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

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

Submitted by the United Nations Group of Experts on Geographical Names

* E/CONF.94/1.
Report Contents

Introduction 1

List of languages, countries and romanization systems 3

Romanization systems:
Languages/scripts covered by systems recommended by the United Nations 9
(Amharic 9 • Arabic 11 • Assamese 15 • Bengali 16 • Bulgarian 18 • Chinese 19
• Greek 20 • Gujarati 22 • Hebrew 24 • Hindi 25 • Kannada 28 • Khmer 30 •
Macedonian Cyrillic 34 • Malayalam 35 • Marathi 37 • Mongolian (in China) 38
• Nepali 39 • Oriya 39 • Persian 41 • Punjabi 43 • Russian 44 • Serbo-Croatian
Cyrillic 45 • Tamil 46 • Telugu 47 • Thai 49 • Tibetan 52 • Uighur 54 • Urdu 55)

Other languages/scripts 57
(Armenian 57 • Burmese 58 • Byelorussian 60 • Dzongkha 61 • Georgian 63 •
Japanese 64 • Kazakh 66 • Kirghiz 67 • Korean 67 • Lao 70 • Maldivian 72 •
Mongolian (Cyrillic) 73 • Pashto 74 • Sinhalese 75 • Tajik 76 • Tigrinya 77 •
Ukrainian 79)

Annex: Languages that have recently adopted Roman alphabets 81
(Azerbaijani 81 • Turkmen 81 • Uzbek 82)

INTRODUCTION

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

It has been agreed from the beginning that one of the general aims of international
standardization is "to arrive at an agreement on a single romanization system, based on
scientific principles, from each non-Roman alphabet or script, for international application"
(Resolution 9 of the first UN Conference on the Standardization of Geographical Names, 1967). The Fourth United Nations Conference on the Standardization of Geographical Names (1982) also adopted Resolution 15 which called for guidelines governing the consideration of romanization systems. The resolution stated that

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

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

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

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

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

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

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

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

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

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

**Updates**

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

**LIST OF LANGUAGES, COUNTRIES AND ROMANIZATION SYSTEMS**


Languages are arranged alphabetically according to 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. The numerals that follow in parentheses refer to the United Nations Conference on the Standardization of Geographical Names (preceding the slash), and to the resolution number (following the slash).

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

<table>
<thead>
<tr>
<th>Language</th>
<th>Writing System</th>
<th>Country (Countries)</th>
<th>Romanization Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>United Nations</td>
</tr>
<tr>
<td>Amharic</td>
<td>Ethiopic</td>
<td>Ethiopia</td>
<td>1967 (I/17)</td>
</tr>
<tr>
<td>Arabic</td>
<td>Perso-Arabic</td>
<td>(General)</td>
<td>1972 (II/8) = BGN/PCGN 1956</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Algeria</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bahrain</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chad</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comoros (the)</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Djibouti</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Egypt</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eritrea</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Iraq</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Israel</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Jordan</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kuwait</td>
<td>(see General)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BGN/PCGN 1967</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>I.G.N. System 1973</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ISO 233:1984 (transliteration)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ISO 233-2:1993 (simplified transliteration)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>National: Survey of Egypt</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>National: Royal Jordanian Geographic Centre System</td>
</tr>
<tr>
<td>Language</td>
<td>Language</td>
<td>Country</td>
<td>Year</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Lebanon</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Libya</td>
<td>(see General)</td>
<td></td>
<td>National: 1963</td>
</tr>
<tr>
<td>Mauritania</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morocco</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oman</td>
<td>(see General)</td>
<td></td>
<td>National: 1972</td>
</tr>
<tr>
<td>Qatar</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somalia</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sudan</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syria</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UAE</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yemen</td>
<td>(see General)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armenian</td>
<td>Armenian</td>
<td>Armenia</td>
<td>---</td>
</tr>
<tr>
<td>Assamese</td>
<td>Bengali</td>
<td>India</td>
<td>1972 (II/11),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1977 (III/12)</td>
</tr>
<tr>
<td>Bengali</td>
<td>Bengali</td>
<td>Bangladesh</td>
<td>1972 (II/11),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1977 (III/12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>India</td>
<td>1972 (II/11),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1977 (III/12)</td>
</tr>
<tr>
<td>Bulgarian</td>
<td>Cyrillic</td>
<td>Bulgaria</td>
<td>1977 (III/10)</td>
</tr>
<tr>
<td>Burmese</td>
<td>Burmese</td>
<td>Myanmar</td>
<td>---</td>
</tr>
<tr>
<td>Byelorussian</td>
<td>Cyrillic</td>
<td>Belarus</td>
<td>---</td>
</tr>
<tr>
<td>Chinese</td>
<td>Chinese</td>
<td>China</td>
<td>1977 (III/8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>= BGN/PCGN 1979; = ISO 7098:1991</td>
</tr>
<tr>
<td>Dari</td>
<td>Perso-Arabic</td>
<td>Afghanistan</td>
<td>1967 (I/13)</td>
</tr>
<tr>
<td>Dzongkha</td>
<td>Tibetan</td>
<td>Bhutan</td>
<td>---</td>
</tr>
<tr>
<td>Georgian</td>
<td>Georgian</td>
<td>Georgia</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>= BGN/PCGN 1979; = ISO 7098:1991</td>
</tr>
<tr>
<td>ISO 9984:1996</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greek</td>
<td>Greek</td>
<td>Cyprus</td>
<td>Date</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Gujarati</td>
<td>Gujarati</td>
<td>India</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td>Hebrew</td>
<td>Hebrew</td>
<td>Israel</td>
<td>1977 (III/13)</td>
</tr>
<tr>
<td>Hindi</td>
<td>Devanagari</td>
<td>India</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td>Kannada</td>
<td>Kannada</td>
<td>India</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td>Kazakh</td>
<td>Cyrillic</td>
<td>Kazakhstan</td>
<td>---</td>
</tr>
<tr>
<td>Khmer</td>
<td>Khmer</td>
<td>Cambodia</td>
<td>1972 (II/10)</td>
</tr>
<tr>
<td>Kirghiz</td>
<td>Cyrillic</td>
<td>Kyrgyzstan</td>
<td>---</td>
</tr>
<tr>
<td>Korean</td>
<td>Korean</td>
<td>(General)</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lao</td>
<td>Lao</td>
<td>Lao People's Democratic Republic (the)</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macedonian</td>
<td>Cyrillic</td>
<td>The former Yugoslav Republic of Macedonia</td>
<td>1977 (III/11)</td>
</tr>
<tr>
<td>Malayalam</td>
<td>Malayalam</td>
<td>India</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td>Marathi</td>
<td>Devanagari</td>
<td>India</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td>Language</td>
<td>Script</td>
<td>Country</td>
<td>Code</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
<td>---------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Mongolian</td>
<td>Cyrillic</td>
<td>Mongolia</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Mongolian</td>
<td>China</td>
<td>1977 (II/8)</td>
</tr>
<tr>
<td>Nepali</td>
<td>Devanagari</td>
<td>Nepal</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td>Oriya</td>
<td>Oriya</td>
<td>India</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td>Pashto</td>
<td>Perso-Arabic</td>
<td>Afghanistan</td>
<td>—</td>
</tr>
<tr>
<td>Persian</td>
<td>Perso-Arabic</td>
<td>Iran (Islamic Republic of)</td>
<td>1967 (I/13)</td>
</tr>
<tr>
<td>Punjabi</td>
<td>Gurmukhi</td>
<td>India</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td>Russian</td>
<td>Cyrillic</td>
<td>Belarus</td>
<td>1987 (V/18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kazakhstan</td>
<td>1987 (V/18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kyrgyzstan</td>
<td>1987 (V/18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Russia</td>
<td>1987 (V/18)</td>
</tr>
<tr>
<td>Serbo-Croatian</td>
<td>Cyrillic</td>
<td>Bosnia and Herzegovina</td>
<td>1977 (II/11)</td>
</tr>
<tr>
<td>Sinhalese</td>
<td>Sinhalese</td>
<td>Sri Lanka</td>
<td>—</td>
</tr>
<tr>
<td>Tajik</td>
<td>Cyrillic</td>
<td>Tajikistan</td>
<td>—</td>
</tr>
<tr>
<td>Tamil</td>
<td>Tamil</td>
<td>India</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Singapore</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sri Lanka</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td>Telugu</td>
<td>Telugu</td>
<td>India</td>
<td>1972 (II/11), 1977 (III/12)</td>
</tr>
<tr>
<td>Thai</td>
<td>Thai</td>
<td>Thailand</td>
<td>1967 (II/14)</td>
</tr>
<tr>
<td>Tigrinya</td>
<td>Ethiopic</td>
<td>Eritrea</td>
<td>—</td>
</tr>
<tr>
<td>Ukrainian</td>
<td>Cyrillic</td>
<td>Ukraine</td>
<td>—</td>
</tr>
<tr>
<td>Language</td>
<td>Perso-Arabic</td>
<td>Country</td>
<td>Years</td>
</tr>
<tr>
<td>----------</td>
<td>--------------</td>
<td>-----------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Urdu</td>
<td></td>
<td>India</td>
<td>1972 (II/11),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1977 (III/12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pakistan</td>
<td>1972 (II/11),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1977 (III/12)</td>
</tr>
</tbody>
</table>

