i	ts'o		
32	ፈ	ፉ	ፊ	ፋ	ፅ	ፇ	ፈ	ፉ (ፊ ፋ)	ፊ
	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 <sup>3</sup>	ፕ	ፖ	ፗ	ፘ	ፙ	ፚ	፛		
	ve	vu	vi	va	ve	vi	vo		

- <sup>1</sup> the vowel of the sixth order (i) is eliminated in spelling except when the actual pronunciation requires it (e.g. not *Menigisiti* but *Mengist*)
- <sup>2</sup> sounds identical to fourth-order form
- <sup>3</sup> 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 elsewhere in the table

5.

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

UN	BGN/PCGN
e	e
i	i
e	e
i	i

The syllables *gwe*, *hwe*, *kwē*, *k'wē* of the UN system are treated as variations of the syllables *go*, *ho*, *ko*, *k'o* in the BGN/PCGN system.

For documentation purposes the following consonants may be romanized with diacritical marks in the BGN/PCGN system: ḥ (row 3), ṣ (row 5), ḫ (row 13), ḥ (row 18), ts' (row 31).

Arabic

1.

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<sup>5</sup>.

2.

Already in the UN resolution it was exceptionally pointed out that the system was recommended "for the romanization of the geographical names within those Arabic-speaking countries where this system is officially

<sup>5</sup> *Second United Nations Conference on the Standardization of Geographical Names*. London, 10-31 May 1972, Vol. II: *Technical papers*, p. 170.

acknowledged". It cannot be definitely stated which of the Arabic-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 Bahrain, Iraq, Jordan (see also p. 5), Kuwait, the Libyan Arab Jamahiriya, Oman, Qatar, Saudi Arabia, United Arab Emirates and Yemen. For the geographical names of Lebanon and 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. Finally, there is no sufficient evidence to conclude which systems are used for Arabic names in Chad, the Comoros and Somalia.

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".

3.

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, also one must take into account dialectal and idiosyncratic deviations. The romanization is generally reversible though there are some ambiguous character sequences (*dh, kh, sh, th*) which may also point to combinations of Arabic characters in addition to the respective single characters.

4.

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
ء				' <sup>(1)</sup>	ر			ر	r	ف	ف	ف	ف	f
ا			ا	-	ز			ز	z	ق	ق	ق	ق	q
ب	ب	ب	ب	b	س	س	س	س	s	ك	ك	ك	ك	k
ت	ت	ت	ت	t	ش	ش	ش	ش	sh	ل	ل	ل	ل	l
ث	ث	ث	ث	th	ص	ص	ص	ص	ṣ	م	م	م	م	m
ج	ج	ج	ج	j	ض	ض	ض	ض	ḍ	ن	ن	ن	ن	n
ح	ح	ح	ح	ḥ	ط	ط	ط	ط	ṭ	ه	ه	ه	ه	h
خ	خ	خ	خ	kh	ظ	ظ	ظ	ظ	ẓ	و			و	w
د			د	d	ع	ع	ع	ع	'	ي	ي	ي	ي	y
ذ			ذ	dh	غ	غ	غ	غ	gh					

<sup>1</sup> omitted word-initially

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

•	a	•	( <sup>1</sup> )	•	á
••	aw	•	ā	•	a <sup>n</sup>
•••	ay	•	ā	•	i <sup>n</sup>
•	i	•	ī	•	u <sup>n</sup>
•	u	•	ū	•	( <sup>2</sup> )

<sup>1</sup> marks absence of the vowel

<sup>2</sup> 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, ḡ, ḥ, ḷ, l, n*) the *t* of the definite article is assimilated with the first consonant of the word: *الشارقة* ash-Shāriqah.

## 5.

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. *البيصره* 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<sup>6</sup>:

UN	I.G.N.				
a	= a, e, é, e <sup>1</sup>	j	= dj, j <sup>1</sup>	u	= ou, o <sup>1</sup>
ā	= ā, e <sup>1</sup>	n	= n, ne <sup>2</sup>	ū	= oū, ô <sup>1</sup>
á	= a	q	= q, g <sup>1</sup> , gu <sup>3</sup>	w	= ou
ḡ	= ḡ	s	= s, se <sup>4</sup>	y	= i, i <sup>5</sup> , y <sup>7</sup>
ḥ	= ḥ	ṣ	= ṣ, c <sup>5</sup>	z	= z
i	= i, e <sup>1</sup>	sh	= ch	'	(omitted)
ī	= i, e <sup>1</sup>	ṭ	= t	‘	= ‘, aa <sup>8</sup>

<sup>1</sup> used according to the local pronunciation

<sup>2</sup> *ne* is used word-finally after *a, e, i, o*

<sup>3</sup> *gu* is used before *e* and *i*

<sup>4</sup> *se* is used between vowels

<sup>5</sup> *ç* is used in some conventional names

<sup>6</sup> *t* is used as second element of a diphthong

<sup>7</sup> *y* is used word-initially or between vowels

<sup>8</sup> *aa* is specific to Lebanon

The transliteration ISO 233:1984 gives every character and diacritic sign 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 (ā)	= a	ḥ	= ḥ	th	= t
á	= ay	i <sup>a</sup>	= í	u <sup>a</sup>	= ú
a <sup>a</sup>	= á	j	= ğ	y (ي)	= ý
ḡ	= ḡ	kh	= ḫ	y (ي)	= y
dh	= ḡ	ṣ	= ṣ	z	= z
gh	= ğ	sh	= š		
h (h)	= ħ	ṭ	= t		

<sup>6</sup> *Présentation de la variante B du système de translittération de l'arabe "Beirut amendé"*. UNGEGN, 17th Session, New York, 13-24 June 1994. WP No. 61.

The Royal Jordanian Geographic Centre (RJGC) System<sup>7</sup> is essentially the same as the amended Beirut system. The underscore is used instead of cedilla.

UN	RJGC				
-ah (ā)	= -a	ḥ	= h	ṭ	= t
ḍ	= ḍ	š	= š	ẓ	= ḡh

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

UN	SES				
á	= a	ḥ	= h	th	= th (t)
ā	= ā (a)	ī	= i	ū	= ū
-ah (ā)	= -a	j	= g (j)	ẓ	= ẓ (d)
aw	= ō (au)	q	= q (k)	'	= '
ay	= ei (ai)	s	= s (c)		
ḍ	= ḍ	š	= š		
dh	= dh (z)	ṭ	= t		

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

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 forms in the official maps edited since 1969 have been rendered in accordance with a simplified version of the I.G.N. system<sup>8</sup>. 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

### 1.

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<sup>9</sup>.

### 2.

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

### 3.

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

<sup>7</sup> *Activities in Jordan on the Standardization of Geographical Names*. UNGEGN, 18th Session. Geneva, 12-23 August 1996. WP. No. 86.

<sup>8</sup> Report of the Working Group on a Single Romanization System for Each Non-Roman Writing System: Activities from 4 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.

<sup>9</sup> *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.



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).

## 4.

## Independent vowel characters

অ a	ই i	উ u	ঋ r	এ e	ও o
আ ā	ঈ ī	ঊ ū		ঐ ai	ঔ au

## Abbreviated vowel characters (• stands for any consonant character)

• a	ি i	ু u <sup>1</sup>	্ r <sup>3</sup>	ে e	ো o
া ā	ী ī	ূ ū <sup>2</sup>		ৈ ai	ৌ au

<sup>1</sup> exceptions: ঔ gu; ঋ ru; ঞ shu; হ hu; ঞ ntu; ঞ stu

<sup>2</sup> exception: ঋ rū

<sup>3</sup> exception: ঞ hr

## Other symbols

• m	• m̄	• h	• 1
-----	------	-----	-----

<sup>1</sup> pronunciation without a vowel; special form: ৎ t

## Consonant characters

ক ka	চ cha	ট ta	ত ta	প pa	য ja <sup>1</sup>	শ sha
খ kha	ছ chha	ঠ tha	থ tha	ফ pha	ব ra	ষ sha
গ ga	জ ja	ড da <sup>1</sup>	দ da	ভ ba	ল la	স sa
ঘ gha	ঝ jha	ঢ dha <sup>1</sup>	ধ dha	ভ bha	ৱ va	হ ha
ঙ ṅa	ঞ ña	ণ ṅa	ন na	ম ma		

<sup>1</sup> dotted variants of the characters: ৎ ta; ৎ tha; ৎ ya

## Ligatures

See the section on the romanization of Bengali.

## 5.

For differences between the UN system and the ISO transliteration draft ISO CD15919 see the section on the romanization of Hindi.

## Bengali

## 1.

The United Nations recommended system was approved in 1972 (III/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<sup>10</sup>.

<sup>10</sup> *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.

## 2.

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.

## 3.

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).

## 4.

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

## Independent vowel characters

অ a	ই i	উ u	ঋ r	এ e	ও o
আ ā	ঈ ī	ঊ ū		ঐ ai	ঔ au

## Abbreviated vowel characters (• stands for any consonant character)

• a	ি i	ু u <sup>1</sup>	ৃ r <sup>3</sup>	ে e	ো o
া ā	ী ī	ূ ū <sup>2</sup>		ৈ ai	ৌ au

<sup>1</sup> exceptions: ও gu; রু ru; শু shu; হু hu; ণ্ট ntu; ণ্ট stu

<sup>2</sup> exception: রূ rū

<sup>3</sup> exception: হ্র hr

## Other symbols

• r̄ m	• m̄	• : h	• 1
--------	------	-------	-----

<sup>1</sup> pronunciation without a vowel; special form: ৎ t

## Consonant characters

ক ka	চ cha	ট ta	ত ta	প pa	য ja <sup>1</sup>	শ sha
খ kha	ছ chha	ঠ tha	থ tha	ফ pha	র ra	ষ sha
গ ga	জ ja	ড da <sup>1</sup>	দ da	ব ba	ল la	স sa
ঘ gha	ঝ jha	ঢ dha <sup>1</sup>	ধ dha	ভ bha		হ ha
ণ na	ঞ ña	ণ na	ন na	ম ma		

<sup>1</sup> dotted variants of the characters: ড̣ ta; ঢ̣ tha; ণ̣ ya

## Ligatures

## Adscript forms of some consonants

Example

• r-: র̣ rka

• r-: র̣ gra

• J -y: জ̣ tyā

Other ligatures (the list is not complete)

ক	ক্	ক্	ক্	ক্	ক্	ক্	ক্	ক্	ক্	ক্	ক্	ক্
kka	kṛa	kṛa	kna	kma	kra	kla	kva	kṣha	kṣhna	kṣhma	kṣhva	
ক্স	গ	গ	গ্	গ্	গ্	গ্	গ্	ঘ	ক	ক	ক	ক
ksa	gga	gda	gdha	gna	gma	gra	gla	ghra	ṅka	ṅga	chcha	
চ্ছ	চ্ছ	চ্ছ	চ্ছ	চ্ছ	চ্ছ	চ্ছ	চ্ছ	চ্ছ	চ্ছ	চ্ছ	চ্ছ	চ্ছ
chchha	chchva	chña	jjā	jjva	jjha	jjña	jva	fiḥha	fiḥha	fija	fijha	
ট	ড	ট	ঠ	ড	ড	ড	থ	ড	অ	অ	ড	ড
ṭṭa	ḍḍa	ṇṭa	ṇṭha	ṇḍa	ṭṭa	ṭṭva	ṭṭha	ṭṭa	ṭṭa	tra	ṭṭa	
ড	দ	দ	দ	দ	দ	দ	ড	দ	দ	দ	দ	দ
tva	dda	ddva	ddha	ddhva	dna	dva	dbha	dma	dra	dla	dhra	
ঠ	ড	ক	ক	ক	ক	ক	ক	ক	ক	ক	ক	ক
nṭha	ṇḍa	nka	nta	ntra	ntha	nda	ndra	ndha	nna	nma	nva	
প	ঙ	প	প্র	প	ফ	জ	দ	ক	ক	ক	ক	ক
pna	pta	ppa	pra	pla	phra	bja	bda	bdha	bba	bra	bhra	
ম্প	ব	ভ	ম	ম	ম	ম	ক	ক	ক	ক	ক	ক
mpa	mba	mbha	mbhra	mma	mra	mra	ika	lṭa	lḍa	lma	lla	
শ্চ	শ্চ	শ্চ	শ্চ	শ্চ	শ্চ	শ্চ	শ্চ	শ্চ	শ্চ	শ্চ	শ্চ	শ্চ
shcha	shchha	shta	shna	shma	shra	shla	shva	ṣhka	ṣhta	ṣhtra	ṣhṭha	
ক্ষ	ক্ষ	ক্ষ	ক্ষ	ক্ষ	ক্ষ	ক্ষ	ক্ষ	ক্ষ	ক্ষ	ক্ষ	ক্ষ	ক্ষ
ṣhña	ṣhpa	ṣhpha	ska	skra	skha	sta	sna	sma	sra	sva	hna	
হ্ম	হ্ম	হ্ম										
hma	hra	hla										

## 5.

For differences between the UN system and the ISO transliteration draft ISO CD15919 see the section on the romanization of Hindi.

## Bulgarian

## 1.

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 annex to the resolution<sup>11</sup>. The 1977 resolution adopted one of the two systems approved by the UN conference in 1972 (II/5).

## 2.

The system is used in Bulgaria and increasingly in the international cartographic products. It has been mentioned in the information concerning the 15th session of the UNGEGN Division of East Central and South-East Europe (Ljubljana 1999) that the Bulgarian Council of Orthography and Transcription had adopted on April 8, 1999 a modified system of romanization which however needed confirmation and publication. The essence of the changes is not yet known to the Working Group.

## 3.

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. E.g. the romanized *j* will correspond to *ѣ* if used after a consonant, in other cases it represents the Cyrillic *ѣ*. As exception, when *j* is followed by *a* or *u*, the combinations *ja* and *ju* should be reverted to the Cyrillic as *ѣ* and *ѣ*

<sup>11</sup> Third United Nations Conference on the Standardization of Geographical Names. Athens, 17 August - 7 September 1977. Vol. I, Report of the Conference, p. 28

respectively. These romanizations could be ambiguous, if there should exist character sequences *ѣа* and *ѣу* which seems highly unlikely. The romanization of *и* by *st* could be ambiguous if the Cyrillic sequences *иит* may occur as well.

4.

А а	а	И и	и	Р р	р	Ш ш	ш
Б б	б	Й й	й	С с	с	Щ щ	шт
В в	в	К к	к	Т т	т	Ъ ъ	а
Г г	г	Л л	л	У у	у	Ь ь	ј
Д д	д	М м	м	Ф ф	ф	Ю ю	ју
Е е	е	Н н	н	Х х	х	Я я	ја
Ж ж	џ	О о	о	Ц ц	ц		
З з	з	П п	п	Ч ч	ч		

## 5.

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

ж (џ)	zh	ц (ц)	ts	ш (шт)	sht	ю (ју)	yu
ђ (ј)	y	ч (ч)	ch	ъ (џ)	ǔ	я (ја)	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

## 1.

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

## 2.

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

## 3.

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 reverted to its written form.

## 4.

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

## 5.

The Modified Wade-Giles transcription (1912) was previously one of the most well-known systems of rendering Chinese syllables, especially in the English-language texts. The syllables of the Wade-Giles (WG) system may be transformed into Pinyin syllables as given in the table below<sup>12</sup> 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:	ch'	= ch, q'	j	= r	k'	= k	
ch	= zh, j'	hs	= x	k	= g	p	= b

<sup>12</sup> Adapted from Klaus Kaden, *Die wichtigsten Transkriptionssysteme für die chinesische Sprache*. VEB Verlag Enzyklopädie, Leipzig 1975, S. 145.

p'	= p	tz'	= c	i	= i, yi <sup>2</sup>	ü	= u, ü <sup>6</sup>
ss	= s	Syllable endings:		ieh	= ie	üan	= uan
sz	= s	ê	= e	ien	= ian	üeh	= ue, üe <sup>6</sup>
t	= d	eh	= e	ih	= i	uei	= ui
t'	= t	en	= an	iung	= iong	ün	= un
ts	= z	ên	= en	o	= o <sup>3</sup> , uo <sup>4</sup> , e <sup>5</sup>	ung	= ong
ts'	= c	êng	= eng	u	= u, ou <sup>6</sup>		
tz	= z	êrh	= er	ü	= i		

<sup>1</sup> before *i* (except the ending *-ih*) and *ü*

<sup>2</sup> word-initially

<sup>3</sup> after *p, p', m, f, w, y*

<sup>4</sup> after *t, t', n, l, ch, ch', j, ts, ts', s*

<sup>5</sup> after *k, k', h* and syllable-initially

<sup>6</sup> after *n, l*

A modification of the Wade-Giles system which excludes all diacritical marks, including the apostrophe, is being used in Taiwan, province of China (officially approved in 1972).

The Pinyin alphabet has also been adapted to the romanization of names from three minority languages in China: Mongolian (page 32), Tibetan (page 43) and Uighur (page 45).

## Greek

### 1.

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 annex to the resolution<sup>13</sup>.

### 2.

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".

### 3.

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 an underscore to certain ambiguous romanization equivalents.

### 4.

The table below contains the transcription version of the romanization system. The transliteration equivalent, if different from the transcription, is added in curly brackets.

A α	a	γξ	nx {nx}	Z ζ	z	K κ	k
αυ	av <sup>1</sup> , af <sup>2</sup>	γχ	nch {nch}	H η	i {i}	Λ λ	l
	{ay <sup>1</sup> , af <sup>2</sup> }	Δ δ	d	ηυ	iv <sup>1</sup> , if <sup>2</sup>	Μ μ	m
B β	v	Ε ε	e		{iv <sup>1</sup> , if <sup>2</sup> }	μπ	b <sup>3</sup> , mp <sup>4</sup>
Γ γ	g	ευ	ev <sup>1</sup> , ef <sup>2</sup>	Θ θ	th	N ν	n
Υ υ	ng {ng}		{ev <sup>1</sup> , ef <sup>2</sup> }	Ι ι	i	Ξ ξ	x

<sup>13</sup> Fifth United Nations Conference on the Standardization of Geographical Names, Montreal, 18-31 August 1987, Vol. I, Report of the Conference, pp. 42-43.

Ο ο	ο	Ρ ρ	ρ	Υ υ	υ	Ψ ψ	ps
ου	ou	Σ σ ς	s	Φ φ	f	Ω ω	o {Ω}
Π π	p	Τ τ	t	Χ χ	ch		

<sup>1</sup> before the consonants β, γ, δ, ζ, λ, μ, ν, ρ and all vowels

<sup>2</sup> before the consonants θ, κ, ξ, π, σ, τ, φ, χ, ψ and at the end of the word

<sup>3</sup> at the beginning of the word and at the end of the word

<sup>4</sup> in the middle of the word

#### Notes

- Some examples of the ELOT 743 table contain also combinations in parentheses αι, άι, αῑ, γκ, ει, έι, εῑ, υτ, οι, όι, οῑ, υι which are converted according to the conversion rules of each character they consist of, resp. ai, ái, aī, gk, ei, éi, eī, nt, oi, ói, oī, yi.
- 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.
- When the character combinations αυ, ευ, ηυ are stressed in Greek and the accent is transcribed, then in the Roman script the accent is set on the vowel (áv, áf, év, éf, ív, íf).

#### 5.

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

αυ (av/af) = av	γκ (gk) = g <sup>2</sup> , ng <sup>3</sup>	ευ (ev/ef) = ev	ου (ou) = u
γ (g) = g, y <sup>1</sup>	δ (d) = dh, d <sup>4</sup>	ηυ (iv/if) = iv	υ (y) = i
γχ (nch) = nkch	ει (ei) = i	υτ (nt) = d <sup>2</sup> , nd <sup>3</sup>	χ (ch) = kh

<sup>1</sup> before αι, ε, ει, η, ι, οι, υ and υι

<sup>2</sup> initially

<sup>3</sup> medially

<sup>4</sup> between ν and ρ

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 system:

Η η	Ι ι or Ῑ ῑ
Ω ω	Ό ό or Ο̄ ο̄

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 ΑΥ, αυ, ΕΥ, ευ, ΟΥ, ου which are mapped into Latin 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

### 1.

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<sup>14</sup>.

<sup>14</sup> *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.

## 2.

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

## 3.

Gujarati (Gujarāṭī) 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 characters (e/ē, o/ō) no precision has been made on which conditions either of the two given alternatives are to be used. 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).

## 4.

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

**Independent vowel characters**

અ a	ઈ i	ઉ u	૨ r	એ e, ē	ઓ o, ō
અ઼ ā	ઈ઼ ī	ઉ઼ ū		એ ai	ઓ au

**Abbreviated vowel characters (\* stands for any consonant character)**

• a	• i	• u	• r	• e, ē	• o, ō
• ā	• ī	• ū <sup>2</sup>		• ai	• au

<sup>1</sup> exception: ષ jī

<sup>2</sup> exception: ઝ rū

**Other symbols**

• m	• m̄	• h	• ʾ
-----	------	-----	-----

<sup>1</sup> pronunciation without a vowel

Note. These characters were missing in the original table.

**Consonant characters**

ક ka	ચ cha	ટ ta	ત ta	પ pa	ય ya	શ sha
ખ kha	છ chha	ડ tha	ધ tha	ફ fa	ર ra	ષ sha
ગ ga	જ ja	ડ ડા	ઢ ઢા	બ ba	લ la	સ sa
ઘ gha	ઝ jha	ઢ ઢા	ઘ ઘા	ભ bha	વ va	હ ha
ઙ na	ઞ ña	ણ ña	ન na	મ ma		ળ la

**Consonant ligatures****Adscript forms of some consonants****Examples**

•	r-	ક rka, ઘ rma
•	-r	ક kra, ઘ bra, ઙ dra

**Other ligatures (the list is not complete)**

ક	ક	ક	ખ	ખ	ગ	ગ	ઙ	ઙ	ચ	ચ	છ	છ
kka	kra	ksha	khta	khra	gra	ghra	ṅka	ṅkha	ṅga	ṅgha	ṅma	ṅma

च	ज	झ	ञ	ट	ठ	ड	ढ	ण	त	थ	द
chra	jja	jha	jra	ṭṭa	ṭṭha	ṭṭṭha	ḍḍa	ḍḍha	ḍḍḍha	ṭṭa	ṭṭa
न	त्र	थ्र	दग	दद	ददह	दभह	दम	दय	दर	दव	दध्र
na	tra	thra	dga	dda	ddha	dbha	dma	dya	dra	dva	dhra
म	न्र	प्र	प	फ	भ	भ्र	म्र	य्र	व्र	श्र	श्र
mna	nra	pta	pra	phra	bra	bhra	mra	yra	vra	shna	shra
श	श्र	श्र	श्र	स	ह	ह	ह	ह	ह	ह	ह
shva	shvra	shṭa	shṭha	sra	hra	hna	hma	hya	hla	hva	hva

## 5.

For differences between the UN system and the ISO transliteration draft ISO CD15919 see the section on the romanization of Hindi.

## Hebrew

## 1.

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 annex to the resolution<sup>15</sup>. The 1977 resolution was the elaboration of an earlier resolution (II/9) adopted in 1972.

## 2.

This 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 is yet known to have been taken<sup>16</sup>.

## 3.

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 (*ktiv male*) is gaining ground. 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.

## 4.

In the romanization table below five consonants have a special final form which is shown in addition to the main character after a comma.

א	ʾ, - <sup>2</sup>	ז	z	מ, ם	m	ק	q
ב	b	ח	h	נ, ן	n	ר	r
ב	v	ט	t	ס	s	ש	sh
ג (ג) <sup>4</sup>	g	י	y	ע	ʿ	ש	s
ד (ד) <sup>4</sup>	d	כ, ך	k	פ	p	ת (ת) <sup>4</sup>	t
ה	h <sup>7</sup>	כּ, ךּ	kh	פ, ף	f		
ו	w	ל	l	צ, ץ	z		

<sup>15</sup> 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.

<sup>16</sup> 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)

• a	• e	• i	• o
• a	• e, é <sup>5</sup>	• i	• u
• a, o	• e	• o	• u
• e	• e <sup>6</sup>	• o	

#### Notes

1. The addition of a dot (*dagesh hazaq*) 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 between two consecutive vowels or between 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 these letters carry a dot (*dagesh qal*).
5. In some maps the *gere* (•), if accentuated, is transliterated é.
6. The *sheva* (•) is of two kinds: *sheva nah*, which is omitted in transliteration, and *sheva na'*, which occurs at the beginning of a word or syllable.
7. Final *h* (ח), unless vocalized, is omitted in transliteration.

## Hindi

### 1.

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<sup>17</sup>.

### 2.

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<sup>18</sup>.

### 3.

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

### 4.

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

#### Independent vowel characters

अ a	इ i	उ u	ऋ r	ऌ l	ए e	ओ o
आ ā	ई ī	ऊ ū	ॠ ṛ		ऐ ai	औ au

<sup>17</sup> *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.

<sup>18</sup> *Fifth United Nations Conference on the Standardization of Geographical Names*. Montreal, 18-31 August 1987. Vol. I. *Report of the Conference*, p. 30.

## Abbreviated vowel characters (\* stands for any consonant character)

• a	ि i	ु u <sup>2</sup>	ृ r <sup>2</sup>	• e	• ो o
• ी ā <sup>1</sup>	• ी ī	• ू ū <sup>2</sup>		• ai	• ौ au

<sup>1</sup> with an additional mark: • ी ă

<sup>2</sup> exceptions: रु ru; रू rū; हृ hr

## Other symbols

• m	• m̄	• h	• 1
-----	------	-----	-----

<sup>1</sup> pronunciation without a vowel: क् k

## Consonant characters

क ka <sup>1</sup>	च cha	ट ṭa	त ta	प pa	य ya	श sha
ख kha <sup>1</sup>	छ chha	ठ ṭha	थ tha	फ pha <sup>1</sup>	र ra	ष sha
ग ga <sup>1</sup>	ज ja <sup>1</sup>	ड ḍa <sup>1</sup>	द da	ब ba	ल la	स sa
घ gha	झ (झ) jha	ढ ḍha <sup>1</sup>	ध dha	भ bha	व va	ह ha
ङ ṅa	ञ ja	ण ṇa	न na	म ma		

<sup>1</sup> dotted variants of the characters: क् qa; ख̣ kha; ग̣ ga; ज̣ ja; ड̣ ḍa; ढ̣ ḍha; फ̣ pha

## Notes

1. It is recommended that the vowel अ (a) should always be rendered in Roman except when it ends a name. If a name ends with a consonant, the consonant should be underlined. Such cases, however, will be very rare. E.g. कानपुर Kānapur (not Kāṅapura), जगत् Jagat̄.
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.

## Ligatures

Adscript form:

•	r-	Examples
		र्त rta, र्द rdva

Other ligatures (the list is not complete)

क	ख	क	क्	क्य	क्र	क्ल	क्व	क्ष	क्षण	क्षम	क्ष्य	क्ष्व
kka	kkha	kta	ktva	kya	kra	kla	kva	kṣha	kṣhṇa	kṣhma	kṣhya	kṣhva
कस	ख्य	ख्	ग्	ग्य	ग्न	ग्ल	ग्व	ग्य	ग्र	ग्ल	ग्न	ग्य
ksa	khya	kḥsha	gda	gdha	gna	gbha	gma	gya	gra	gla	ghna	ghya
घ	झ	ञ	ङ	ञ	च	छ	च्	च्य	च्र	ज	ज	ञ
ghra	ñka	ñkha	ṅga	ṅgha	chcha	chchha	chchhra	chya	chra	jja	jjha	jña
ज	झ	ञ	ञ (ञ)	ञ	ञ (ञ)	ट	ट	ट	ड	ड	ड	ण
jya	jra	jva	ñcha	ñchha	ñja	ṭta	ṭtha	ṭṭha	ḍda	ḍḍha	ḍḍha	ṇta
ण	ण	ण	त्	त्	त्	त्	त्	त्	त्	त्	त्	त्
ṇha	ṇḍa	ṇya	ṭka	ṭta	ṭtya	ṭtva	ṭtha	ṭṭa	ṭpa	ṭna	ṭnya	ṭya
त्र	त्रय	त्व	त्स	त्स	त्स्य	थ्य	थ्र	द्र	द	द	द	ध
tra	trya	tva	tsa	tsna	tsya	thya	thra	dga	dda	ddha	dbha	dma
घ	घ	घ	घ	घ	घ्य	घ्र	घ्व	न्	न्	न्	न्	न्
dya	dra	dva	dhna	dhma	dhya	dhra	dhva	nta	ntyā	ntra	ntha	nda
द्र	द्य	द्व	ध	ध	न्	न्	न्	न्	न्	न्	न्	न्
ndra	ndha	ndhya	ndhra	nna	nma	nya	nra	nva	nsa	nha	pta	ptya

प	प्य	प्र	प्ल	प्स	पय	फु	ब्द	ब	ब्य	ब्र	भ्य	भ्र
pna	pya	pra	pla	psa	phya	phla	bda	bna	bya	bra	bhya	bhra
म	म्व	मभ	म्ल	म्य	म्र	म्ल	म्व	म्ह	म्य	म्र	ल्क	ल्प
mna	mba	mbha	mmla	mya	mra	mlla	mva	mha	yya	yra	lka	lpa
ल	ल्य	ल्ल	ल्व	ल्ल	ल्य	ल्र	ल्व	ल्य	ल्र	ल्ल	ल्य	ल्र
lma	lya	lla	lva	lsa	vya	vra	shcha	shchya	shchha	shna	shya	shra
श	श्व (श्व)	ष्क	ष्क	ष्ट (ष्ट)	ष्ट	ष्ट	ष्ण	ष्ण	ष्ण	ष्ण	ष्ण	ष्ण
shla	shva	shka	shkra	shṭa	shtra	shtha	shṇa	shṇa	shpra	shma	shya	shva
स्क	स्ख	स्त	स्त्र	स्थ	स्न	स्प	स्फ	स्म	स्म्य	स्य	स्त्र	स्व
ska	skha	sta	stra	stha	sna	spa	spha	sma	smya	sya	sra	sya
ह	ह्य	ह्य	ह	ह	ह							
hna	hma	hya	hra	hla	hva							

## 5.

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

UN Hunterian

ā	= ā, a <sup>1</sup>	m̄	= n, m	sh	= s, sh	v	= w, v <sup>2</sup>
ḍ	= ḍ	ṅ	= ng	ṣh	= ṣh		
ḍh	= ḍh	ṇ	= n	ṭ	= ṭ		
ī	= ī, i <sup>1</sup>	ñ	= ny	ṭh	= ṭh		

<sup>1</sup> word-finally

<sup>2</sup> exceptionally, e.g. before *i*

The draft transliteration standard for Indian languages ISO CD 15919 (1998) gives differently the following transliteration equivalents to the characters of Indian languages.

UN ISO

ā	= ȃ	h	= k <sup>2</sup>	m̄	= m̄, ṁ, ṁ <sup>5</sup>	sh	= ś
ch	= c	j	= y <sup>3</sup>	o	= ȃ, o <sup>1</sup>	ṣh	= ṣ
chh	= ch	kh	= kh	ō	= ȃ	ṭ	= ṭ <sup>4</sup>
e	= ē, e <sup>1</sup>	l	= l	ṛ	= ṛ	y	= y <sup>3,6</sup>
ē	= e	l̄	= l̄	ṝ	= ṝ	ȳ	= ȳ <sup>6</sup>
ḡ	= ḡ	l̄, l̄	= l̄ <sup>2,4</sup>	ṝ	= ṝ		

<sup>1</sup> allowed for Northern Indian languages

<sup>2</sup> used in Tamil

<sup>3</sup> used in Bengali

<sup>4</sup> used in Malayalam

<sup>5</sup> used in Telugu

<sup>6</sup> used in Oriya

## Kannada

## 1.

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<sup>19</sup>.

<sup>19</sup> *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.

## 2.

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

## 3.

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).

## 4.

## Independent vowel characters and others symbols

ಅ a	ಇ i	ಉ u	ಎ ಏ	ಒ ಓ	ಠ ಠ
ಆ ā	ಈ ī	ಊ ū	ಏ e	ಓ o	ಃ ḥ
		ಋ ṛ	ಋ ai	ಌ au	ಌ ṛ

<sup>1</sup> pronunciation without a vowel: ಕ್ ಕ.