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

Amharic

The United Nations recommended romanization system was approved in 1967 (resolution I/17), based on the Amharic to English Transliteration System (2nd revised edition, August 1962) by the Imperial Ethiopian Mapping and Geography Institute. The table has been published in volume II of the conference report1.

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

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

**Romanization**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6(A)</th>
<th>7</th>
<th>(1)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ha</td>
<td>hu</td>
<td>hi</td>
<td>ha</td>
<td>ha</td>
<td>he</td>
<td>hi</td>
<td>ho</td>
<td>hwa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>le</td>
<td>lu</td>
<td>li</td>
<td>la</td>
<td>le</td>
<td>li</td>
<td>lo</td>
<td>lwa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ha</td>
<td>hu</td>
<td>hi</td>
<td>he</td>
<td>he</td>
<td>hi</td>
<td>ho</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>me</td>
<td>mu</td>
<td>mi</td>
<td>ma</td>
<td>me</td>
<td>mj</td>
<td>mo</td>
<td>mwa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>su</td>
<td>si</td>
<td>sa</td>
<td>se</td>
<td>si</td>
<td>so</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rğ</td>
<td>ru</td>
<td>ri</td>
<td>ra</td>
<td>re</td>
<td>rî</td>
<td>ro</td>
<td>rwa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shé</td>
<td>shu</td>
<td>shi</td>
<td>sha</td>
<td>she</td>
<td>shi</td>
<td>sho</td>
<td>shwa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k’ê</td>
<td>k’u</td>
<td>k’i</td>
<td>k’a</td>
<td>k’e</td>
<td>k’i</td>
<td>k’o</td>
<td>k’we</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

Sounds identical to fourth-order form.

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

Other systems of romanization

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

<table>
<thead>
<tr>
<th>UN</th>
<th>BGN/PCGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st order e</td>
<td>e</td>
</tr>
<tr>
<td>1st order a</td>
<td>ā (į hā, ą hā, ı̂ hā, ā, ā, ō 'ā)</td>
</tr>
<tr>
<td>3rd order i</td>
<td>ĭ</td>
</tr>
<tr>
<td>5th order e</td>
<td>ē</td>
</tr>
<tr>
<td>6th order ĭ</td>
<td>ĭ</td>
</tr>
</tbody>
</table>

Other differences:
1. The vowel characters in row 16 should be romanized in the BGN/PCGN System ā, ū, ā, ē, ĭ, and ŏ in all other positions.
2. The characters in row 20 should be romanized in the BGN/PCGN System ą, ū, ĭ, ē, ĭ, ĕ, and ŏ in all other positions.
3. The syllables gwe, hwe, kwe, k’we of the UN system are treated as variations of the syllables go, ho, ko, k’o in the BGN/PCGN system.
4. For documentation purposes the following consonants may be romanized with diacritical marks in the BGN/PCGN system: ĥ (row 3), s (row 5), ĥ (row 13), ĥ (row 18), ũ (row 31).

Arabic

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

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

---

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

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

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

### Romanization

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

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>ع</td>
<td>م</td>
<td>م</td>
<td>م</td>
<td>م</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
</tr>
<tr>
<td>ل</td>
<td>ل</td>
<td>ل</td>
<td>ل</td>
<td>ل</td>
<td>ز</td>
<td>ز</td>
<td>ز</td>
<td>ز</td>
<td>ز</td>
</tr>
<tr>
<td>ب</td>
<td>ب</td>
<td>ب</td>
<td>ب</td>
<td>ب</td>
<td>س</td>
<td>س</td>
<td>س</td>
<td>س</td>
<td>س</td>
</tr>
<tr>
<td>ت</td>
<td>ت</td>
<td>ت</td>
<td>ت</td>
<td>ت</td>
<td>ش</td>
<td>ش</td>
<td>ش</td>
<td>ش</td>
<td>ش</td>
</tr>
<tr>
<td>ث</td>
<td>ث</td>
<td>ث</td>
<td>ث</td>
<td>ث</td>
<td>ص</td>
<td>ص</td>
<td>ص</td>
<td>ص</td>
<td>ص</td>
</tr>
<tr>
<td>ج</td>
<td>ج</td>
<td>ج</td>
<td>ج</td>
<td>ج</td>
<td>ض</td>
<td>ض</td>
<td>ض</td>
<td>ض</td>
<td>ض</td>
</tr>
<tr>
<td>ح</td>
<td>ح</td>
<td>ح</td>
<td>ح</td>
<td>ح</td>
<td>ط</td>
<td>ط</td>
<td>ط</td>
<td>ط</td>
<td>ط</td>
</tr>
<tr>
<td>خ</td>
<td>خ</td>
<td>خ</td>
<td>خ</td>
<td>خ</td>
<td>ه</td>
<td>ه</td>
<td>ه</td>
<td>ه</td>
<td>ه</td>
</tr>
<tr>
<td>د</td>
<td>د</td>
<td>د</td>
<td>د</td>
<td>د</td>
<td>غ</td>
<td>غ</td>
<td>غ</td>
<td>غ</td>
<td>غ</td>
</tr>
<tr>
<td>ذ</td>
<td>ذ</td>
<td>ذ</td>
<td>ذ</td>
<td>ذ</td>
<td>ف</td>
<td>ف</td>
<td>ف</td>
<td>ف</td>
<td>ف</td>
</tr>
<tr>
<td>ن</td>
<td>ن</td>
<td>ن</td>
<td>ن</td>
<td>ن</td>
<td>ن</td>
<td>ن</td>
<td>ن</td>
<td>ن</td>
<td>ن</td>
</tr>
<tr>
<td>ه</td>
<td>ه</td>
<td>ه</td>
<td>ه</td>
<td>ه</td>
<td>ه</td>
<td>ه</td>
<td>ه</td>
<td>ه</td>
<td>ه</td>
</tr>
<tr>
<td>و</td>
<td>و</td>
<td>و</td>
<td>و</td>
<td>و</td>
<td>و</td>
<td>و</td>
<td>و</td>
<td>و</td>
<td>و</td>
</tr>
<tr>
<td>ي</td>
<td>ي</td>
<td>ي</td>
<td>ي</td>
<td>ي</td>
<td>ي</td>
<td>ي</td>
<td>ي</td>
<td>ي</td>
<td>ي</td>
</tr>
</tbody>
</table>