## Abbreviated vowel characters (\* stands for any consonant character)

• a <sup>1</sup>	• i <sup>2</sup>	• u <sup>3</sup>	• ಏ	• ಓ
• ā	• ī	• ū <sup>4</sup>	• e	• o
		• ṛ	• ai	• au

<sup>1</sup> some consonants do not have the diacritic, e.g. ಖ kha, ಜ ja, ಣ na, ಬ ba, ಲ la

<sup>2</sup> variations of the diacritic, e.g. ಗಿ gi, ದಿ di

<sup>3</sup> variations of the diacritic: ಪು pu, ಪು phu, ವು vu

<sup>4</sup> variations of the diacritic: ಪು ಪು, ಪು phū, ವು ವು

## Consonant characters

ಕ ka	ಚ cha	ಟ ṭa	ತ ta	ಪ pa	ಯ ya	ಶ sha
ಖ kha	ಛ chha	ಠ ṭha	ಥ tha	ಫ pha	ರ ra	ಷ ṣha
ಗ ga	ಜ ja	ಡ ḍa	ದ da	ಬ ba	ಲ la	ಸ sa
ಘ gha	ಞ jha	ಢ ḍha	ಧ dha	ಭ bha	ವ va	ಹ ha
ಣ ṇa	ಠ ṭha	ನ na	ನ na	ಮ ma		ಳ ḷa

## Subscript consonant characters

• k (• kr,	• chh	• dh	• n	• r	• ḷ
• kṣh)	• j	• ṇ	• p (• pr)	• l	
• kh	• jh	• t (• tr,	• ph	• v	
• ḅ	• ṅ	• ty)	• b	• sh	
• gh	• ṭ (• ṭr)	• th	• bh	• ṣh	
• ṅ	• ṭh	• d	• m	• s	
• ch	• ḍ	• ḍh	• y	• h	

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

## 5.

For differences between the UN system and the ISO transliteration draft ISO CD15919 see the section on the romanization of Hindi.

## Khmer

## 1.

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<sup>20</sup>.

## 2.

The system is used in many international cartographic products. In 1994 a Gazetteer of Cambodia was produced using the same 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.

## 3.

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.

## 4.

## Consonant characters

ក ká	ច chá	ដ dá	ត tá	ប bá <sup>1</sup>	យ yó	ស sá
ខ khá	ឆ chhá	ថ thá	ឌ thá	ផ phá	រ ró	ល lá
ក kó	ច chó	ដ dó	ត tó	ប pó	យ yó	ស sá
ក khó	ច chhó	ដ thó	ត thó	ប phó	យ vó	ស á
ង ngó	ន nhó	ន ná	ន nó	ម mó		

<sup>1</sup> if the character is used with a subscript character the romanization of the character will be p

## Subscript consonant characters

ក k	ច chh	ត th	ត t	ប ph	ស s
ខ kh	ច ch	ដ d	ត th	ម m	ល l
ក k	ច chh	ត th	ត t	យ y	ស s
ខ kh	ន nh	ន n	ប b	រ r	
ង ng	ន nh	ត t	ប ph	ល l	
ច ch	ដ d	ត th	ប p	យ y	

<sup>20</sup> 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.

Independent vowel characters

ភ ៉េ	ឌ ៉ូ, ៉ូ	ឃ ្រែ	ញ ្រៃ	ឌ ៉េ	ឌី(ឃ្រ) ៉ាវ
ឃ្រ ៉ៃ	ឌី ៉ាវ	ឃ្រ ្រែ	ញ ្រៃ	ឃ្រ ៉ៃ	ឌី ៉ាវ

Vocalic nuclei (• stands for any consonant character)

ភ <sup>1</sup> ៉ា—៉ា <sup>1</sup>	• ៉េ	• ៉ូ—៉ូ	ឃ្រ ៉ៃ	• ៉េ	ឃ្រ ៉ាវ—៉ាវ
• ៉េ—៉ៃ	• ៉ែ	• ៉ូ—៉ូ	ឃ្រ ្រៃ	• ៉ៃ—៉ៃ	
• ៉ៃ—៉ៃ	• ៉ូ—៉ូ	• ៉ៃ—៉ៃ	• ៉ៃ	• ៉ៃ—៉ៃ	

<sup>1</sup> with the character ឃ ្រ a ligature ឃ្រ is used

Note. Where two variants in romanization are given the first is to be used in the ៉-series and the second in the ៉-series

Shortened syllables and vocalic nuclei with anuswara or visarga

• ៉ា—៉ា	• ៉ៃ—៉ៃ, ៉ៃ <sup>2</sup>	• ៉ៃ—៉ៃ	• ៉ៃ—៉ៃ	• ៉ៃ	• ៉ៃ ៉ៃ—៉ៃ
• ៉ៃ—៉ៃ, ៉ៃ <sup>1</sup>	• ៉ៃ—៉ៃ	• ៉ៃ—៉ៃ	• ៉ៃ—៉ៃ	• ៉ៃ	• ៉ៃ ៉ៃ—៉ៃ

<sup>1</sup> ៉ៃ if followed by k, ng, h

<sup>2</sup> ៉ៃ after a short vowel or if followed by k, ng, h

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

Notes

1. Khmer consonants are divided into 2 series, the ៉-series and the ៉-series, as indicated in the consonant table. With vocalic nuclei these consonants may produce different romanizations, as shown in the abbreviated vowels table: ភ ៉ៃ, ឃ្រ ៉ៃ, ភ ៉ៃ, ឃ្រ ៉ៃ. A Khmer consonant in syllable-final position, not accompanied by a vowel marker or by •, should generally be romanized without a vowel letter following: ភ ភ ៉ៃ kək, ភ ៉ៃ ៉ៃ 'ăngk (exception: ឃ្រ ៉ៃ pōngrō, also written ឃ្រ ៉ៃ pōngrōr and ឃ្រ ៉ៃ pōngrōh).
2. The Khmer diacritical mark • or • written above an ៉-series consonant (except ឃ and ឃ្រ; see note 4) changes it to the ៉-series: ឃ្រ ៉ៃ héang. The diacritical mark • written above an ៉-series consonant changes it to the ៉-series: ឃ្រ ៉ៃ nhang. When either of these marks would conflict with another symbol written above a character, the mark • or • may be written in its place: ឃ្រ ៉ៃ hi, ឃ្រ ៉ៃ dāmrei. (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": ឃ្រ ៉ៃ khnang. There is no foot for the character ឃ្រ ៉ៃ. The "feet" • and • usually represent the characters ៉ៃ and ៉ៃ respectively, rather than ៉ៃ and ៉ៃ: ឃ្រ ៉ៃ kdei, ភ ៉ៃ ៉ៃ kânthéay, but ឃ្រ ៉ៃ ៉ៃ kântráb. A "foot" determines the series of the following vocalic nucleus unless it is a nasal (ឃ ៉ៃ ng, ឃ្រ ៉ៃ nh, ឃ ៉ៃ n, ឃ ៉ៃ m) or ៉ៃ y, ៉ៃ r, ឃ ៉ៃ l, ៉ៃ v, ឃ ៉ៃ s. In that case, the base consonant determines the vocalic series: ឃ្រ ៉ៃ khpōng, ឃ្រ ៉ៃ l'ă, ឃ្រ ៉ៃ thmă, ឃ្រ ៉ៃ svay. Syllable-final ៉ៃ and ៉ៃ sometimes appear as "feet": ឃ្រ ៉ៃ ៉ៃ or ឃ្រ ៉ៃ ៉ៃ, ៉ៃ ៉ៃ or ៉ៃ ៉ៃ teăng. This practice appears to be optional and such irregular Khmer spellings are not reflected

- as such in romanization.
4. The combination ឃ plus ្រ is written ហា *ba*. The latter character is a graphic device designed to prevent confusion with ហា *hā*. The characters ឃ and ហា with the diacritical mark ្រ are romanized *p* in the *ā*-series, rather than as *b* in the *ō*-series: ឃ្រ *pāng*, ហា្រ *patau*. The diacritical mark ្រ or ្រ is substituted where a conflict with another symbol written above a character would occur: ឃ្រ *pei*. The characters ឃ and ហា when accompanied by a "foot" are also romanized as *p* in the *ā*-series, although the Khmer diacritical mark is generally omitted: ឃ្រ *pléng*, ឃ្រ *p'ā*, ហា្រ *prā*.
  5. The *ā*-series consonant អ is romanized by means of an apostrophe ('): អ្រ *h'ek*, ចង្កៀត *chāng vā*, រើល *rō'ā*, រើ *'vei*, រើ *'āng*. In word-final position before a vowel, ' may be omitted: រើ *āng*.
  6. The Khmer diacritical mark ្រ appears only in two combinations: ្រ (examples: ហា្រ *bāt*, ខ្រ *khpōs*) and ្រ. The diacritical mark ្រ appears only in the combination ្រ. In the *ā*-series both ្រ and ្រ are romanized *ā*: ហា្រ and ហា្រ are both romanized *chā*. In the *ō*-series both ្រ and ្រ are romanized *ēā* when followed by *k*, *ng* or *h*; otherwise, they are romanized *ōā*: រើ *rōpeā*, រើ *mōā*, រើ *veāngk*, រើ *phōāpv*.
  7. The combination ្រ is romanized *r* before the consonant: រើ *thōrm*. The combination ្រ is romanized as *r* before the consonant preceded by a shortened vocalic nucleus: រើ *koār*.
  8. The symbol ្រ in syllable-initial position is ignored in romanization: រើ *sā*, រើ *sāsās*. In syllable-final position ្រ indicates that the consonant is vowelled, i.e. followed by *ā* in the *ā*-series, by *ō* in the *ō*-series: រើ *āmnā*, រើ *pōmō*.
  9. The diacritical mark ្រ (which appears above characters and/or vowel markers which are not vocalized) is ignored in romanization: រើ *bōry*, រើ *poūhī*, រើ *phumī*.
  10. The independent character រើ is romanized either *ō* or *ū*. A reference source should be consulted where doubt arises.

## Macedonian Cyrillic

### 1.

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 annex to the resolution<sup>21</sup>.

### 2.

The system is used in The former Yugoslav Republic of Macedonia<sup>22</sup> and in international cartographic products.

### 3.

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 *f* and *k*. These are romanized as *g* and *k*.

<sup>21</sup> 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.

<sup>22</sup> 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.

respectively, before e and н. But also r and к are occasionally used before e and н, so e.g. the name Makedonija goes back to Македонџа, not Македонџа in the Cyrillic alphabet.

А а	a	З з	z	4.	Н н	n	У у	u
Б б	b	С с	dz		Њ њ	nj	Ф ф	f
В в	v	И и	i		О о	o	Х х	h
Г г	g	Ј ј	j		П п	p	Ц ц	c
Д д	d	К к	k		Р р	r	Ч ч	č
Ѓ ѓ	d, g <sup>1</sup>	Л л	l		С с	s	Џ џ	dž
Е е	e	Љ љ	lj		Т т	t	Ш ш	š
Ж ж	ž	М м	m		Ќ к	ć, k <sup>1</sup>		

<sup>1</sup> the second variant is used before e and н

## Malayalam

### 1.

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<sup>23</sup>.

### 2.

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

### 3.

Malayalam (Malayālam) 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 1980'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.

### 4.

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.

#### Independent vowel characters

അ	a	ഇ	i	ഉ	u	ഈ	r	ഏ	e	ഈ	o	ഔ	au
ആ	ā	ഈ	ī	ഊ	ū	ഐ	č	ഐ	ai	ഓ	o		

#### Abbreviated vowel characters (\* stands for any consonant character)

*	a	ഈ	i	*ു	u <sup>1</sup>	*ു	r <sup>2</sup>	ഏ	e	ഐ	o	*ൗ	au <sup>4</sup>
*ഓ	ā	ഈ	ī	*ു	ū <sup>2</sup>	ഐ	č	ഐ	ai	ഓ	o		

<sup>1</sup> traditional script form: ു - ഐ khu, ഐ mu, etc. Exceptions in the older spelling: ക്കു (കകു kku, ക്കു kku), ഗുഗു, etc.

<sup>23</sup> *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.



കു chhu, കു ju, ഞ nu, ത tu, ന nu (ന nu), ദ bhu, ഞ ru, ശു shu, ഹ hu

<sup>2</sup> traditional script form: ് - ഖ khū, മ mū, etc. Exceptions in the older spelling: ക kū, ത tū, ച chhū, കു jū, ഞ nu, ത tū, ന nū, ദ dhū, ശ rū, ശു shū, ഹ hū

<sup>3</sup> traditionally written in conjunction with the consonant character: ക്ക kr

<sup>4</sup> in the older spelling: ഹ്ഹ

**Other symbols**

° o    ṁ    ° ː    ḥ    ˆ    (1)

<sup>1</sup> indicates absence of the inherent short [a] vowel: ക k , ന n. Special forms of consonants with that symbol:

ൺ n	ര r	ശ (s)
ൻ n	രി l	

**Consonant characters**

ക	ka	ച	cha	ട	ta	ത	ta	പ	pa	യ	ya	സ	sa
ഖ	kha	ഛ	chha	ഠ	tha	ഥ	tha	ഫ	pha	ര	ra	റ	ra
ഗ	ga	ജ	ja	ഡ	da	ഭ	da	ബ	ba	ല	la	ക്ഷ	ksha
ഘ	gha	ശ്ശ	jha	ഢ	ḍha	ധ	ḍha	ഭ	bha	വ	va	ള	la
ങ	ṅa	ഞ	ṅa	ണ	ṇa	ന	na	മ	ma	ശ	sha	ഴ	ra <sup>1</sup>
										ഷ	ṣha	ഠ	ra <sup>1</sup>

<sup>1</sup> special ligature: ഠാ ṭa

**Ligatures**

Adscript forms of some consonants

Example

- ° r-      റ rga (in modern usage it is substituted: റ്റ rrga)
- ° -r      ക kra (older style: ക്ര)
- ° ː y      യ ya
- ° ː v      വ va

**Other ligatures**

കക	കു	കത	ക്ഷ	കി	കു	ഗ	ഗദ്ധ	ഗ	ഗ	ഗ	ഗ	ഗ	ക
kka	kta	kta	ksha	kla	gga	gda	gdda*	gna	gma	gla	nka	ka	nka
ച	ച	ച	ഛ	ഛ	ജ	ജ	ജ	ജ	ജ	ജ	ജ	ജ	ജ
chha	chha	chchha	jjha	jjha	ṅha	ṅchha*	ṅja*	ṅna	ṅma	ṅta	ṅta	ṅta	ṅta
ഞ	ണ	ണ്ണ	മ്മ	ത്ത	ഥ	ത	ത	ത	ത	ത	ത	ത	ത
ṅtha*	ṅda	ṅṅa	ṅma	ṅta	ttha	tta*	tbha	tma	tta	dda	dda	dda	dda
ന	ന	ന	ന്ദ	ന്ന	ന	ന	ന്റ	പ	പ	പ	പ	പ	പ
nta	ntha	nda	ndha	nna	nma	nra	pta*	pna*	ppa	pla	psa*	pta*	pta*
ബ	ബ	ബ	ബ	ബ	മ്പ	മ്മ	മ	മ	മ	മ	മ	മ	മ
bda	bdha	bbha	bba	mna*	mpa	mma	mha	yka*	ykka	yta*	yta*	yta*	yta*
യ	യ	യ	ല	ല	ല	ല	ല	ല	ല	ല	ല	ല	ല
ypa*	ympa*	yya	lka*	lkka*	lpa	lma*	lla	vva	shcha	shla	shsha	shpa*	shpa*
ഷ	ഷ	ഷ	ഷ	ഷ	ഷ	ഷ	ഷ	ഷ	ഷ	ഷ	ഷ	ഷ	ഷ
ṣha*	ṣhta	ṣhha*	ṣhna*	ṣhpa*	ṣhpha*	ṣhma*	ska*	sta*	stha	sna*	sna*	sna*	sna*
സ	സ	സ	സ	സ	ഹ	ഹ	ഹ	ഹ	ഹ	ഹ	ഹ	ഹ	ഹ
spha*	spha*	sma*	sla	ssa	hna	hma	hla	lla	lka*	lcha*	lpa*	lpa*	lpa*
ല	ല	ല	ല	ല	ല	ല	ല	ല	ല	ല	ല	ല	ല
lma*	lpa*												

Ligatures marked with an asterisk are thought to be obsolete since the introduction of the new spelling.