A Not romanized word-initially.
B Not romanized, but see romanizations accompanying alif (i) in the table for vowels.
C In certain endings, an original tā’ (ذ) is written s, i.e., like hā’ (ح) with two dots, and is known as tā’ marbūfah. It is romanized h, except in the construct form of feminine nouns, where it
is romanized t, instead.

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

<table>
<thead>
<tr>
<th>1</th>
<th>١</th>
<th>a</th>
<th>5</th>
<th>٥</th>
<th>(A)</th>
<th>11</th>
<th>١١</th>
<th>١١</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>٢</td>
<td>٢</td>
<td>٢</td>
<td>٢</td>
<td>٢</td>
<td>٢</td>
<td>٢</td>
<td>٢</td>
</tr>
<tr>
<td>3</td>
<td>٣</td>
<td>٣</td>
<td>٣</td>
<td>٣</td>
<td>٣</td>
<td>٣</td>
<td>٣</td>
<td>٣</td>
</tr>
<tr>
<td>4</td>
<td>٤</td>
<td>i</td>
<td>9</td>
<td>٩</td>
<td>٩</td>
<td>٩</td>
<td>٩</td>
<td>٩</td>
</tr>
<tr>
<td>5</td>
<td>٥</td>
<td>u</td>
<td>10</td>
<td>١٠</td>
<td>١٠</td>
<td>١٠</td>
<td>١٠</td>
<td>١٠</td>
</tr>
</tbody>
</table>

^ Marks absence of the vowel.  
^ 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, s, š, ẓ, l, n*) the *l* of the definite article is assimilated with the first consonant of the word: أش-شريقة.

**Other systems of romanization**

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

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

<table>
<thead>
<tr>
<th>UN</th>
<th>I.G.N.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>a, e, è, e^A</td>
</tr>
<tr>
<td>å</td>
<td>å, e^A</td>
</tr>
<tr>
<td>ã</td>
<td>å</td>
</tr>
<tr>
<td>ð</td>
<td>ð</td>
</tr>
<tr>
<td>h</td>
<td>h</td>
</tr>
<tr>
<td>i</td>
<td>i, e^A</td>
</tr>
<tr>
<td>î</td>
<td>î, ê^A</td>
</tr>
</tbody>
</table>

^ Used according to the local pronunciation.  
^ ne is used word-finally after a, e, i, o.  
^ gu is used before e and i.  
^ ss is used between vowels.

\( c \) is used in some conventional names.

\( i \) is used as second element of a diphthong.

\( y \) is used word-initially or between vowels.

\( aa \) is specific to Lebanon.

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

\[
\begin{array}{llll}
\text{UN} & \text{ISO} \\
\text{ā (ī)} &= \dot{a} & h &= \dot{h} & \text{th} &= \dot{t} \\
\dot{a} &= ay & i^n &= i & u^n &= û \\
a^n &= \acute{a} & j &= \dot{g} & y (\ddot{a}) &= \dot{y} \\
\dd &= \dd & kh &= h & y (\dddot{a}) &= y \\
\ddh &= \dd & s &= \dot{s} & z &= \ddot{z} \\
gh &= \acute{g} & sh &= \dot{s} & \dd &= \ddh \\
h (s) &= \dd & t &= \dot{t} \\
\end{array}
\]

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

\[
\begin{array}{llll}
\text{UN} & \text{RJGC} \\
\text{-ah (a)} &= -a & h &= h & \dot{t} &= \dot{t} \\
\dd &= \dd & s &= \dot{s} & z &= \ddh \\
\end{array}
\]

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

\[
\begin{array}{llll}
\text{UN} & \text{SES} \\
\acute{a} &= a & h &= h & \text{th} &= \text{th} (t) \\
\dd &= \acute{a} (a) & i &= i & û &= û \\
\text{-ah (a)} &= -a & j &= g (j) & z &= z (d) \\
\ddh &= \ddh (au) & q &= q (k) & \dd &= \dd \\
\dd &= \ddh (ai) & s &= s (c) & \dd &= \dd \\
\end{array}
\]

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

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

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

Assamese

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

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

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

Romanization

I. Independent vowel characters

<table>
<thead>
<tr>
<th>1</th>
<th>आ</th>
<th>इ</th>
<th>ई</th>
<th>उ</th>
<th>ऊ</th>
<th>ऋ</th>
<th>ऌ</th>
<th>ऍ</th>
<th>ऎ</th>
<th>ए</th>
<th>ऐ</th>
<th>ऒ</th>
<th>ओ</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>आ</td>
<td>इ</td>
<td>ई</td>
<td>उ</td>
<td>ऊ</td>
<td>ऋ</td>
<td>ऌ</td>
<td>ऍ</td>
<td>ऎ</td>
<td>ए</td>
<td>ऐ</td>
<td>ऒ</td>
<td>ओ</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>1</th>
<th>आ</th>
<th>इ</th>
<th>ई</th>
<th>उ</th>
<th>ऊ</th>
<th>ऋ</th>
<th>ऌ</th>
<th>ऍ</th>
<th>ऎ</th>
<th>ए</th>
<th>ऐ</th>
<th>ऒ</th>
<th>ओ</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>आ</td>
<td>इ</td>
<td>ई</td>
<td>उ</td>
<td>ऊ</td>
<td>ऋ</td>
<td>ऌ</td>
<td>ऍ</td>
<td>ऎ</td>
<td>ए</td>
<td>ऐ</td>
<td>ऒ</td>
<td>ओ</td>
</tr>
</tbody>
</table>

A: Exceptions: गु; रु; शु; छु; न्तु; न्तु.
B: Exception: रु.
C: Exception: रु.