## 5.

For differences between the UN system and the ISO transliteration draft ISO CD15919 see the section on the romanization of Hindi.

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

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

## Marathi

## 1.

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<sup>24</sup>.

## 2.

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

## 3.

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

## 4.

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 la

ऐ ā (abbreviated: कै kā)

## 2. Different shapes of some characters and ligatures:

Hindi	Marathi	Roman
ख	ख	kha
ल	ळ	la
क्त	क्त	kta
क्क	क्क	kka

3. Ligatures may be formed with *r* as the first component: च rya, र्हा rha4. Typical are ligatures ending in *h*: र्ह nh, म्ह mh, ल्ह lh, व्ह vh

## 5. The following independent characters are sometimes used: सि i, औ ī, उ u, ऊ ū, ऐ e, औ ai.

## Mongolian

(in China)

## 1.

The United Nations resolution III/8 in 1977 recognized the Scheme for a Chinese Phonetic Alphabet (Pinyin) as

<sup>24</sup> 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.

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*<sup>25</sup>.

## 2.

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

## 3.

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.

## 4.

ᠠ	a	ᠬ	h	ᠣ	o	ᠠ	w
ᠪ	b	ᠢ	i	ᠮ	p	ᠬ	x
ᠴ	c	ᠵ	j	ᠬ	q	ᠠ	y
ᠳ	d	ᠬ	k	ᠠ	r	ᠵ	z
ᠡ	e	ᠠ	l	ᠰ	s	ᠣ	o [ō]
ᠮ	f	ᠮ	m	ᠲ	t	ᠤ	u [ū]
ᠭ	g	ᠨ	n	ᠮ	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 square brackets) is used for recording the pronunciation of place-names while the first form is for general use.
3. Due to the image quality of the original table there may occur some misinterpretations of certain Mongolian characters.

## Nepali

## 1.

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<sup>26</sup>.

## 2.

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.

## 3.

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

## 4.

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

<sup>25</sup> Fourth United Nations Conference on the Standardization of Geographical Names. Geneva, 24 August - 14 September 1982. Vol. II. *Technical papers*, pp. 121-125.

<sup>26</sup> 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

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

## 5.

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

UN	BGN/PCGN				
<i>m̐</i>	= <i>m̐</i>	<i>r̐</i>	= <i>ri</i> , <i>ir</i> <sup>1</sup>	<i>ṣh</i>	= <i>ṣ</i>
<i>n̐</i>	= <i>ng</i>	<i>r̐̄</i>	= <i>rī</i> , <i>īr</i> <sup>1</sup>	<i>v</i>	= <i>w</i>
<i>f</i>	= <i>r̐</i>	<i>sh</i>	= <i>ś</i>		

<sup>1</sup> the second variant is used in word-medial position

## Oriya

## 1.

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<sup>27</sup>.

## 2.

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

## 3.

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).

## 4.

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

## Independent vowel characters

ଅ a	ଇ i	ଉ u	ଋ ṛ	ଏ e	ଓ o <sup>1</sup>
ଆ ā	ଈ ī	ଊ ū	(ଋ ṛ, ୠ ḷ)	ଐ ai	ଔ au

<sup>1</sup> special form: ଌ va (not present in the original table)

## Abbreviated vowel characters (\* stands for any consonant character)

* a	* i	* u	* ṛ	୦ e	୦ o
* ā	* ī	* ū		୦ (୦) ai	୦ (୦) au

## Other symbols

* m̐	* m̐̄	* ṣ h	* ṛ
------	-------	-------	-----

<sup>1</sup> pronunciation without a vowel: ୠ k

<sup>27</sup> *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.

**Consonant characters:**

କ ka	ଚ cha	ଟ ta	ଡ ta	ଢ pa	ଢ ya <sup>2</sup>	ଶ sha
ଖ kha	ଛ chha	ଠ tha	ଢ tha	ଢ pha	ର ra	ଷ sha
ଗ ga	ଜ ja	ଢ଼a <sup>1</sup>	ଢ da	ଭ ba	ଲ la	ସ sa
ଘ gha	ଞ jha	ଢ଼ha <sup>1</sup>	ଧ dha	ଭ bha	ଞ ja	ହ ha
ଣ na	ଢ na	ଶ na	ନ na	ମ ma		କ୍ଷ sha

<sup>1</sup> dotted forms: ଢ ta, ଢ tha

<sup>2</sup> special form: ଢ ya

**Subscript consonant characters**

କ୍ କ	ଚ୍ ଚ	ଟ୍ ଟ	ଢ୍ ଢ	ଢ୍ ଢ	ଭ୍ ଭ	ଶ୍ ଶ
ଖ୍ ଖ	ଜ୍ ଜ	ଢ଼୍ ଢ଼	ଢ଼୍ ଢ଼	ଭ୍ ଭ	ର୍ ର	ଷ୍ ଷ
ଗ୍ ଗ	ଞ୍ ଞ	ନ୍ ନ	ନ୍ ନ	ମ୍ ମ	ସ୍ ସ	
ଘ୍ ଘ	ଞ୍ ଞ	ତ୍ ତ	ପ୍ ପ	ୟ୍ ଯ		
ଚ୍ ଚ	ଢ୍ ଢ	ଥ୍ ଥ	ପ୍ ପ	ର୍ ର		

**Ligatures (the list is not complete)**

କ୍ କ	କ୍ କ	କ୍ କ	କ୍ କ	କ୍ କ	କ୍ କ	କ୍ କ	କ୍ କ	କ୍ କ	କ୍ କ	କ୍ କ	କ୍ କ
kka	kta	kla	ksha	kshna	ksa	nka	nksha	nkha	nga	ngha	chcha
ଚ୍ ଚ	ଚ୍ ଚ	ଚ୍ ଚ	ଚ୍ ଚ	ଚ୍ ଚ	ଚ୍ ଚ	ଚ୍ ଚ	ଚ୍ ଚ	ଚ୍ ଚ	ଚ୍ ଚ	ଚ୍ ଚ	ଚ୍ ଚ
chhchha	jja	jjha	jja	ncha	nchha	nja	njha	tta	dda	nta	ntha
ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ
nda	ndha	nna	tka	tta	tpa	tra	tta	dga	dda	ddha	dbha
ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ	ନ୍ ନ
nta	ntra	ntha	nda	ndha	nna	pta	psa	bda	bba	mpa	mpha
ମ୍ ମ	ମ୍ ମ	ମ୍ ମ	ମ୍ ମ	ମ୍ ମ	ମ୍ ମ	ମ୍ ମ	ମ୍ ମ	ମ୍ ମ	ମ୍ ମ	ମ୍ ମ	ମ୍ ମ
mba	mbha	mma	lka	lpa	lpha	lbha	shcha	shchha	shka	shtha	shtha
ଷ୍ ଷ	ଷ୍ ଷ	ଷ୍ ଷ	ଷ୍ ଷ	ଷ୍ ଷ	ଷ୍ ଷ	ଷ୍ ଷ	ଷ୍ ଷ	ଷ୍ ଷ	ଷ୍ ଷ	ଷ୍ ଷ	ଷ୍ ଷ
shna	shpa	shpha	ska	skha	sta	stra	stha	spa	spha	hna	hla

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

**5.**

For differences between the UN system and the ISO transliteration draft ISO CD15919 see the section on the romanization of Hindi.

**Persian****1.**

The United Nations recommended system was approved in 1967 (U/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 has not been published among the materials of the United Nations conferences on the standardization of geographical names, nor was the original table available at the time of writing the report. Based on indirect sources, incl. national reports of the Islamic Republic of Iran and the lists of geographical names supplied thereby, it is assumed that the system is identical to that used by the BGN/PCGN.

**2.**

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

3.

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 character sequences (*kh, sh, th, zh*) which may also point to combinations of Persian characters in addition to the respective single characters.

4.

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

ا	(omitted)	خ	kh	ص	ṣ	ك	k
ب	b	د	d	ض	ẓ	گ	g
پ	p	ذ	ẓ	ط	ṭ	ل	l
ت	t	ر	r	ظ	ẓ	م	m
ث	ṯ	ز	z	ع	ʿ	ن	n
ج	j	ژ	zh	غ	gh	و	v
چ	ch	س	s	ف	f	ه	h
ح	ḥ	ش	sh	ق	q	ی	y

Vowels (\* stands for any consonant)

•	a	آ / اِی	á	•	آ / اِی	ey	•	(doubling of consonant)
•	e	آ	'ā, ā <sup>1</sup>	•	آ / اِو	ow	•	-e <sup>2</sup> , -ye <sup>3</sup>
•	o	آ / اِی	ī	•		(omitted)	•	
•	ā	آ / اِو	ū	•			•	

<sup>1</sup> word-initially

<sup>2</sup> after a consonant (excl. -ah)

<sup>3</sup> after a vowel (see also note 2)

Notes

1. The adjectival ending of Arabic origin *-یة* in Persian is romanized *-yeh*. 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 *t* or *o* it is written with *ی*, e.g. *پای آب* Pā-ye Āb. After final *ی* and *•* 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.

5.

The Persian language used in Afghanistan is officially referred to as *Darī* (in other sources also *Kabuli* or *Farsi-Kabuli*). Different from Persian, the *Darī* alphabet contains a character *خ* *ts* borrowed from *Pashto*. The consonant character *و* has the sound value [w] instead of [v] and the word-final combination *••* is pronounced [a] instead of [e]. In romanization also vowels *i, o* or *u* may occur.

The simplified transliteration of Persian in the standard ISO 233-3:1999 has the following equivalents to the UN system:

UN	ISO						
ā	= ā	ḥ	= h	ṣ	= ṣ	ṭ	= ṭ
ch	= č*	ī	= y?	ṣ	= ṣ	ū	= v?
gh	= ġ	kh	= k	sh	= š	ẓ	= z

z = ż                      ż = z\*                      zh = z\*

Note. The ISO publication contains some evident 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

### 1.

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<sup>28</sup>.

### 2.

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

### 3.

Punjabi (Panjābī) 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 exist some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (combinations with subscript consonants vs. character sequences).

### 4.

ੳ (u,ū) <sup>1</sup>	ਕ ka	ਚ cha	ਟ ta	ਤ ta	ਪ pa	ਯ ya
ਅ a (ā,ai,au) <sup>1</sup>	ਖ kha <sup>2</sup>	ਛ chha	ਠ tha	ਥ tha	ਫ pha <sup>2</sup>	ਰ ra
ੲ (i,ī,e) <sup>1</sup>	ਗ ga <sup>2</sup>	ਜ ja <sup>2</sup>	ਡ da	ਦ da	ਬ ba	ਲ la
ਸ s <sup>2</sup>	ਘ gha	ਝ jha	ੳ dha	ੲ dha	ਭ bha	ਵ va
ਹ h	ਙ na	ਞ na	ਣ na	ਨ na	ਮ ma	ੜ za

<sup>1</sup> the characters are used only in combination with vowel marks: ਲ਼ u, ਲ਼ u, ਅ a, ਆ ā, ਐ ai, ਐ au, ਏ i, ਏ i, ਏ e; exceptionally: ਲ਼ o

<sup>2</sup> dotted variants: ਸ sha, ਖ kha, ਗ ga, ਜ za, ਫ fa

#### Abbreviated vowel characters (\* stands for any consonant character)

• a	• i	• u	• e	• o
• ā	• ī	• ū	• ai	• au

#### Other symbols

• m <sup>1</sup>	• m <sup>2</sup>	• s
------------------	------------------	-----

<sup>1</sup> used if it is preceded by short vowels (a, i, u) and ū, excluding the independent vowel character ਲ਼ (u, ū)

<sup>2</sup> used in all other occasions: ਸਿਰ sarig, ਸਿਰ sitig, ਬੁਰ burtā, ਬੁਰ būrd, ਸਿਰ sijnjan; ਸਿਰ jhomkā

<sup>3</sup> marks doubling of the following consonant: ਨਿਕ nikkā, ਲੱਭ ਲabbhbhā, ਕੁੱਟ kuttā, ਹਿੱਸ hissā, ਲੱਮ lamma

#### Subscript consonant characters

• -h	• -v	• -r	• -y
------	------	------	------

Note. Absence of the inherent vowel (-a) is not marked in the spelling in any way except for the combinations

<sup>28</sup> 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.

with subscript characters and those which are doubled by <sup>29</sup> (*adhaka*).

## 5.

For differences between the UN system and the ISO transliteration draft ISO CD15919 see the section on the romanization of Hindi.

## Russian

## 1.

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 annex to the resolution<sup>29</sup>.

## 2.

The system is used in the Russian Federation and increasingly in international cartographic products. It was noted in the materials of the 3rd session of the Eastern Europe, Northern and Central Asia Division of UNGEGN in Moscow (June 1999) that a new version of the rules for the transliteration of the geographical names of the Russian Federation had been prepared. This would not change the content of the romanization table.

## 3.

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. E.g. a geographical name Ojusardah of the Sakha Republic (Yakutia) reverts to Cyrillic Ойусардах, not Оюсардах as would seem logical.

## 4.

А а	а	И и	и	С с	с	Ъ ъ	ъ
Б б	б	Й й	й	Т т	т	Ы ы	ы
В в	в	К к	к	У у	у	Ь ь	ь
Г г	г	Л л	л	Ф ф	ф	Э э	э
Д д	д	М м	м	Х х	х	Ю ю	ю
Е е	е	Н н	н	Ц ц	ц	Я я	я
Ё ё	ё	О о	о	Ч ч	ч		
Ж ж	ж	П п	п	Ш ш	ш		
З з	з	Р р	р	Щ щ	щ		

## 5.

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

е (е)	е, ye <sup>1</sup>	й (й)	у	ч (ч)	ch	э (э)	е
ё (ё)	ё, ye <sup>1</sup>	х (х)	kh	ш (ш)	sh	ю (ю)	yu
ж (ж)	zh	ц (ц)	ts	щ (щ)	shch	я (я)	ya

<sup>1</sup> initially, after the vowel characters and after й, ъ, and ь

The transliteration of Cyrillic letters contained in the standard ISO 9:1995 provides for the romanization, as a single block, of the following letters differently from the UN system:

щ (щ)	š	ю (ю)	û	я (я)	â
-------	---	-------	---	-------	---

<sup>29</sup> Fifth United Nations Conference on the Standardization of Geographical Names. Montreal, 18-31 August 1987. Vol. I. Report of the Conference, pp. 40-41.



## Serbian

### 1.

The United Nations recommended system was approved in 1977 (III/11), based on the romanization of the Serbo-Croatian Cyrillic alphabet that had long been employed in official gazetteers and maps of Yugoslavia. The table was published as annex to the resolution<sup>30</sup>.

### 2.

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

### 3.

Serbian uses the Cyrillic script which is alphabetic. The romanization table is unambiguous and can be applied automatically. The system is reversible.

### 4.

А а	а	З з	з	Ђ ђ	њ	Ф ф	ф
Б б	б	И и	и	О о	о	Х х	х
В в	в	Ј ј	ј	П п	п	Ц ц	ц
Г г	г	К к	к	Р р	р	Ч ч	ч
Д д	д	Л л	л	С с	с	Џ џ	џ
Ђ ђ	ђ (Đ)	Љ љ	љ	Т т	т	Ш ш	ш
Е е	е	М м	м	Ђ ђ	џ		
Ж ж	ж	Н н	н	У у	у		

## Tamil

### 1.

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<sup>31</sup>.

### 2.

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

### 3.

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.

### 4.

#### Independent vowel characters

அ a	இ i	உ u	ஏ e	ஐ ai	ஓ o	ஃ h
ஆ ā	ஈ ī	ஊ ū	ஏ e	ஔ o	ஔ au	

<sup>1</sup> absent in the original table

<sup>30</sup> *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.

<sup>31</sup> *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.