III. Other symbols ( stands for any consonant character)

1. m 2. ṭ m 3. ṭ h 4. (A)

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

IV. Consonant characters

2. ख kha 7. छ cha 12. ठ tha 17. ठ tha 22. फ pha 27. र ra 32. झ sha
3. ग ga 8. ज ja 13. झ ja(A) 18. ज da 23. झ ba 28. झ la 32. स sa
4. घ gha 9. झ jha 14. झ dha(A) 19. झ dha 24. झ bha 29. झ va 33. ह ha
5. ङ na 10. ङ fi(a) 15. ङ na 20. ङ na 25. म ma

A Dotted variants of the characters: र ra; र ha; र ya.

V. Ligatures

See the section on the romanization of Bengali.

Other systems of romanization

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

Bengali

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

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

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

Romanization

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

I. Independent vowel characters

\[
\begin{align*}
1 & \text{ a } & 3 & \text{ i } & 5 & \text{ u } & 7 & \text{ } & 8 & \text{ e } & 10 & \text{ o } \\
2 & \text{ ã } & 4 & \text{ ē } & 6 & \text{ ū } & 9 & \text{ ë } & 11 & \text{ au } \\
\end{align*}
\]

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

\[
\begin{align*}
1 & \ast & 3 & \text{ ì } & 5 & \text{ u}^\text{A} & 7 & \text{ } & 8 & \text{ e } & 10 & \text{ o } \\
2 & \text{ ã } & 4 & \text{ ē } & 6 & \text{ ū}^\text{B} & 9 & \text{ ë } & 11 & \text{ au } \\
\end{align*}
\]

\(^{A}\) Exceptions: ꔖ gu; ꔖ ru; ꔖ shu; ꔖ hu; ꔖ ntu; ꔖ stu.

\(^{B}\) Exception: ꔖ rū.

\(^{C}\) Exception: ꔖ hr.

III. Other symbols (\(\ast\) stands for any consonant character)

\[
\begin{align*}
1 & \text{ m } & 2 & \text{ ṁ } & 3 & \text{ ḫ } & 4 & \text{ ḫ } & (\text{A}) \\
\end{align*}
\]

\(^{A}\) Pronunciation without a vowel; special form: ḫ t.

IV. Consonant characters

\[
\begin{align*}
1 & \text{ k } & 6 & \text{ c } & 11 & \text{ ja } & 16 & \text{ t } & 21 & \text{ p } & 26 & \text{ ja}^\text{A} & 29 & \text{ sh } & 34 & \text{ s } \\
2 & \text{ kh } & 7 & \text{ ch } & 12 & \text{ th } & 17 & \text{ t } & 22 & \text{ ph } & 27 & \text{ r } & 30 & \text{ sh } & 35 & \text{ s } \\
3 & \text{ g } & 8 & \text{ j } & 13 & \text{ da}^\text{A} & 18 & \text{ d } & 23 & \text{ b } & 28 & \text{ la } & 31 & \text{ sa } & 36 & \text{ s } \\
4 & \text{ gh } & 9 & \text{ jh } & 14 & \text{ dh}^\text{A} & 19 & \text{ d } & 24 & \text{ b } & 29 & \text{ ha } & 32 & \text{ h } & 37 & \text{ s } \\
5 & \text{ ṛ } & 10 & \text{ ṛ } & 15 & \text{ ṇ } & 20 & \text{ na } & 25 & \text{ ma } &  \\
\end{align*}
\]

\(^{A}\) Dotted variants of the characters: ḫ rā; ḫ rha; ḫ ya.

V. Ligatures

Adscript forms of some consonants

Example

\(\acute{r}\): ꔖ rka

\(\acute{r}\): ꔖ gra

\(\text{j}\): ꔖ tyā

Other ligatures (the list is not complete)

\[
\begin{align*}
\text{kka} & \text{kta} & \text{kna} & \text{kma} & \text{kra} & \text{kla} & \text{kva} & \text{kṣa} & \text{kṣa} & \text{kṣa} & \text{kṣa} & \text{kṣa} \\
\text{kṣa} & \text{gga} & \text{gda} & \text{gdha} & \text{gma} & \text{gra} & \text{gla} & \text{gha} & \text{ṅka} & \text{ṅga} & \text{chha} \\
\text{c} & \text{c} & \text{c} & \text{c} & \text{c} & \text{c} & \text{c} & \text{c} & \text{c} & \text{c} & \text{c} & \text{c} & \text{c} \\
\text{chēhha} & \text{chēhha} & \text{chē} & \text{jjā} & \text{jjvā} & \text{jjhā} & \text{jñā} & \text{jva} & \text{ñcha} & \text{ñcha} & \text{ñcha} & \text{ñjhā} & \text{ñjhā}
\end{align*}
\]
Other systems of romanization

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

Bulgarian

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

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

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

---

seems highly unlikely.

Romanization

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

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

Other systems of romanization

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

ж (ж) zh  ц (с) c  щ (шт) šht  ю (ю) yu
й (й) y  ч (ч) ch  ъ (ъ) â  я (я) ya
х (х) h  ш (ш) šh  ъ (ъ) â

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

ж (ž) zh  ц (с) ts  щ (št) šht  ю (ju) yu
й (j) y  ч (č) ch  ъ (á) ũ  я (ja) ya
х (h) kh  ш (š) sh  ъ (j)

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

Chinese

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

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

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

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

Other systems of romanization

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

<table>
<thead>
<tr>
<th>WG</th>
<th>UN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syllable beginnings:</td>
<td>sz = s</td>
</tr>
<tr>
<td>ch = zh, ji&lt;sup&gt;A&lt;/sup&gt;</td>
<td>t = d</td>
</tr>
<tr>
<td>ch’ = ch, qi&lt;sup&gt;q&lt;/sup&gt;</td>
<td>t’ = t</td>
</tr>
<tr>
<td>hs = x</td>
<td>ts = z</td>
</tr>
<tr>
<td>j’ = r</td>
<td>ts’ = c</td>
</tr>
<tr>
<td>k = g</td>
<td>tz = z</td>
</tr>
<tr>
<td>k’ = k’</td>
<td>tz’ = c</td>
</tr>
<tr>
<td>p = b</td>
<td>Syllable endings:</td>
</tr>
<tr>
<td>p’ = p</td>
<td>ê = e</td>
</tr>
<tr>
<td>ss = s</td>
<td>eh = e</td>
</tr>
</tbody>
</table>

<sup>A</sup> Before i (except the ending -ih) and ü.
<sup>B</sup> Word-initially.
<sup>C</sup> After p, p’, m, f, w, y.
<sup>D</sup> After t, t’, n, l, ch, ch’, j, ts, ts’, s.
<sup>E</sup> After k, k’, h and syllable-initially.
<sup>F</sup> After n, l.

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

Greek

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

---

<sup>9</sup> Adapted from Klaus Kaden, Die wichtigsten Transkriptionssysteme für die chinesische Sprache. VEB Verlag Enzyklopädie, Leipzig 1975, S. 145.

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

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

**Romanization**

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

<table>
<thead>
<tr>
<th>1</th>
<th>A α a</th>
<th>5</th>
<th>E ε e</th>
<th>11</th>
<th>Λ λ l</th>
<th>17</th>
<th>P ρ r</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>B β v</td>
<td>6</td>
<td>Z ζ z</td>
<td>12</td>
<td>M μ m</td>
<td>18</td>
<td>Σ σ σ s</td>
</tr>
<tr>
<td>3</td>
<td>Γ γ g</td>
<td>7</td>
<td>Η η i</td>
<td>13</td>
<td>Σ ν η n</td>
<td>20</td>
<td>Υ υ υ γ</td>
</tr>
<tr>
<td>4</td>
<td>Δ δ d</td>
<td>10</td>
<td>Κ κ k</td>
<td>16</td>
<td>Π π ρ</td>
<td>24</td>
<td>Ω ω ω o</td>
</tr>
</tbody>
</table>

^A Before the consonants β, γ, δ, ζ, λ, μ, ν, ρ and all vowels.
^B Before the consonants θ, ξ, π, σ, τ, ϕ, χ, ψ and at the end of the word.
^C At the beginning of the word and at the end of the word.
^D In the middle of the word.

**Notes**

1. In order to ensure reversibility a sub-macron may be added to the following characters and character combinations: 1 aw ay/af, 3 γ ng, γ η nx, γ ch, 5 η i, η iv/if, 7 η i, η iv/if, 24 ω ο.
2. Some examples of the ELOT 743 table contain also combinations in parentheses αι, αι, αι, γκ, ει, ει, ει, οι, οι, υ which are converted according to the conversion rules of each character they consist of, resp. ai, ai, ai, gi, ei, ei, ei, ni, oi, oi, oi, yi.
3. The combinations aw, eu, ηυ, ou are converted according to the conversion rules of each character when the vowel before u has an accent or when u has the diaeresis sign.
4. When the character combinations aw, eu, ηυ, ou are stressed in Greek, in the Roman script the accent is set on the vowel (άυ, áf, év, éf, iv, if).

**Other systems of romanization**

The BGN/PCGN 1962 System that continued to be used until recently, provided for the romanization, as a single block, of the following characters or character combinations differently from the UN System (the Greek character is followed in parentheses by the romanization according to the UN System):
\( \alpha (av/af) \text{ av} \quad \delta (d) \quad dh, d^p \quad \mu \pi (b/mp) \text{ b}^b, \text{mb}^c \quad \chi (ch) \quad \text{kh} \)

\( \gamma (g) \quad g, \gamma^A \quad \epsilon (eI) \quad i^b \quad \nu (nt) \quad d^b, nd^c, nt^f \quad \nu (y) \quad i \quad \nu (yi) \quad i \)

\( \gamma^A \text{ (nch) } \text{ nhkh} \quad \epsilon (ev/ef) \text{ ev} \quad \eta (iv/iv) \text{ iv} \quad \nu (yi) \quad i \)

\( \gamma^A \text{ (gk) } \text{ g}^b, \text{ng}^c \quad \eta (iv/ivf) \text{ iv} \quad \nu (yi) \quad i \)

\( ^A \) Before \( \alpha, \varepsilon, \eta, \iota, \omicron, \upsilon \) and \( \upsilon I \).

\( ^B \) Initially.

\( ^C \) Medially.

\( ^D \) Between \( \nu \) and \( \rho \).

\( ^E \) But \( \epsilon I = \epsilon i \) in both systems.

\( ^F \) In the combination \( \nu \tau \zeta \).

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

- \( \Lambda \eta \quad I \iota \text{ or I}^\gamma \text{ i}^- \)
- \( \Omega \omega \quad O \delta \text{ or O}^- \text{ o}^- \)

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 \( \text{AY}, \alpha \nu, \text{EY}, \epsilon \nu, \text{OY}, \omicron \nu \) which are mapped into Roman script as \( \text{AU}, \alpha u, \text{EU}, \epsilon u, \text{OU}, \omicron u \) respectively.

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

**Gujarati**

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

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

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

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

**Romanization**

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

**I. Independent vowel characters**

<table>
<thead>
<tr>
<th>1</th>
<th>a</th>
<th>3</th>
<th>i</th>
<th>5</th>
<th>u</th>
<th>7</th>
<th>ē</th>
<th>8</th>
<th>e, ē</th>
<th>10</th>
<th>ā</th>
<th>o, ō</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>ā</td>
<td>4</td>
<td>ē</td>
<td>6</td>
<td>ō</td>
<td>9</td>
<td>ē</td>
<td>11</td>
<td>ai</td>
<td>11</td>
<td>ā</td>
<td>au</td>
</tr>
</tbody>
</table>

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

| 1 | * | a | 3 | * | i | 5 | * | u | 7 | * | ē | 8 | * | e, ē | 10 | * | ā | o, ō |
| 2 | * | ā | 4 | * | ē | 6 | B | * | ō | 9 | * | ai | 11 | * | ā | au |

^ Exception: ꞏ jī.
B Exception: ꞑ rū.

**III. Other symbols**

(These symbols were missing in the original table.)

| 1 | ō | m | 2 | ō | m | 3 | * | ḷ | 4 | * | (A) |

^ Pronunciation without a vowel.

**IV. Consonant characters**

| 1 | े | ka | 6 | ḍ | cha | 11 | े | ṭa | 16 | ṭ | ta | 21 | ḩ | pa | 26 | ṭ | ya | 30 | ḫ | sha |
| 2 | ṽ | kha | 7 | ṽ | cha | 12 | े | ṭha | 17 | ṭ | tha | 22 | ṽ | fa | 27 | ṭ | ra | 31 | ṽ | sha |
| 3 | ṽ | ga | 2 | ṽ | ja | 13 | े | ṭda | 18 | ṭ | da | 23 | ṽ | ba | 28 | ṭ | la | 32 | ṽ | sa |
| 4 | ṽ | gha | 9 | ṽ | jha | 14 | े | ṭđha | 19 | ṭ | dha | 24 | ṽ | bha | 29 | ṭ | va | 33 | ṽ | ha |
| 5 | ṽ | ṇa | 10 | ṽ | ṇa | 15 | े | ṭṇa | 20 | ṭ | na | 25 | ṽ | ma | 34 | ṽ | ṭa |