## Abbreviated vowel characters (\* stands for any consonant character)

• a	ஏ <sup>2</sup>	ஔ <sup>4</sup>	ஔ <sup>6</sup>	ஔ <sup>6</sup>	ஔ <sup>6</sup>
• ா <sup>1</sup>	ஔ <sup>3</sup>	ஔ <sup>3</sup>	ஔ <sup>3</sup>	ஔ <sup>3</sup>	ஔ <sup>3</sup>

<sup>1</sup> exceptions in older and Classical literature: ஔ<sup>3</sup> நா, ஔ<sup>3</sup> ரா, ஔ<sup>3</sup> நா

<sup>2</sup> variation: ஔ<sup>3</sup> தி

<sup>3</sup> variation: ஔ<sup>3</sup> தி

<sup>4</sup> variations: ஔ<sup>3</sup> ku, ஔ<sup>3</sup> nu, ஔ<sup>3</sup> chu, ஔ<sup>3</sup> nu, ஔ<sup>3</sup> tu, ஔ<sup>3</sup> nu, ஔ<sup>3</sup> pu, ஔ<sup>3</sup> mu, ஔ<sup>3</sup> yu, ஔ<sup>3</sup> ru, ஔ<sup>3</sup> lu, etc.

<sup>5</sup> variations: ஔ<sup>3</sup> kī, ஔ<sup>3</sup> nū, ஔ<sup>3</sup> chū, ஔ<sup>3</sup> nū, ஔ<sup>3</sup> tū, ஔ<sup>3</sup> nū, ஔ<sup>3</sup> pū, ஔ<sup>3</sup> mū, ஔ<sup>3</sup> yū, etc.

<sup>6</sup> exceptions in older and Classical literature: ஔ<sup>3</sup> lai, ஔ<sup>3</sup> lai, ஔ<sup>3</sup> nai, ஔ<sup>3</sup> nai

## Other symbols and syllables

• ஔ<sup>3</sup> shrī

<sup>1</sup> pronunciation without a vowel: ஔ<sup>3</sup> k

## Consonant characters

க ka	ட ta	ப pa	ய ya	ழ ja	ஜ ja <sup>1</sup>	ஹ ha <sup>1</sup>
ங na	ண na	ம ma	ர ra	ள la	ஷ sha <sup>1</sup>	க்ஷ ksha
ச cha	த ta		ல la	ற ra	ஷ sha	(ற ra)
ந na	ந na		வ va	ன na	ஸ sa <sup>1</sup>	(ன na)

<sup>1</sup> characters borrowed from the Grantha script

## 5.

For differences between the UN system and the ISO transliteration draft ISO CD15919 see the section on the romanization of Hindi.

## Telugu

## 1.

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<sup>32</sup>.

## 2.

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

## 3.

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).

## 4.

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

<sup>32</sup> *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.

**Independent vowel characters**

అ a	ఆ ā	ఐ ī	ఎ e	ఒ o
ఇ i	ఊ ū	ఋ ṛ	ఏ e	ఓ o
			ఐ ai	ఔ au

**Vowel characters (• stands for any consonant character)**

• a	• ā	• ī	• e	• o
• i	• ū	• ṛ	• e	• o
			• ai	• au

Note. Due to technical reasons, the diacritics 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	క్రు	కె	కే	కై	కో	కొ	కౌ
ka	kā	ki	kī	ku	kū	kr	kē	ke	kai	kō	ko	kau

<sup>1</sup> variations: చి chi, చీ chī, యి yi, లి li, etc.

<sup>2</sup> variations: జు ju, జూ ju, వు vu, etc.

<sup>3</sup> variations: మొ mō, యొ yō

<sup>4</sup> variations: మో mo, యో yo

**Other symbols**

◌ṁ	◌ṁ <sup>1</sup>	◌ḥ	◌ <sup>2</sup>
----	-----------------	----	----------------

<sup>1</sup> signified historic nasal sound, now obsolete

<sup>2</sup> end-of-syllable mark (i.e. a consonant without a vowel): రాంగపూర్ Raṅgapūr

**Consonant characters**

క ka	చ cha	ట ṭa	ఠ ṭha	ప pa	య ya	ష ṣha
ఖ kha	ఛ chha	థ ṭha	ఢ ḍha	ఫ pha	ర ra <sup>1</sup>	స sa
గ ga	జ ja	డ ḍa	ఢ ḍa	బ ba	ల la	హ ha
ఘ gha	ఝ jha	ఢ ḍha	ఢ ḍha	భ bha	వ va	ళ ḷa
ఞ ṅa	ఞ ṅa	ణ ṇa	న na	మ ma	శ sha	

<sup>1</sup> variant: ర ra

**Subscript consonant characters**

◌k	◌n	◌y	◌l
◌c	◌m	◌r	◌v

**5.**

For differences between the UN system and the ISO transliteration draft ISO CD15919 see the section on the romanization of Hindi.

**Thai****1.**

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<sup>33</sup>.

2.

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.

3.

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. For transliteration a different system is used.

4.

ก k (k)	ช ch (t)	ด d (t)	ต d (t)	บ b (p)	ย y	ส s (t)
ข kh (k)	ฌ ch (t)	ฎ t (t)	ฏ t (t)	ป p (p)	ร r (r)	ษ s (t)
ค kh (k)	ช ch (t)	ฐ th (t)	ถ th (t)	ผ ph (p)	ฤ ru, ri, ree <sup>2</sup>	ศ s (t)
ก kh (k)	ซ s (t)	ฑ th, d (t)	ฒ th (t) <sup>1</sup>	ฝ f (p)	ฎ ru <sup>2</sup>	ห h
ก kh (k)	ฌ ch (t)	ฒ th (t)	ฐ th (t)	พ ph (p)	ล l (n)	ฬ l (n)
ข kh (k)	ญ y (n)	ณ n (n)	น n (n)	ฟ f (p)	ฤ lu <sup>2</sup>	อ o
ง ng (ng)				ภ ph (p)	ฎ lu <sup>2</sup>	ฮ h
				ม m (m)	ว w	

<sup>1</sup> combination ฬ is romanized s if so pronounced

<sup>2</sup> these Thai characters are traditionally considered as vowel characters; for revised romanizations see below  
 Note. 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 brackets.

Vocalic-nuclei (\* stands for any consonant character)

๑๒	a	๑๓	i	๑๔	e	๑๕	o	๑๖	ua*	๑๗	ai
๑๘	a	๑๙	u*	๑๒๐	ae	๑๒๑	o	๑๒๒	ua*	๑๒๓	ai
๑๒๔	a	๑๒๕	u*	๑๒๖	ae	๑๒๗	oe	๑๒๘	ua	๑๒๙	ai
๑๓๐	a	๑๓๑	u	๑๓๒	ae	๑๓๓	oe	๑๓๔	ua	๑๓๕	ao
๑๓๖	an	๑๓๗	u	๑๓๘	o	๑๓๙	oe	๑๔๐	ua	๑๔๑	ao
๑๔๒	am	๑๔๓	e	๑๔๔	o	๑๔๕	ia	๑๔๖	ai	๑๔๗	ui
๑๔๘	i	๑๔๙	e	๑๕๐	o	๑๕๑	ia	๑๕๒	ai	๑๕๓	oi

<sup>33</sup> *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.*

๑๐๒	oi	๑๐๒๒	uai*	๑๐๓	iu*	๑๐๔	eo	๑๐๕	aeo
๑๐๒	oei	๑๐๓	uai	๑๐๔	eo	๑๐๕	aeo	๑๐๖	ieo*

\* romanization equivalents marked with an asterisk have been revised in the 1997 version

Notes<sup>34</sup>

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 which are combinations of words and the last consonant of words pronounced should be transliterated as single words to retain their original pronunciation and forms, e.g. *Ratchakhrām* is transcribed from the Thai words which, without the above notation, would be *rat khram*; *Aranyaprathet* is transcribed instead of *Aran Prathet*.
3. 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*.
4. A word or a group of words followed by a *matyamok* (๑) is written twice according to the reading rules, e.g. *ท่าบอย ๑* *tham bot-boi*, *ไฟไหม้ ๑* *fai mai fai mai*.
5. A word or a group of words followed by a *paiyanoi* (๑), a shortened form of well-known word or expression, is to be written in full according to the reading rules, e.g. *กรุงเทพมหานคร* *Krung Thep Maha Nakhon*.
6. A Thai consonant character with no accompanying vowel mark or diacritical mark may be romanized with the vowel letter *a* or *o*: ปทม *pathom*, พลบ *phlop*.
7. ㅎ at the beginning of a syllable and serving as a tone mark should not be romanized: ท๑ wo, หลบ *lup*; otherwise, ㅎ should be romanized *h*: ท๑ huai, โหว๑ wohan.
8. The vowel shortening mark ๑ and the tone marks ๑, ๑, ๑, and ๑ should not be romanized: ก๑ ko, ย๑ yih๑, กระ๑ krat๑.

5.

The 1997 revised version<sup>35</sup> of the official Thai romanization system gives different Roman equivalents to the following Thai characters:

๑	rue,ri,roe	๑	lue	๑	ue	๑๐๓	uea	๑๐๒	ueai	๑๐๓	iao
๑๑	rue	๑๑	lue	๑	ue	๑๐๒	uea	๑๐๓	io		

The transliteration of Thai in the standard ISO 11940:1998<sup>36</sup> 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

<sup>34</sup> Notes 1-3 have been taken from the original paper (1967), 4 and 5 from the revised version (1997) and 6-8 have been adapted from the BGN/PCGN 1970 romanization system, to cover the most essential features of the Thai script.

<sup>35</sup> 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.

<sup>36</sup> 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.

starting by the uppermost mark and finishing with the subscript mark. As an example, here are some transliteration equivalents: ཀ k, ཕ kh, བ kh̄, ཎ kh, ཏ k'h, ཡ kh, འ ng; ཨ a, ཨ̄ a, འ̄ a, འ̄̄ a, འ̄̄̄ a, འ̄̄̄̄ a, etc.

## Tibetan

### 1.

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<sup>37</sup>.

### 2.

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

### 3.

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.

### 4.

#### Consonant characters

ཀ ga	ཅ ja	ཉ da	བ ba	ཅ za	མ xa	ལ la
ཁ ka	ཆ qa	ཐ ta	པ pa	ཚ ca	ཞ sa	ལ xa
ཀ ka, ga <sup>1</sup>	ཅ qa, ja <sup>1</sup>	ཉ ta, da <sup>1</sup>	བ pa, ba <sup>1</sup>	ཅ ca, za <sup>1</sup>	མ a	ལ sa
ང nga	ཧ nya	ཨ na	མ ma	ཡ wa	ལ ya	ཧ ha
					ར ra	ལ a

<sup>1</sup> if the letter is not accompanied by a prefixed letter or a superadded 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)

ཀྱ gya	ཀྲ kya, gya <sup>1</sup>	ཅྱ cha	ཅྲ sha
ཁྱ kya	ཁྲ zha	ཐྱ cha, zha <sup>1</sup>	ཐྲ lha

<sup>1</sup> if the combination is not accompanied by a prefixed letter or a superadded 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

#### Vowel characters (ཀ stands for any consonant character)

ཀ a	ཀི i	ཀུ u	ཀེ e	ཀོ o
-----	------	------	------	------

#### Other symbols

• syllable separation character, e.g. འགུ Nag Qu

#### Syllable endings

ལ a	ལ་ལ au	ལག ag	ལག་ལ ag
-----	--------	-------	---------

<sup>37</sup> Fourth United Nations Conference on the Standardization of Geographical Names. Geneva, 24 August - 14 September 1982. Vol. II. Technical papers, pp. 121-125.

ཨ	ang	ཨ	i	ཨ	ur	ཨ	ér
ཨ་	ang	ཨ་	i	ཨ་	ü	ཨ་	én
ཨམ	ab	ཨམ	ing	ཨམ་	ü	ཨམ་	o
ཨམ་	ab	ཨམ་	ing	ཨམ་	ü	ཨམ་	og
ཨམམ	am	ཨམམ	ib	ཨམམ་	ü	ཨམམ་	og
ཨམམ་	am	ཨམམ་	ib	ཨམམ་	ün	ཨམམ་	ong
ཨར	ar	ཨར	im	ཨར་	é	ཨར་	ong
ཨར་	ai [ä]	ཨར་	im	ཨར་	é	ཨར་	ob
ཨརམ	ai [ä]	ཨརམ	ir	ཨརམ་	é	ཨརམ་	ob
ཨར་	ai [ä]	ཨར་	in	ཨརམ་	ég	ཨརམ་	om
ཨརམ	ai [ä]	ཨརམ	u	ཨརམ་	ég	ཨརམ་	om
ཨརམ་	ain [än]	ཨརམ་	ug	ཨརམ་	é	ཨརམ་	or
ཨི	i	ཨི་	ug	ཨི་	é	ཨི་	oi [ö]
ཨི་	i	ཨི་	ung	ཨི་	éng	ཨི་	oi [ö]
ཨིམ	i	ཨིམ	ung	ཨིམ་	éng	ཨིམ་	oi [ö]
ཨིམ་	iu	ཨིམ་	ub	ཨིམ་	éb	ཨིམ་	oi [ö]
ཨིམམ	iu	ཨིམམ	ub	ཨིམམ་	éb	ཨིམམ་	oin [ön]
ཨིམམ་	ig	ཨིམམ་	um	ཨིམམ་	ém		
ཨིམམམ	ig	ཨིམམམ	um	ཨིམམམ་	ém		

## Notes

1. Prefixed consonant characters ཀ (k), ཉ (l), ར (p), མ (m), འ (a), and superadded consonant characters ར་ (r), འ་ (l), མ་ (s) remain mostly unpronounced and are not reflected in the romanization.
2. Subscript characters ར་ (w), ར་ (y), ར་ (r), ར་ (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 letters འ and ར are treated as without initials when transcribed with Pinyin.
5. When the prefixed letters མ and འ or the superadded consonant ར 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 square brackets) is used for recording the pronunciation of place-names while the first form is for general use.

## Uighur

## 1.

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*<sup>38</sup>.

## 2.

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

<sup>38</sup> Fourth United Nations Conference on the Standardization of Geographical Names: Geneva, 24 August - 14 September 1982. Vol. II. *Technical papers*, pp. 121-125.

## 3.

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 *o*, *h*, *k*, *ə*, *ø*, *z* were replaced by *g*, *h*, *k*, *a* (or *e*), *o* and *y* (or *j*). The scheme is unambiguous but does not distinguish in general use some of the Uighur characters.

## 4.

ا،ئا <sup>1</sup>	a	ر	r	گ	g	ا،ئۆ،ۆ	o [ō]
ا،ئە <sup>2</sup>	a, e <sup>2</sup> [ā]	ز	z	ڭ	ng	ا،ئۇ،ۇ	ū
ب	b	ژ	y, j <sup>3</sup>	ل	l	ه	h [h̄]
پ	p	س	s	م	m	ئە،ئەي <sup>1</sup>	e [ê]
ت	t	ش	x	ن	n	ئى،ئى(ءى)	i
ج	j	غ	g [ġ]	و	w, v	ي	y
چ	q	ف	f	ا،ئو،و	o		
خ	h	ق	k [k̂]	ا،ئۇ،ۇ	u		
د	d	ك	k				

<sup>1</sup> form used word-initially

<sup>2</sup> e is used where confusion of a word's meaning arises

<sup>3</sup> y is used at the beginning of a syllable and j at the end

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

## Urdu

## 1.

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<sup>39</sup>.

## 2.

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<sup>40</sup>. The resolutions III/12 (1977) and IV/17 (1982) recommended association, inter alia, with Pakistan, in carrying out further studies on the system.

## 3.

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 rest undifferentiated in the romanization which means that the system is not fully reversible.

## 4.

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

<sup>39</sup> *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.

<sup>40</sup> 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.