**V. Consonant ligatures**

Adscript forms of some consonants

Examples

| 4 | r- | ṽ | rka, ṽ | rma |
| 5 | r- | ṽ | kra, ṽ | bra, ṽ | dra |

Other ligatures (the list is not complete)

| kka | kra | kṣha | khta | khra | gra | ghra | ṇka | ṇkha | ṇga | ṇgha | ṇma |
| kṣ | kṣ | kṣ | kṣ | kṣ | kṣ | kṣ | kṣ | kṣ | kṣ | kṣ | kṣ |
| chra | jja | jḥa | jra | jhra | ṭṭa | ṭṭha | ṭṭha | ṭḍa | ṭḍha | ṭḍhha | ṭṭa |
| tṛ | tṛ | tṛ | tṛ | tṛ | tṛ | tṛ | tṛ | tṛ | tṛ | tṛ | tṛ |
| tna | tra | thra | ḍga | ḍḍa | ḍḍha | ḍḥha | ḍḥha | ṭṭa | ṭṭa | ṭṭa | ṭṭa |

24
Other systems of romanization

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

Hebrew

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

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

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

Romanization

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

| 1 | ש | ' | (ך) | 2 | ר | ז |
| 2 | ב | ב | ג |
| 3 | ד | ד | ח |
| 4 | ג | ג | י |
| 5 | ד | ד | ק |
| 6 | ח | ח | ק
| 7 | w | ל |

---


Vowels (* stands for any consonant character)

1. a
2. a
3. a, o
4. e
5. e
6. e
7. e
8. e
9. i
10. i
11. i
12. i
13. o
14. u
15. u

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 two consecutive vowels or a consonant and a vowel; at the beginning of a word it is omitted in transliteration.
3. Does not occur at the beginning of a name.
4. At the beginning of a name or of a syllable after sheva nah these letters carry a dot (dagesh qal).
5. In some maps the zere (*), if accentuated, is transliterated é.
6. The sheva (*) is of two kinds: sheva nah, which is omitted in transliteration, and sheva na', which occurs at the beginning of a word or syllable.
7. Final h (ฏ), unless vocalized, is omitted in transliteration.
8. The sign † ('ayin) slightly separates two consecutive vowels or a consonant and a vowel; it is also retained at the beginning and end of a word.
9. In practical application נ (') and י (') are transliterated by the same symbol.

Hindi

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

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.

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).

Romanization

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

I. Independent vowel characters

<table>
<thead>
<tr>
<th>1</th>
<th>a</th>
<th>3</th>
<th>i</th>
<th>5</th>
<th>u</th>
<th>7</th>
<th>v</th>
<th>9</th>
<th>l</th>
<th>10</th>
<th>e</th>
<th>12</th>
<th>o</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>ā</td>
<td>4</td>
<td>ī</td>
<td>6</td>
<td>ū</td>
<td>8</td>
<td>ū̄</td>
<td>11</td>
<td>āi</td>
<td>13</td>
<td>ō</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>1</th>
<th>• a</th>
<th>3</th>
<th>i</th>
<th>5</th>
<th>u</th>
<th>7</th>
<th>v</th>
<th>9</th>
<th>l</th>
<th>10</th>
<th>e</th>
<th>12</th>
<th>• o</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>ă</td>
<td>4</td>
<td>ĭ</td>
<td>6</td>
<td>ū</td>
<td>8</td>
<td>ū̆</td>
<td>9</td>
<td>ai</td>
<td>11</td>
<td>• e</td>
<td>13</td>
<td>au</td>
</tr>
</tbody>
</table>

^ With an additional mark: • yā.

b Exceptions: r u; r ū; h ṛ.

III. Other symbols

<table>
<thead>
<tr>
<th>1</th>
<th>• m</th>
<th>2</th>
<th>• m</th>
<th>3</th>
<th>h</th>
<th>4</th>
<th>(A)</th>
</tr>
</thead>
</table>

^ Pronunciation without a vowel: k k.

IV. Consonant characters

<table>
<thead>
<tr>
<th>1</th>
<th>k</th>
<th>6</th>
<th>c</th>
<th>11</th>
<th>t a</th>
<th>16</th>
<th>t a</th>
<th>21</th>
<th>p a</th>
<th>26</th>
<th>y a</th>
<th>30</th>
<th>s h a</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>kha</td>
<td>7</td>
<td>c</td>
<td>12</td>
<td>t h a</td>
<td>17</td>
<td>t h a</td>
<td>22</td>
<td>pha</td>
<td>27</td>
<td>r</td>
<td>31</td>
<td>s h a</td>
</tr>
<tr>
<td>3</td>
<td>g</td>
<td>8</td>
<td>j</td>
<td>13</td>
<td>ḍ</td>
<td>18</td>
<td>ḍ</td>
<td>23</td>
<td>b</td>
<td>28</td>
<td>la</td>
<td>32</td>
<td>s</td>
</tr>
<tr>
<td>4</td>
<td>gh</td>
<td>9</td>
<td>(j j)</td>
<td>14</td>
<td>ḍha</td>
<td>19</td>
<td>ḍha</td>
<td>24</td>
<td>bha</td>
<td>29</td>
<td>va</td>
<td>33</td>
<td>h</td>
</tr>
<tr>
<td>5</td>
<td>ḍ</td>
<td>10</td>
<td>ḍ</td>
<td>15</td>
<td>ō</td>
<td>20</td>
<td>ō</td>
<td>25</td>
<td>ma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^ Dotted variants of the characters: ḍ qa; ḍ kha; ḍ ga; ḍ za; ḍ ra; ḍ rha; ḍ fa.

V. Ligatures

Adscript form:

Examples

r- t rā, ā rdva

Other ligatures (the list is not complete)

kka kkha kta ktva kya kra kla kva kṣha kṣhna kṣhma kṣhya kṣhva
kṣ̄ ḍa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa ṭa
ksa khya khsha gda gdha gna gbha gma gya gra gla ghna ghya
bra br̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍra r̥ ḍr
Notes

1. It is recommended that the vowel अ (a) should always be romanized except when it ends a name. If a name ends with a consonant, the consonant should carry a sub-macron. Such cases, however, will be very rare. For example, कानपुर Kānapur (not Kānapura), जगत Jagat.

2. If each letter of a digraph or any two parts of a trigraph has a distinct independent sound then it should be indicated by a hyphen, thus d-h.

Other systems of romanization

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

<table>
<thead>
<tr>
<th>UN</th>
<th>Hunterian</th>
</tr>
</thead>
<tbody>
<tr>
<td>ā</td>
<td>= ā, a&lt;sup&gt;A&lt;/sup&gt;</td>
</tr>
<tr>
<td>d</td>
<td>= d</td>
</tr>
<tr>
<td>dh</td>
<td>= dh</td>
</tr>
<tr>
<td>i</td>
<td>= i, i&lt;sup&gt;A&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>A</sup> Word-finally.
<sup>B</sup> Exceptionally, e.g. before i.

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

<table>
<thead>
<tr>
<th>UN</th>
<th>ISO</th>
</tr>
</thead>
<tbody>
<tr>
<td>ā</td>
<td>= ē</td>
</tr>
<tr>
<td>ō</td>
<td>= ŏ</td>
</tr>
<tr>
<td>ch</td>
<td>= c</td>
</tr>
<tr>
<td>chh</td>
<td>= h, k&lt;sup&gt;B&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

"
\[ \begin{array}{cccc}
\hat{\imath} & = & \hat{r} & \text{sh} = \hat{s} & \gamma = \hat{y}^C \\
\hat{r} & = & \hat{r} & \text{sh} = \hat{s} & \gamma = \hat{y}^E \\
\tilde{i} & = & \hat{i} & \tilde{i} = \hat{i}^D
\end{array} \]

\text{A} Non-uniform vowels option: if the only scripts in view are without short \( e, o \), then long \( e, o \) may be transliterated \( e, o \), respectively.

\text{B} Used in Tamil.
\text{C} Used in Bengali.
\text{D} Used in Malayalam.
\text{E} Used in Telugu.
\text{F} Used in Oriya.

**Kannada**

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

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

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

**Romanization**

I. Independent vowel characters

\[ \begin{array}{ccccccc}
1 & \hat{a} & 3 & \hat{e} & i & 5 & \hat{u} & u & 8 & \hat{e} & \hat{c} & 11 & \hat{a} & \hat{o} \\
2 & \hat{a} & 4 & \hat{i} & 6 & \hat{u} & \hat{u} & 9 & \hat{e} & e & 12 & \hat{o} & o \\
 & 7 & \hat{e} & \hat{a} & 10 & \hat{a} & \hat{a} & 13 & \hat{e} & \hat{a} & \hat{u}
\end{array} \]

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

1 1. a^A 3 3. j^B 5 υ υ^C 8 • 6 11 • 8 0 14 o m
2 2. a 4 4. i 6 6. u^D 9 • 6 e 12 • 10 o 15 • h
7 7. i 10 • 8 ai 13 • 8 au 16 • (E)

A Some consonants do not have the diacritic, e.g. h kha, t ja, r na, t ba, s la.
B Variations of the diacritic, e.g. h gi, o di.
C Variations of the diacritic: o pu, o phu, o vu.
D Variations of the diacritic: o pü, o phu, o vü.
E Pronunciation without a vowel: s k.

III. Consonant characters

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka</td>
<td>kha</td>
<td>ga</td>
<td>gha</td>
<td>na</td>
</tr>
<tr>
<td>cha</td>
<td>chha</td>
<td>ja</td>
<td>jha</td>
<td>na</td>
</tr>
<tr>
<td>ta</td>
<td>tha</td>
<td>da</td>
<td>dha</td>
<td>ma</td>
</tr>
<tr>
<td>pa</td>
<td>pha</td>
<td>ba</td>
<td>bha</td>
<td>ma</td>
</tr>
<tr>
<td>ya</td>
<td>ra</td>
<td>la</td>
<td>va</td>
<td>la</td>
</tr>
<tr>
<td>sha</td>
<td>sha</td>
<td>sa</td>
<td>ha</td>
<td></td>
</tr>
</tbody>
</table>

IV. Subscript consonant characters

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>k^A</td>
<td>kh</td>
<td>g</td>
<td>gh</td>
<td>n</td>
</tr>
<tr>
<td>ch</td>
<td>chh</td>
<td>j</td>
<td>jh</td>
<td>n</td>
</tr>
<tr>
<td>t^B</td>
<td>th</td>
<td>d</td>
<td>dh</td>
<td></td>
</tr>
<tr>
<td>t^C</td>
<td>th</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p^D</td>
<td>ph</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y</td>
<td>r</td>
<td>l</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>sh</td>
<td>sh</td>
<td>s</td>
<td>h</td>
<td></td>
</tr>
</tbody>
</table>

A Subscript combinations: kr, ksh.
B Subscript combination: tr.
C Subscript combinations: tr, ty.
D Subscript combination: pr.

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

Other systems of romanization

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

The United Nations recommended system was approved in 1972 (resolution II/10), based on the system used by the BGN/PCGN (1972), this being a modified version of the Service Géographique Khmère (SGK) 1959 system. The table and notes on its usage were published in volume II of the conference report17.

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

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

Romanization

I. Consonant characters

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
</tr>
</thead>
<tbody>
<tr>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td>ṇ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A If used with a subscript character the lower element of the character is omitted: ṇ nhchō.
B If used with a subscript character the romanization of the character will be p (see note 4).
C See note 5.

---

II. Subscript consonant characters (see also note 3)
Character numbers correspond to those in Table I.

\[
\begin{array}{ccccccccccc}
1 & \text{k} & 6 & \text{ch} & 11 & \text{d} & 16 & \text{t} & 21 & \text{b} & 26 & \text{y} & 30 & \text{i} & 5 \\
2 & \text{kh} & 7 & \text{chh} & 12 & \text{th} & 17 & \text{th} & 22 & \text{ph} & 27 & \text{i} & 31 & \text{i} & 1 \\
3 & \text{k} & 8 & \text{ch} & 13 & \text{d} & 18 & \text{t} & 23 & \text{p} & 28 & \text{i} & 33 & \text{i} & 1 \\
4 & \text{kh} & 9 & \text{chh} & 14 & \text{th} & 19 & \text{th} & 24 & \text{ph} & 29 & \text{v} \\
5 & \text{ng} & 10 & \text{nh} & 15 & \text{n} & 20 & \text{n} & 25 & \text{m} \\
\end{array}
\]

\(^A\) With the same base character, \(\text{gh}\) is used (\(\text{gh}\)), in all other cases only the lower element is used as subscript, e.g. \(\text{gh} \text{khn\text{h}}\).

III. Independent vowel characters

\[
\begin{array}{ccccccccccc}
1 & \text{h} & \text{e} & 3 & \text{a} & \text{u} & 5 & \text{r} & \text{e} & 7 & \text{e} & \text{a} & 9 & \text{e} & \text{e} & 11 & \text{a} & \text{a} & \text{e} \\
2 & \text{ei} & 4 & \text{au} & 6 & \text{ro} & 8 & \text{e} & \text{e} & 10 & \text{a} & \text{ai} & 12 & \text{au} \\
\end{array}
\]

\(^A\) See note 10.

\(^B\) Variation: \(\text{gh}\)

IV. Vocalic nuclei (* stands for any consonant character)
Where variants in romanization separated by a dash are given the first is to be used in the \(\text{a}\)-series and the second in the \(\text{e}\)-series.

\[
\begin{array}{ccccccccccc}
1 & \text{a\text{--}ea} & 5 & \text{ae} & 9 & \text{eau\text{--}eu} & 13 & \text{e} & \text{e} \\
2 & \text{e\text{--}i} & 6 & \text{o\text{--}u} & 10 & \text{aa} & 14 & \text{ai\text{--}ey} \\
3 & \text{ei\text{--}i} & 7 & \text{o\text{--}u} & 11 & \text{i} & \text{i} & 15 & \text{ao\text{--}oU} \\
4 & \text{ae} & 8 & \text{uO} & 12 & \text{e} & 16 & \text{au\text{--}oU} \\
\end{array}
\]

\(^A\) With the character \(\text{p}\) a ligature \(\text{gh}\) is used (see note 4).