ا	a, i, u	د	d	ض	z	م	m
ب	b	ڈ	d	ط	t	ن	n
پ	p	ذ	z	ظ	z	و	v, v <sup>2</sup>
ت	t	ر	r	ع	' <sup>1</sup>	و (ہ)	h, h <sup>3</sup>
ٹ	t'	ڑ	r'	غ	gh	ہ	h <sup>4</sup>
ث	s	ز	z	ف	f	ی	y
ج	j	ذ	y	ق	q	ہ	e, ai
چ	ch	س	s	ک	k		
ح	h	ش	sh	گ	g		
خ	kh	ص	s	ل	l		

<sup>1</sup> if preceded by short *a*, it is transliterated 'ā, e.g. مَعْمَل M'āmūl

<sup>2</sup> when *v* is imperceptible, e.g. in a few words of Persian origin when preceded by ع (kh)

<sup>3</sup> word-finally after a short vowel

<sup>4</sup> marks aspiration of consonants

<sup>5</sup> used only word-finally

Note. The Arabic *h* has two distinct shapes in Urdu: *h* (ہ) and *h* (ہ). The latter denotes aspiration after consonants b, ch, d, ḍ, j, k, p, t, ṭ, ṭ.

#### Vowels

ا	a	ی	i	ای	ai <sup>1</sup>	ہ	(doubling)
آ	ā	ی	ī	او	o		
اُ	ā	و	u	او	au		
اَ	ā	و	ū		(omitted)		
آ	ā	ای	e <sup>1</sup>		- (hyphen)		

<sup>1</sup> word-final form

#### 5.

The Hunterian system of romanization marks the following sounds differently:

UN	Hunterian						
-ah (ا)	= a	gh	= gh	n	= n	v	= w, v <sup>2</sup>
ā	= ā, a <sup>1</sup>	h	= h	t	= r	y	= zh
ḍ	= d	ī	= l, i <sup>1</sup>	ṭ	= t		
ḍh	= dh	kh	= kh	th	= th		

<sup>1</sup> word-finally

<sup>2</sup> exceptionally, e.g. before *i*

## B. OTHER LANGUAGES

## Armenian

## 1.

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.

## 5.

The BGN/PCGN 1981 System gives the following romanization:

Ա ա	a	Ի ի	i	Յ յ	y	Տ տ	t
Բ բ	b	Լ լ	l	Ն ն	n	Ր ռ	r
Գ գ	g	Խ խ	kh	Շ շ	sh	Ց զ	ts'
Դ դ	d	Մ մ	ts	Ո ո	o, vo <sup>2</sup>	Ու ու	u
Ե ե	e, ye <sup>1</sup>	Կ կ	k	Չ չ	ch'	Փ փ	p'
Զ զ	z	Հ հ	h	Պ պ	p	Զ զ	k'
Է է	e	Զ ձ	dz	Ջ յ	j	Եւ էւ	ev, yev <sup>3</sup>
Ը ը	y	Ղ ղ	gh	Ռ ռ	rr	Օ օ	o
Թ թ	t'	Ճ ճ	ch	Ս ս	s	Ֆ ֆ	f
Ժ ժ	zh	Մ մ	m	Վ վ	v		

<sup>1</sup> ye is romanized initially and after the vowel characters ա, ե, է, ը, ի, ո, ու, and օ

<sup>2</sup> vo is romanized initially except in the word ույ which should be romanized ov

<sup>3</sup> yev is romanized initially, in isolation, and after the vowel characters ա, ե, է, ը, ի, ո, ու, and օ

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 brackets):

Arm.	ISO						
ե (e/ye)	e	ծ (ts)	ɟ	ո (o/vo)	o	ու (u)	ow
է (e)	ē	ձ (dz)	j	չ (ch')	č	օ (o)	ò
ը (y)	ē	ղ (gh)	ǧ	յ (j)	j		
ժ (zh)	ž	ճ (ch)	č	ռ (rr)	ɾ		
խ (kh)	x	շ (sh)	š	ց (ts')	c'		

## Belarusian

## 1.

No romanization systems for Belarusian (Byelorussian) 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 Belarusian was approved by the State Committee on Land Resources, Geodesy and Cartography, Republic of Belarus, on 20 March, 1998<sup>41</sup>. This scheme was also supported by the Y. Kolas Institute of Linguistics and the Republic Committee on Toponymy at the Belarusian Academy of Sciences. While the system is still based on GOST 1983, it takes more precisely into account the peculiarities of the

<sup>41</sup> The romanization table and notes on its application with the copy of the official act were distributed at the 3rd Session of the Eastern Europe, Northern and Central Asia Division of UNGEGN in June 1999 (Moscow).

Belarusian orthography. The system is reversible though there may exist some ambiguous consonant combinations.

## 5.

The national system of romanization (1998) is as follows.

А а	a	І і	i	С с	s		(omitted)
Б б	b	Ў ў	j	Т т	t	Ы ы	y
В в	v	К к	k	У у	u	Ь ь	'
Г г	h	Л л	l	Ў ў	ú	Э э	e
Д д	d	М м	m	Ф ф	f	Ю ю	ju, 'u <sup>1</sup>
Е е	je, 'e <sup>1</sup>	Н н	n	Х х	ch	Я я	ja, 'a <sup>1</sup>
Ё ё	jo, 'o <sup>1</sup>	О о	o	Ц ц	c		
Ж ж	ž	П п	p	Ч ч	č		
З з	z	Р р	r	Ш ш	š		

<sup>1</sup> the second variant is used after consonant letters

The BGN/PCGN 1979 System gives, as a single block, different Roman equivalents to the following Belarusian letters (the national equivalents are in brackets):

e (je/'e) ye	ђ (j) y	ц (c) ts	
ё (jo/'o) yo	ў (ü) w	ч (č) ch	ю (ju/'u) yu
ж (ž) zh	х (ch) kh	ш (š) sh	я (ja/'a) ya

## Burmese

## 1.

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.

## 5.

The BGN/PCGN 1970 System is given below.

## Consonant characters

ka	sa	qa	ta	pa	ya	
ka	sa	qa	ta	pa	ya	ha
ga	za	da	da	ba	la	la
ga	za	da	da	ba	wa	a
nga	nyā	na	na	ma	tha	

<sup>1</sup> at the end of a syllable also special character ꞑ may occur

<sup>2</sup> at the end of a syllable also special character ꞑꞑ may occur

**Independent vowel characters**

ဣ i	ဥ u	ဦ u	ဧ e
ဩ i	ဪ u	ဣ e	ဩ aw

**Abbreviated vowel characters (ဆ stands for any consonant character)**

ဆ an	ဆိ i	ဆု, ဆု u	ဆဲ e	ဆော aw	ဆေါ aw	ဆေ e
ဆာ, ဆါ a	ဆီ i	ဆု, ဆု u	ဆော aw	ဆော် aw	ဆို o	

**Other symbols (က stands for any consonant character)**

ကြ -y	ကျ -y	ကွ -w	ကု h-	ကိ <sup>1</sup> in-	ကိ <sup>2</sup>
-------	-------	-------	-------	---------------------	-----------------

Note. The hyphen in the first five character combinations stands for any consonant. With က k the resulting romanizations should read *ky, ky, kw, hk* and *ink*, respectively (the latter two do not exist in reality).

<sup>1</sup> adds *-in* to the end of the preceding syllable: ဘင်ဂလာအော် *Bin-gala Aw*

<sup>2</sup> end-of-syllable mark (marks the absence of the inherent vowel)

The following consonant combinations have different values:

ဠ, ဠ	cha
ရှ, ဝါ	sha

**Syllable endings (ဆ stands for any initial consonant character)**

ဆတ် et	ဆိတ် eik	ဆိန် ein	ဆွမ် un
ဆိတ် eik	ဆုတ် ók	ဆုန် ón	ဆမ် wun
ဆုတ် ók	ဆုတ် ut	ဆုန် un	ဆတ် è
ဆိုတ် aik	ဆတ် wut	ဆန် wun	ဆတ် an
ဆောတ် auk	ဆော် it	ဆက် an	ဆိုလ် aing
ဆတ် et	ဆင့် at	ဆိက် ein	ဆတ် at
ဆင် in	ဆင့် wut	ဆက် ón	ဆိတ် eik
ဆိင် ein	ဆင့် wut	ဆိုက် aing	ဆင့် in
ဆိုင် aing	ဆင် at	ဆတ် at	ဆတ် an
ဆောင် aung	ဆိင် eik	ဆိပ် eik	ဆိ ein
ဆပ် it	ဆင် ók	ဆပ် ók	ဆုတ် ón
ဆုပ် ók	ဆင် at	ဆုတ် ut	ဆုတ် un
ဆပ် it	ဆပ် at	ဝိတ် wut	ဝတ် wun
ဆိပ် eik	ဆိပ် eik	ဆပ် an	ဆတ် at
ဆည် i, in, e <sup>1</sup>	ဆပ် at	ဆိပ် ein	
ဆတ် at	ဆန် an	ဆပ် ón	

<sup>1</sup> choice depends on 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:  $\text{မာမာ}$  *madama*.
3.  $\text{အ}$  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:  $\text{အာ}$  *ma-u*,  $\text{အိဝ်ဒိဒိ}$  *Thibin-aing*.
4. The independent vowel characters should be romanized with a hyphen at the beginning of medial and final syllables:  $\text{အိဝ်}$  *kye-e*.
5. In consonant clusters the first component is romanized according to the main table, not according to the table of syllable endings:  $\text{မိဗ္ဗ}$  *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:  $\text{အင်ဂုတ်}$  *in-gut* (but  $\text{အင်ဂန်}$  *shwengan*).
7. Tone marks, and  $\text{း}$  should not be romanized.

## Dzongkha

### 1.

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.

### 5.

The official romanization (1994) is as follows<sup>2</sup>.

#### Consonant characters:

ཀ ka	ཅ ca	ཉ ta	པ pa	ཌ tsa	འ zh'a (zha)	ལ la ('la)
ཁ kha	ཆ cha	ཐ tha	ཕ pha	ཎ tsha	མ z'a (za)	ཤ sha
ཀ g'a (ga)	ཇ j'a (ja)	ཉ d'a (da)	པ b'a (ba)	ཎ dza	འ a	ལ sa
ང nga ('nga)	ཉ nya ('nya)	འ na ('na)	འ ma ('ma)	འ wa	འ ya	ཉ ha
					ར ra	ལ 'a

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

<sup>2</sup> Adapted from *Dzongkha* by George van Driem with the collaboration of Karma Tshering of Gaseló. Research School CNWS, School of Asian, African, and Amerindian Studies, Leiden, The Netherlands 1998.

**Consonant combinations at the beginning of a syllable (the list is not complete)**

ཀ	k, c <sup>1</sup>	ཇ	j' (j), bj' (bj)	ཏ	tr	ཌ	dr
ཁ	kh, ch <sup>1</sup>	ཉ	ny	ཐ	thr	ཌྱ	hra
ག	g' (g), j' (j) <sup>1</sup>	ཊ	tr	ཎ	dr' (dr)	ཌྱ	wa
ང	c, pc	ཏྱ	thr	ཏྱ	tr	ཌྱ	d
ཅ	ch, pch	ཏྱ	dr' (dr)	ཏྱ	thr	ཌྱ	lh

<sup>1</sup> palatal variants *j, j', c, ch* are generally used before *a (a), o (o)* and *u (u)*

Note. Where there are multiple romanizations, the romanization in brackets is to be generally used if the character is accompanied by a prefixed letter or a superadded consonant.

**Vowel characters (ཀ stands for any consonant character)**

ཀ a	ཀི i	ཀུ u	ཀེ e	ཀོ o
-----	------	------	------	------

**Other symbols**

· syllable separation character

**Syllable endings (the list is not complete)**

ཨ	a	ཨིགས	i, i, ik	ཨུགས	um	ཨེགས	em
ཨག	á, a, ak	ཨི	i, i	ཨུ	ú	ཨེ	é
ཨགས	á, a, ak	ཨིས	i, i	ཨུས	ü	ཨེས	en
ཨང	a, ang	ཨིང	i, ing	ཨུང	ü	ཨེ	o, ó
ཨངས	a, ang	ཨིངས	i, ing	ཨུངས	ü	ཨེག	ó, ok
ཨཔ	ap	ཨིཔ	ip	ཨུཔ	ü	ཨེགས	ó, ok
ཨཔས	ap	ཨིཔས	ip	ཨུཔས	ü	ཨི	o, ong
ཨམ	am	ཨིམ	im	ཨུམ	e, i	ཨིས	o, ong
ཨམས	am	ཨིམས	im	ཨེ	é	ཨི	op
ཨའ	á	ཨི	i	ཨུའི	iu	ཨིའི	op
ཨའི	ai, é	ཨི	in	ཨེཀ	é, ek	ཨི	om
ཨའུ	á	ཨུ	u, ú	ཨེཀས	é, ek	ཨིཀས	om
ཨའུ	é, á	ཨུཀ	ú, u, uk	ཨི	é	ཨི	ó
ཨའུ	á, á	ཨུཀས	ú, u, uk	ཨི	é	ཨི	ó
ཨའུ	en	ཨིཀ	u, ung	ཨི	c, eng	ཨི	ó
ཨི	i	ཨིཀས	u, ung	ཨིཀས	c, eng	ཨི	ó, ó
ཨིའི	t, i	ཨིཔ	up	ཨི	cp	ཨི	ón
ཨིའི	tu	ཨིཔས	up	ཨིཔས	cp		
ཨིག	t, i, ik	ཨིམ	um	ཨི	em		

Note. 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 པ (p), མ (m) or འ (w) at the end, e.g. ཨགཔ ap, ཨིགཔ ip, ཨིགམ op, ཨའཔ áp, ཨིའཔ íp, ཨིའའི ép, ཨའམ ám, ཨུའི ú, etc. These are not covered here.

## Georgian

### 1.

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. No information has been received on whether there exist any national systems of romanization in Georgia.

### 5.

The BGN/PCGN 1981 System gives the following romanization:

ა	a	კ	k	ც	t	ძ	dz
ბ	b	ლ	l	უ	u	ჭ	ts
გ	g	მ	m	პ	p'	ჩ	ch
დ	d	ნ	n	კ	k'	ხ	kh
ე	e	ო	o	გ	gh	ჯ	j
ვ	v	პ	p	ყ	q	ჰ	h
ზ	z	ჟ	zh	შ	sh		
რ	r'	ს	r	ჩ	ch'		
ი	i	ს	s	ც	ts'		

The transliteration standard ISO 9984:1996 gives, as a single block, different Roman equivalents to the following Georgian letters (the BGN/PCGN equivalents are in brackets):

ა (zh)	z	ჩ (ch')	č'	ც (ts)	c	ჯ (j)	j
ბ (gh)	g	ც (ts')	c'	ჩ (ch)	č		
შ (sh)	š	ძ (dz)	j	ხ (kh)	x		

## Japanese

### 1.

No romanization systems for Japanese have 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.

### 5.

Below is a list of Japanese syllables with Roman equivalents. In the table line 1 gives *katakana*, line 2 *hiragana* and line 3 romanization in *Kunrei-siki*. In line 4 romanization in the Modified Hepburn System in italics is given, if different from the *Kunrei-siki*.

Combined forms:

1	ア	イ	ウ	エ	オ	カ
2	あ	い	う	え	お	か
3	a	i	u	e	o	ka
4						<i>ka</i>

1	カ	キ	ク	ケ	コ	キ+	キユ	キユウ	キョ	キョウ	コウ
2	か	き	く	け	こ	き*	きゆ	きゆう	ぎょ	ぎょう	こう
3	ka	ki	ku	ke	ko	kya	kyu	kyū	kyo	kyō	kō
4				ke (ga)				kyū	kyō	kyō	kō
1	サ	シ	ス	セ	ソ	シ+	シユ	シユウ	ショ	ショウ	ソウ
2	さ	し	す	せ	そ	し*	しゆ	しゆう	じょ	じょう	そう
3	sa	si	su	se	so	sya <sup>4</sup>	syu	syū	syo <sup>4</sup>	syō	sō
4		shi				sha	shu	shū	sho	shō	sō
1	タ	チ	ツ	テ	ト	チ+	チユ	チユウ	チョ	チョウ	トウ
2	た	ち	つ	て	と	ち*	ちゆ	ちゆう	ちょ	ちょう	とう
3	ta	ti	tu <sup>1</sup>	te	to	tya	tyu	tyū	tyo	tyō	tō
4		chi	tsu			cha	chu	chū	cho	chō	tō
1	ナ	ニ	ヌ	ネ	ノ	ニ+	ニユ	ニユウ	ニョ	ニョウ	ノウ
2	な	に	ぬ	ね	の	に*	にゆ	にゆう	にょ	にょう	のう
3	na	ni	nu	ne	no	nya	nyu	nyū	nyo	nyō	nō
4								nyū	nyō	nyō	nō
1	ハ	ヒ	フ	ヘ	ホ	ヒ+	ヒユ	ヒユウ	ヒョ	ヒョウ	ホウ
2	は	ひ	ふ	へ	ほ	ひ*	ひゆ	ひゆう	ひょ	ひょう	ほう
3	ha	hi	hu	he	ho	hya	hyu	hyū	hyo	hyō	hō
4	ha (wa)		fu	he (e)				hyū		hyō	hō (ō)
1	マ	ミ	ム	メ	モ	ミ+	ミユ	ミユウ	ミョ	ミョウ	モウ
2	ま	み	む	め	も	み*	みゆ	みゆう	みょ	みょう	もう
3	ma	mi	mu	me	mo	mya	myu	myū	myo	myō	mō
4								myū	myō	myō	mō
1	ヤ		ユ		ヨ						ヨウ
2	や		ゆ		よ						よう
3	ya		yu		yo						yō
4											yō
1	ラ	リ	ル	レ	ロ	リ+	リュ	リュウ	リョ	リョウ	ロウ
2	ら	り	る	れ	ろ	り*	りゆ	りゆう	りょ	りょう	ろう
3	ra	ri	ru	re	ro	rya	ryu	ryū	ryo	ryō	rō
4								ryū		ryō	rō
1	ワ				ヲ						
2	わ				を						
3	wa				o						
4											
1	ン										
2	ん										
3	n <sup>2</sup>										
4	n, n', m										
1	ガ	ギ	グ	ゲ	ゴ	ギ+	ギユ	ギユウ	ギョ	ギョウ	ゴウ
2	が	ぎ	ぐ	げ	ご	ぎ*	ぎゆ	ぎゆう	ぎょ	ぎょう	ごう
3	ga	gi	gu	ge	go	gya	gyu	gyū	gyo	gyō	gō
4								gyū		gyō	gō
1	ザ	ジ	ズ	ゼ	ゾ	ジ+	ジユ	ジユウ	ジョ	ジョウ	ゾウ
2	ざ	じ	ず	ぜ	ぞ	じ*	じゆ	じゆう	じょ	じょう	ぞう
3	za	zi	zu	ze	zo	zya	zyu	zyū	zyo	zyō	zō
4		ji				ja	ju	jū	jo	jō	zō
1	ダ			デ	ド						ドウ
2	だ			で	ど						どう
3	da			de	do						dō
4											dō



1	バ	ビ	ブ	ベ	ボ	ビヤ	ビユ	ビユウ	ビョ	ビョウ	ボウ
2	ば	び	ぶ	べ	ぼ	びや	びゆ	びゆう	びょ	びょう	ぼう
3	ba	bi	bu	be	bo	bya	byu	byū	byō	byō	bō
4								byū		byō	bō
1	パ	ピ	プ	ペ	ポ	ピヤ	ピユ	ピユウ	ピョ	ピョウ	ポウ
2	ぱ	ぴ	ぷ	ぺ	ぽ	ぴや	ぴゆ	ぴゆう	ぴょ	ぴょう	ぽう
3	pa	pi	pu	pe	po	pya	pyu	pyū	pyō	pyō	pō
4								pyū		pyō	pō

Notes

1. A small-script *tu/su* form (ツ / ヅ) is inserted between kana symbols to indicate a double consonant (*kk, ss, tt, pp* in Kunrei-siki; *kk, ss, ssh, tt, ts, tch, pp* in Hepburn).
2. ん / 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 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 (ˆ) over a vowel; in Hepburn they are indicated by a macron (¯).
4. The romanization of syllables シヤ / シャ and ショ / ショ is given as *sha* and *sho*, respectively, in the Gazetteer of Japan, published by the Geographical Survey Institute and Hydrographic Department, Maritime Safety Agency for the Government of Japan in 1997, while in earlier sources these syllables are romanized as *syā* and *syō*.

## Kazakh

1.

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. No information has been received on whether there exist any national systems of romanization in Kazakhstan. There have been reports suggesting that the Kazakh language will move to use the Roman alphabet in the future but until now no specific information on it has been made available.

5.

The BGN/PCGN 1979 System gives the following romanization:

А а	a	И и	i	Р р	r	Ш ш	sh
Ә ә	ä	Ү ү	y	С с	s	Щ щ	shch
Б б	b	Қ қ	k	Т т	t	Ъ ъ	"
В в	v	К к	q	У у	u	Ы ы	y
Г г	g	Л л	l	Ұ ұ	u	І і	i
Ғ ғ	gh	М м	m	Ү ү	ü	Ь ь	'
Д д	d	Н н	n	Ф ф	f	Ә ә	e
Е е	e	Ң ң	ng	Х х	kh	Ю ю	yu
Ә ә	yo	О о	o	Һ һ	h	Я я	ya
Ж ж	zh	Ө ө	ö	Ц ц	ts		
З з	z	П п	p	Ч ч	ch		

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

ә (ä)	ä	қ (q)	q	х (kh)	h	і (i)	i
ғ (gh)	g	ң (ng)	ŋ	һ (h)	h	ә (e)	e
ё (yo)	e	ө (ö)	ö	ц (ts)	c	ю (yu)	ü
ж (zh)	z	ү (ü)	u	ч (ch)	ç	я (ya)	ä
и (i)	i	у (u)	u	ш (sh)	š		
й (y)	j	ұ (ü)	ü	щ (shch)	š		

## Kirghiz

### 1.

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. There have been reports suggesting that the Kirghiz language will move to use the Roman alphabet in the future but until now no specific information on it has been made available.

### 5.

The BGN/PCGN 1979 System gives the following romanization:

А а	а	И и	и	П п	р	Ч ч	ch
Б б	б	Й й	у	Р р	г	Ш ш	sh
В в	в	К к	к	С с	с	Щ щ	shch
Г г	г	Л л	л	Т т	т	Ъ ъ	"
Д д	д	М м	м	У у	у	Ы ы	у
Е е	е	Н н	н	Ү ү	ü	Ь ь	'
Ё ё	yo	Ң ң	ng	Ф ф	f	Э э	е
Ж ж	ж	О о	о	Х х	kh	Ю ю	yu
З з	з	Ө ө	ö	Ц ц	ts	Я я	ya

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

ё (yo)	ë	ө (ö)	ö	ч (ch)	č	ю (yu)	ü
ж (j)	ž	ү (ü)	ü	ш (sh)	š	я (ya)	ä
й (y)	j	х (kh)	h	щ (shch)	š		
н (ng)	ñ	ц (ts)	c	э (e)	è		

## Korean

### 1.

No romanization systems for Korean have 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<sup>43</sup>.

In the Republic of Korea the Ministry of Education had adopted on 13 January, 1984, guidelines for the romanization of Korean, presented to the Fifth UN Conference on the Standardization of Geographical Names in 1987<sup>44</sup>. The Conference recommended that the guidelines be referred to UNGEGN for consideration (resolution V/20).

Apart from national systems the most widely used international system is that of McCune-Reischauer (1939) which

<sup>43</sup> *Guidelines for Romanization of Korean* (Democratic People's Republic of Korea). UNGEGN, 17th Session. New York, 13-24 June 1994, WP.No. 57.

<sup>44</sup> Fifth United Nations Conference on the Standardization of Geographical Names. Montreal, 18-31 August 1987. Document E/CONF.79/INF/68.

has also been adopted by the BGN and the PCGN<sup>45</sup>.

A transliteration system that was provisionally agreed upon by the experts of the Democratic People's Republic of Korea and the Republic of Korea, is the standard ISO TR 11941:1996 (a provisional technical report, not enforced as a standard)<sup>46</sup>.

In linguistics also the so-called Yale system of romanization is widely accepted<sup>47</sup>.

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 판문-점.

5.

The transliteration of Korean script into Latin characters in the standard ISO TR 11941:1996 is as follows.

ㄱ	k/g	ㅋ	ch/c	ㅣ	eo	ㅛ	yo
ㅋ	kh/k	ㆁ	cc/jj	ㅜ	o	ㅠ	yu
ㆁ	kk/gg	ㅅ	s	ㅡ	u	ㅣ	yae
ㄷ	t/d	ㅆ	ss	ㅡ	eu	ㅣ	ye
ㅌ	th/t	ㅎ	h	ㅣ	i	ㅜ	wa
ㄸ	tt/dd	ㅇ	' <sup>1</sup> , - <sup>1</sup> , ng <sup>2</sup>	ㅣ	ae	ㅜ	weo
ㅍ	p/b	ㄴ	n	ㅣ	e	ㅜ	wi
ㅑ	ph/p	ㄹ	l, r <sup>3</sup>	ㅜ	oe	ㅜ	wae
ㅓ	pp/bb	ㅁ	m	ㅣ	ya	ㅜ	we
ㅕ	c/j	ㅂ	a	ㅣ	yeo	ㅜ	yi

<sup>1</sup> word-initially.

<sup>2</sup> at the end of a syllable

<sup>3</sup> at the beginning of a syllable

Notes

1. Where there are pairs of romanizations, the first corresponds to Method I and the second to Method II.
2. To avoid ambiguity an apostrophe is used to denote the beginning of the next syllable in a polysyllabic word.

By Method I it is needed in the following cases:

- A. if the consonant <sup>0</sup> is at the initial position of a non-initial syllable of a polysyllabic word: ㅋㅌ ʼkkoč'i, ㄱㅌ ʼkang'e, ㄴㅌ ʼanc'ara;
- B. if the five double consonants appear at the initial positions of the non-initial syllables of polysyllabic words: ㄱㅌ ʼgkka, ㅌㅌ ʼhyin'tteok;
- C. if the four aspirated consonants appear at the initial position of the non-initial syllables of polysyllabic words: ㅍㅍ ʼyu khwaehata;

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

- A. if the consonant <sup>0</sup> is at the initial position of a non-initial syllable of a polysyllabic word: ㄱㅌ ʼggoc'i, ㄱㅌ ʼgang'e, ㄴㅌ ʼanj'ara;
- B. if the five double consonants appear at the initial positions of the non-initial syllables of polysyllabic words: ㄱㅌ ʼg'gga, ㅌㅌ ʼhyin'ddeog.

<sup>45</sup> *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.

<sup>46</sup> The summary of it was published at the Internet site <http://www.phon.ucl.ac.uk/home/kskim/11941.html>.

<sup>47</sup> 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.

The McCune-Reischauer 1939 System is as follows.

ㄱ	k	ㄷ	t'	ㅣ	yō	ㅛ	oe
ㄴ	n	ㄸ	p'	ㅜ	o	ㅜ	wi
ㄷ	t	ㅌ	h	ㅠ	yo	ㅜ	ūi
ㄹ	r,n,l	ㅍ	kk	ㅡ	u	ㅜ	wa
ㅁ	m	ㅊ	tt	ㅝ	yu	ㅜ	wō
ㅂ	p	ㅍ	pp	ㅡ	ū	ㅜ	wae
ㅅ	s	ㅍ	ss	ㅣ	i	ㅜ	we
ㅇ	ng, (¹)	ㅊ	tch	ㅣ	ae		
ㅈ	ch	ㅊ	a	ㅣ	yae		
ㅊ	ch'	ㅊ	ya	ㅣ	e		
ㅋ	k'	ㅊ	ō	ㅣ	ye		

¹ not romanized syllable-initially

Note. There are complicated rules based on pronunciation that determine the conversion of Korean syllables into Roman. Most changes concern consonants which are often assimilated when used in combinations. For reasons of economy these rules are not reproduced in this report.

The national system of the Democratic People's Republic of Korea largely coincides with Method I of the provisional standard ISO TR 11941:1996. In addition, rules on reflecting the assimilation of consonants and other phonetic changes have been supplied.

The national system of the Republic of Korea largely coincides with the McCune-Reischauer system. Differences lie in the notes on its application and in the conversion of some syllables, e.g. the syllable ㅅ is romanized *shi*, not *si*: ㅅ *Shinan*.

## Laotian

### 1.

No romanization systems for Laotian have 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.

### 5.

The BGN/PCGN 1966 System gives the following romanization<sup>48</sup>:

Consonant characters

ກ k (k)	ຈ ch	ດ d (t)	ບ b (p)	ຢ y	ຫ h
ຂ kh	ສ s	ຕ t	ປ p	ຮ r	ອ (ʰ)
ຄ kh	ຊ x	ຖ th	ຜ ph	ລ l²	ຮ h
ງ ng (ng)	ຍ gn (y)¹	ຸ th	ຝ f	ວ v, o³	
		ນ n (n)	ຸ ph		
			ຸ f		
			ມ m		

¹ variation: ງ (used at the end of a syllable: ຂາງ *khay*; after ຫ: ຫາງ *gnap*; and between consonants: ສາງ *siang*)

<sup>48</sup> 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.

<sup>2</sup> variation: ຫຼາມ (*lam*)

<sup>3</sup> see note 4

<sup>4</sup> see note 3

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

Vocalic nuclei (\* stands for any consonant character)

ອ*	a	ເ*	é	ເື*	eu	ເືວະ	oua
ອ້*	a	ແອ*	è	ເື້*	eu	ເືວ້*	oua
ອາ	a	ແອ້*	è	ເື່*	eu	ເືວ່*	oua
ອີ	i	ແອ*	è	ເື໊*	eu	ເືວ໊*	oua
ອີ້	i	ໄອະ	ó	ເື້ວະ	ia	ໄອ*	ai
ອື	u	ອີ*	ó	ເື້ວ້*	ia	ໄອ້*	ai
ອື້	u	ໄອ*	ó	ເື້ວ່*	ia	ເື່າ	ao
ອູ	ou	ເອະ	o	ເືວ໊*	ia	ອາ	am
ອູ້	ou	ອັອ*	o	ເື້ອະ	ua		
ເອະ	é	ອີ	o	ເື້ອ໊*	ua		
ເອ້*	é	ອັອ*	o	ເື໊*	ua		

Reduplicating mark:

ງ (e.g. ມອອງ *xəoxəo*)

Notes

1. Tone marks ະ, ື, ື້, ື່ should not be romanized.
2. At the beginning of a syllable, the character ທ unaccompanied by a vowel or tone mark and occurring immediately before ອ *gn*, ນ *n*, ມ *m*, ຣ *r*, ລ *l*, or ວ *v* should generally not be romanized: ທລວງພຣະບາງ *Louāngphrabang*. Often ligatures ພ, ທຸ, ທູ are used in such cases.
3. At the beginning of a word ອ should not be romanized: ອີກ *ik*. At the beginning of a syllable within a word the letter should be romanized by a hyphen: ຊະເອນ *xa-ém*, ບັນທຶກອີ *Ban Kəng-i*.
4. The character ອ at the beginning of a syllable should be romanized *v*: ທັນວາ *thanva*. As the second character of a combination, ອ should be romanized *o*. At the end of a syllable the character is romanized *o* generally (ນາວ *nao*, ມາວ *mao*, ດາວ *diao*) but after *i* it is romanized *ou* (ຄົວ *khiov*, ຕົວ *tiov*).
5. In consonant combinations where the second character is ຣ, ຣ, ລ or ວ, it should be romanized character by character, even if the second character may represent no pronounced sound: ບັນທຶກພຣະໄຊ *Ban Vaphraxai*, ຖືສຳນຽນ *Khūtsān*, ສວຍ *soy*, ມອວງ *khoəng*.
6. In some words of foreign origin the variant ຣ *r* occurs: ບີຣ໌ *beur*.

## Maldivian

### 1.

No romanization systems for Maldivian (Divehi) have 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, this was in 1988 also approved by the BGN and the PCGN.

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

### 5.

The official romanization is as follows.

#### Consonants

ހ	h	ކ	k	ތ	th	ޅ	z
ޞ	sh	ލ	( <sup>2</sup> )	ލ	l	ޑ	t
ނ	n <sup>1</sup>	ވ	v	ގ	g	ޕ	y
ރ	r	މ	m	ޅ	gn	ފ	p
ބ	b	ފ	f	ސ	s	ޖ	j
ދ	dh	ޑ	dh	ޔ	d	ޝ	ch

<sup>1</sup> also used without any vowel or auxiliary sign: ބ ޗ ޝ bandos

<sup>2</sup> not romanized (see note 1)

#### Vowels (\* stands for any consonant character)

އ	a	ޯ	ey	އ	o	ޯ	oo
ޯ	aa	އ	i	ޯ	oa		
އ	e	ޯ	ee	އ	u		

#### Auxiliary symbol

• (marks absence of the vowel)

#### Notes

- The character ލ is not romanized. If it bears a vowel character, that vowel character alone is romanized, e.g. ލ ޯ ޯ ޯ adhaifi. If it does not bear a vowel sign ( ލ ) the romanized value of the following consonant character is doubled, e.g. ލ ލ ލ ލ ލ ލ ލ ލ behettumash.
- In words of Arabic origin also consonants with the following diacritics are used (romanization values of Arabic are in brackets): ڪ (kh), ھ (h), ڃ (gh), ڀ (ʔ), ڦ (w), ڦ (dh), ڦ (th), ڦ (t), ڦ (q), ڦ (sh), ڦ (d), ڦ (s).

## Mongolian (Cyrillic)

## 1.

No romanization systems for Mongolian Cyrillic have 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.

## 5.

The BGN/PCGN 1964 System gives the following romanization:

А а	a	И и	i	Р р	r	Ш ш	sh
Б б	b <sup>1</sup>	Й й	y <sup>2</sup>	С с	s	Щ щ	shch <sup>4</sup>
В в	v <sup>2</sup>	К к	k <sup>4</sup>	Т т	t	Ъ ъ	' <sup>5</sup>
Г г	g	Л л	l	У у	u	Ы ы	f <sup>6</sup>
Д д	d	М м	m	Ү ү	ü	Ь ь	l
Е е	yö	Н н	n	Ф ф	f <sup>4</sup>	Э э	e
Ё ё	yo	О о	o	Х х	h	Ю ю	yu, yü <sup>7</sup>
Ж ж	j	Ө ө	ö	Ц ц	ts	Я я	ya
З з	dz	П п	p <sup>4</sup>	Ч ч	ch		

<sup>1</sup> the character б occurs initially and after x, m and ö

<sup>2</sup> the character s occurs medially in Mongolian words and initially in foreign words

<sup>3</sup> the character й occurs only as the second element of diphthongs

<sup>4</sup> the characters x, n, ф and ш occur only in foreign words

<sup>5</sup> the character ъ occurs only in Russian words

<sup>6</sup> the character ы occurs only in suffixes to words containing the vowel character a, o or y

<sup>7</sup> 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 l, e, ö, or ü

## Pashto

## 1.

No romanization systems for Pashto (Afghan) have 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.

## 5.

The BGN/PCGN 1968 System gives the following romanization:

ا	-	ب	b	س	s	ق	q
ب	b	ك	kh	ش	sh	ك	k
پ	p	د	d	س	sh	گ (گ)	g
ت	t	ذ	z	ص	s	ل	l
ث	t	ر	r	ض	z	م	m
ج	j	ز	z	ط	t	ن	n
چ	ch	ژ	zh	ظ	z	و	w
س	s	ح	ch	ع	e	ه	h
ش	sh	خ	gh	غ	gh	ي (ي)	y
ص	s	ج	j	ف	f		

Vowels (\* stands for any consonant)

• a	اى ay	اے ey	آ ā
آ ā	ای ey	• u	•• ' (omitted)
• i	ئ ey	• u	• (omitted)
آ ī	ی ē	• ow	• (doubling)
آ ī	• ē	• uy	

## Sinhalese

### 1.

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<sup>49</sup>.

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.

### 5.

To the table prepared by D. N. Sharma character variations and the list of ligatures have been added.

#### Independent vowel characters

අ a	ඉ i	ආ ා	ඒ ඒ <sup>1</sup> ai <sup>1</sup>
ආ ā	ඊ ī	ඃ ා ා	ඔ ඔ
උ උ <sup>1</sup>	ඌ u	ඒ ඒ	ඹ ඹ
උ උ <sup>1</sup>	ඍ ū	ඔ ඔ	ඹ ඹ au

<sup>1</sup> the character අඒ was romanized as *āi* and the character අඔ as *ai* in Sharma's table, but there was no romanization for the character ඒ 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.

#### Abbreviated vowel characters (\* stands for any consonant character)

• a	ඞ i	• ා ා	ඞ• ai
•) ā	ඞ ī	• ා ා ා	ඞ•) ඔ
•උ උ <sup>1</sup>	ඞ u <sup>2</sup>	ඞ• ඒ	ඞ•) ඹ
•උ උ <sup>1</sup>	ඞ ū <sup>3</sup>	ඞ• ඔ	ඞ•) au

<sup>1</sup> exceptions: ඞ ා ා, ඞ ා ා

<sup>2</sup> exceptions: ඞ ku, ඞ gu, ඞ tu, ඞ bhu, ඞ ru, ඞ shu

<sup>3</sup> exceptions: ඞ kū, ඞ gū, ඞ tū, ඞ bhū, ඞ rū, ඞ shū

<sup>49</sup> *Second United Nations Conference on the Standardization of Geographical Names*. London, 10-31 May 1972. Vol. II. Technical papers, pp. 151-152.



**Consonant characters:**

ක ka	ච cha*	ච්චා චා	ත ta	ප pa	ය ya	ශ sha*
ඛ kha*	ඡ chha*	ච්චා චා	ථ tha*	ඵ pha*	ර ra	ඝ gha*
ග ga	ඣ ja	ධ da	ඳ da	ඞ ba	ල la	ස sa
ඝ gha*	ඞ jha*	ධ්ඞා ඞා	ධ්ධා ඞා	ඞ්ඞා ඞා	ව va	හ ha
ඞ ඞා*	ඞ් ඞා*	ඞ් ඞා	ඞ na	ම ma	(ඞ fa)	ඞ ja

\* consonants marked with an asterisk are used only in Sanskrit and Pali loanwords

**Other dependent characters**

◌̆ ◌̇ ◌̈ ◌̉ ◌̊ ◌̋ ◌̌ ◌̍ ◌̎ ◌̏ ◌̐ ◌̑ ◌̒ ◌̓ ◌̔ ◌̕ ◌̖ ◌̗ ◌̘ ◌̙ ◌̚ ◌̛ ◌̜ ◌̝ ◌̞ ◌̟ ◌̠ ◌̡ ◌̢ ◌̣ ◌̤ ◌̥ ◌̦ ◌̧ ◌̨ ◌̩ ◌̪ ◌̫ ◌̬ ◌̭ ◌̮ ◌̯ ◌̰ ◌̱ ◌̲ ◌̳ ◌̴ ◌̵ ◌̶ ◌̷ ◌̸ ◌̹ ◌̺ ◌̻ ◌̼ ◌̽ ◌̾ ◌̿ ◌̿̆ ◌̿̇ ◌̿̈ ◌̿̉ ◌̿̊ ◌̿̋ ◌̿̌ ◌̿̍ ◌̿̎ ◌̿̏ ◌̿̐ ◌̿̑ ◌̿̒ ◌̿̓ ◌̿̔ ◌̿̕ ◌̖̿ ◌̗̿ ◌̘̿ ◌̙̿ ◌̿̚ ◌̛̿ ◌̜̿ ◌̝̿ ◌̞̿ ◌̟̿ ◌̠̿ ◌̡̿ ◌̢̿ ◌̣̿ ◌̤̿ ◌̥̿ ◌̦̿ ◌̧̿ ◌̨̿ ◌̩̿ ◌̪̿ ◌̫̿ ◌̬̿ ◌̭̿ ◌̮̿ ◌̯̿ ◌̰̿ ◌̱̿ ◌̲̿ ◌̳̿ ◌̴̿ ◌̵̿ ◌̶̿ ◌̷̿ ◌̸̿ ◌̹̿ ◌̺̿ ◌̻̿ ◌̼̿ ◌̿̽ ◌̿̾ ◌̿̿

◌̆ marks absence of the inherent vowel. Special forms: ඛ kh, ඞ n, ඞ ch, ඞ r, ඞ d, ඞ dh, ඞ b, ඞ m, ඞ v.

**Ligatures****Examples**

◌̆ r: ඞ rka, ඞ rta, ඞ rma, ඞ rva, ඞ rsha

◌̆ r: ඞ kra, ඞ gra, ඞ ghra, ඞ jra, ඞ tra, ඞ dra, ඞ tra, ඞ dra, ඞ dhra, ඞ pra, ඞ bra, ඞ bhra, ඞ sra

◌̆ y: ඞ kya, ඞ tya, ඞ vya

Prenasalized sounds: ඞ nga, ඞ nja, ඞ nda, ඞ nda, ඞ mba

Other ligatures: ඞ ksha, ඞ jna, ඞ dva, ඞ nda

**Tajik**

1.

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.

5.

The BGN/PCGN 1979 System gives the following romanization:

А а	a	Й й	y	У у	u	Я я	ya
Б б	b	К к	k	Ф ф	f	Ғ ғ	gh
В в	v	Л л	l	Х х	kh	Ҳ ҳ	h
Г г	g	М м	m	Ц ц	ts	Қ қ	q
Д д	d	Н н	n	Ч ч	ch	Ү ү	ü
Е е	e	О о	o	Ш ш	sh	Ҷ ҷ	j
Ё ё	yo	П п	p	Ъ ъ	'	Ҷ ҷ	j
Ж ж	zh	Р р	r	Ь ь	'		
З з	z	С с	s	Э э	e		
И и	i	Т т	t	Ю ю	yu		

Since 1998 the alphabet has been officially rearranged. The changes principally involve the alphabetical resorting of six characters (Ғ, ғ, Қ, қ, Ҷ, ҷ) and the abolition of four characters (У, у, Ъ, ъ).

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

ё (yo)	ë	Ҷ (ch)	č	ю (yu)	ü	Ҷ (ü)	ü
ж (zh)	ž	ш (sh)	š	я (ya)	ä	х (h)	h
й (y)	j	ъ (')	'	ғ (gh)	ğ	ҷ (j)	ç
х (kh)	h	ь ('')	'	и (i)	ı		
ц (ts)	c	э (e)	e	қ (q)	q		

## Tigrinya

### 1.

No romanization systems for Tigrinya have 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.

The BGN and the PCGN agreed in 1994 to use for Tigrinya the same romanization system as for Amharic, because these two languages share the same syllabic alphabet.

### 5.

Tigrinya uses the Ethiopic syllabic alphabet which is the same as used for Amharic (see the section on the romanization of Amharic). In addition at least the following consonant character is used (given here in the BGN/PCGN system for vowels):

ቆ	ቆ	ቆ	ቆ	ቆ	ቆ	ቆ
k'e	k'u	k'i	k'a	k'e	k'i	k'o

The consonant character may be romanized k' for documentation purposes. There may be other modified characters of the Ethiopian alphabet in use (e.g. ጒ).

## Turkmen

### 1.

No romanization systems for Turkmen 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. A new Roman alphabet was officially decreed for the Turkmen language, to take effect in 1996. This is now in parallel use with the Cyrillic alphabet although there are some variations of letters.

### 5.

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

А а	И и	Р р	Ъ ъ
а	и	р	(omitted)
Б б	Й й	С с	Ы ы
б	й (ʏ)	с	у
В в	К к	Т т	Ь ь
в	к	т	(omitted)
Г г	Л л	У у	Э э
г	л	у	е
Д д	М м	Ү ү	Ә ә
д	м	ü	ä
Е е	Н н	Ф ф	Ю ю
е, ye <sup>1</sup>	н	ф	ју
Ё ё	Ң ң	Х х	Я я
yo	ñ	х	ya
Ж ж	О о	Ц ц	
ʒ (ʒ)	о	ts	
Ж ж	Ө ө	Ч ч	
ʒ	ö	ç	
З з	П п	Ш ш	
z	р	ç (\$)²	

Note. Where the capital letter unconventionally differs from the small letter, the capital letter is given in brackets.

<sup>1</sup> the Cyrillic letter е is pronounced [je] (in the new orthography ye) word-initially and after vowels, й, ъ and ъ.

<sup>2</sup> instead of \$ ç also a variant Ş ş is used

The BGN/PCGN 1979 System gives, as a single block, different Roman equivalents to the following Turkmen letters (the official Roman equivalents are in brackets):

е (e/ʏe) e, ye <sup>1</sup>	ӥ (ʏ) y	ш (ç) sh	ю (ʏu) yu
ё (ʏo) yo	ӥ (ñ) ng	ъ	я (ʏa) ya
ж (ʒ) zh	ч (ç) ch	ь	

<sup>1</sup> ye should be romanized initially, after the vowel characters а, е, ё, и, о, ө, у, ү, ы, э, ю, and я, and after й, ъ, and ъ

## Ukrainian

### 1.

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<sup>30</sup>.

The current national system of romanization was approved by the Minister of Justice (also Chairman of the Committee on Law Terminology) in 1996<sup>31</sup>. 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 ambiguity and some characters (ї) are not distinguished in romanization. Also, for practical use further simplifications are allowed which would make the romanization imprecise.

### 5.

The national system of romanization (1996) is as follows.

А а	a	З з	z	О о	o	Ч ч	ch
Б б	b	И и	y	П п	p	Ш ш	sh
В в	v	І і	i	Р р	r	Щ щ	sch
Г г	h, gh <sup>1</sup>	Ї ї	i, yi <sup>2</sup>	С с	s	Ь ь	'
Ґ ґ	g	Й й	i, y <sup>2</sup>	Т т	t	Ю ю	iu, yu <sup>2</sup>
Д д	d	К к	k	У у	u	Я я	ia, ya <sup>2</sup>
Е е	e	Л л	l	Ф ф	f	.	*
Є є	ie, ye <sup>2</sup>	М м	m	Х х	kh		
Ж ж	zh	Н н	n	Ц ц	ts		

<sup>1</sup> gh is used in the romanization of ґr (zgh)

<sup>2</sup> the second variant is used at the beginning of a word

The BGN/PCGN 1965 System gives, as a single block, different Roman equivalents to the following Ukrainian letters (the national equivalents are in brackets):

r (h/gh) h	l (i/yi) yi	ш (sch) shch	я (ia/ya) ya
e (ie/ye) ye	ї (i/y) y	ю (iu/yu) yu	

## Uzbek

### 1.

No romanization systems for Uzbek 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 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 although there are some variations of letters. Full implementation of it has been postponed to the year 2005<sup>32</sup>.

<sup>30</sup> *Provisional Rules of Reproducing Letters of the Ukrainian Alphabet with Latin (English) Characters*. UNEGN, 17th Session, New York, 13-24 June 1995, WP No. 73.

<sup>31</sup> *Romanization of Ukrainian Geographical Names*. UNEGN, 18th Session, Geneva, 12-23 August 1996, WP No. 78.

<sup>32</sup> *Report by the Convenor of the Working Group on Romanization Systems*. UNEGN, 18th Session, Geneva, 12-23 August 1996, WP No. 21.

## 5.

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

А а	а	И и	і	С с	s	Ь ь	(omitted)
Б б	b	Й й	y	Т т	t	Э э	e
В в	v	К к	k	У у	u	Ю ю	yu
Г г	g	Л л	l	Ф ф	f	Я я	ya
Д д	d	М м	m	Х х	x	Ў ў	o' <sup>2</sup>
Е е	e, ye <sup>1</sup>	Н н	n	Ц ц	ts	Қ қ	q
Ё ё	yo	О о	o	Ч ч	ch	Ғ ғ	g' <sup>2</sup>
Ж ж	j	П п	p	Ш ш	sh	Ҳ ҳ	h
З з	z	Р р	r	Ъ ъ	(not known)		

<sup>1</sup> the Cyrillic letter e is pronounced [je] (in the new orthography ye) word-initially and after vowels, й, ъ and ё.

<sup>2</sup> these letters are decreed to have a different shape when handwritten:

Printed	Handwritten
g'	ḡ (or ḡ)
o'	ō (or ð)

The BGN/PCGN 1979 System gives, as a single block, different Roman equivalents to the following Uzbek letters (the official Roman equivalents are in brackets):

в (v)	w	ъ	ʔ (o')	ў
х (x)	kh	ь	ғ (g')	gh