V. Shortened syllables and vocalic nuclei with anusvara or visarga
Where variants in romanization separated by a dash are given the one before the dash is to be used in the \(\text{a}\)-series and the one(s) after the dash in the \(\text{e}\)-series.

\[
\begin{array}{ccccccccccc}
1 & \text{a\text{--}o} & 4 & \text{am\text{--}um} & 7 & \text{ah\text{--}eah} & 10 & \text{ao\text{--}oh\text{--}oah} \\
2 & \text{a\text{--}oa\text{,ea}^A} & 5 & \text{om\text{--}um} & 8 & \text{oh\text{--}uh} & 11 & \text{a\text{,}ang\text{--}eang} \\
3 & \text{a\text{--}oa\text{,ea}^B} & 6 & \text{am\text{--}oam} & 9 & \text{eh} \\
\end{array}
\]

\(^A\) \text{ea} if followed by \(k, ng, h\) (see note 6).

\(^B\) \text{ed} after a short vowel or if followed by \(k, ng, h\) (see note 6).
Notes
1. Khmer consonants are divided into 2 series, the á-series and the ð-series, as indicated in the consonant table. With vocalic nuclei these consonants may produce different romanizations, as shown in the abbreviated vowel table: េ kā, េ krā, េ kō, េ krō. A Khmer consonant in syllable-final position, not accompanied by a vowel marker or by *, should generally be romanized without a vowel letter following: េ kāk, េ anguish (exception: េ pōngrō, also written េ pōngrō and េ pōngrōh).
2. The Khmer diacritical mark ~ or ~ written above an á-series consonant (except ū and ū; see note 4) changes it to the ð-series: េ hēang. The diacritical mark ~ written above an ð-series consonant changes it to the á-series: េ nhrâng. When either of these marks would conflict with another symbol written above a character, the mark ~ or ~ may be written in its place: េ hi, េ dâmrei. (These marks are frequently omitted in Khmer writing, particularly in words of Indic provenance.)
3. The second consonant of a Khmer graphic cluster is generally written below the base consonant in the special form called a “foot”: េ khnâng. There is no foot for the character េ lâ. The "feet" ~ and ~ usually represent the characters ū and ū respectively, rather than ū and ū: េ kdei, េ kântheay, but េ kânrâb.
A "foot" determines the series of the following vocalic nucleus unless it is a nasal (េ ng, េ nh, េ n, េ n, េ m) or េ y, េ r, េ l, េ v, េ s. In that case, the base consonant determines the vocalic series: េ khoông, េ lâ, េ thmâ, េ sây. Syllable-final ū and ū sometimes appear as "feet": េ khrù or េ sây, េ or េ tâng. This practice appears to be optional and such irregular Khmer spellings are not reflected as such in romanization.
4. The combination ū plus 1 is written េ ba. The latter character is a graphic device designed to prevent confusion with េ hâ. The characters ū and ū with the diacritical mark ~ are romanized p in the á-series, rather than as b in the ð-series: េ pang, េ patau. The diacritical mark ~ or ~ is substituted where a conflict with another symbol written above a character would occur: េ pei. The characters ū and ū when accompanied by a "foot" are also romanized as p in the á-series, although the Khmer diacritical mark is generally omitted: េ plâng, េ p'ā, េ prâb.
5. The á-series consonant េ is romanized by means of an apostrophe ('). េ k'ēk, េ châng'ër, េ rô'ër, េ vei, េ ang. In word-initial position before a vowel, ' may be omitted: េ ang.
6. The Khmer diacritical mark * appears only in two combinations: ** (examples: ពត bát, គឹម kʰpós) and ṭ. The diacritical mark ' appears only in the combination **. In the ā-series both ṭ and ** are romanized ā: ដាក khâk and ពៃដ្ឋ pûdpôr are both romanized châk. In the ō-series both ṭ and ** are romanized ea when followed by k, ng or h; otherwise, they are romanized oā: កូម ropeāk, មុន moāt, ស្មេ្ veângk, ស្មេ្ phoâpv.

7. The combination ṭ is romanized r before a consonant: គ្ thôrm. The combination ' is romanized as r before a consonant preceded by a shortened vocalic nucleus: គ្ koârr.

8. The symbol * in syllable-initial position is ignored in romanization: ស្ sâ, ស្មេ្ sâsâs.

   In syllable-final position * indicates that the consonant is vowelised, i.e. followed by ā in the ā-series, by ō in the ō-series: តម្ន្ támnâ, ប្ pômô.

9. The diacritical mark ** (which appears above characters and/or vowel markers which are not vocalized) is ignored in romanization: ប្ bûn, ឃូយ poâhî, ឃូយ phumi.

10. The independent character ុ is romanized either ō or ū. A reference source should be consulted where doubt arises.

**Other systems of romanization**

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

<table>
<thead>
<tr>
<th>UN</th>
<th>Provisional</th>
</tr>
</thead>
<tbody>
<tr>
<td>a, ā</td>
<td>= aa</td>
</tr>
<tr>
<td>-āh</td>
<td>= ah</td>
</tr>
<tr>
<td>-ām</td>
<td>= am</td>
</tr>
<tr>
<td>-āng</td>
<td>= ang</td>
</tr>
<tr>
<td>ā, ā</td>
<td>= a</td>
</tr>
<tr>
<td>aō</td>
<td>= ao</td>
</tr>
</tbody>
</table>

| ē   | = e         |
| -ēh | = -eh       |
| -ēn | = -en       |
| ē ō | = eae       |
| i, i| = i         |

\(^{A}\) The first romanization is applied in the ā-series, the second in the ō-series.

\(^{B}\) The romanization oeo is not verified.

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

The United Nations recommended system was approved in 1977 (III/11), based on the romanization of the Serbo-Croatian and Macedonian Cyrillic alphabets that had long been employed in official gazetteers and maps of Yugoslavia. The table was published as an annex to the resolution. The system is used in The former Yugoslav Republic of Macedonia and in international cartographic products.

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

### Romanization

| 1 | А а a | 9 | З ż z | 17 | Н н n | 25 | У у u |
| 2 | Б б b | 10 | С s sz | 18 | Ъ ъ nj | 26 | Ф ф f |
| 3 | В в v | 11 | И i i | 19 | О o o | 27 | Х х h |
| 4 | Г г g | 12 | Ј j j | 20 | П p p | 28 | Ц ц c |
| 5 | Д д d | 13 | К k k | 21 | Р r r | 29 | Ч ч č |
| 6 | Ѓ у d, g^ | 14 | Л l l | 22 | С s s | 30 | Ј ј u u |
| 7 | Е е e | 15 | Љ л l | 23 | Т t t | 31 | Щ щ š |
| 8 | Ж ж ž | 16 | М m m | 24 | Ъ ъ ć, ćk^ | 32 | '' |

^ The second variant is used before е and и.

Note. Cursive forms of some characters might be formed differently: Аа Бь Вь Гь Дь Ѓ у Ее Ж[жь] Зз Ии Јј Кк Лл Љ ль Мм Ни Њ[њь] Оо Пп Рр Сс Тт Ть Уу Фф Хх Цц Чч Џ[њ] Шш Щ[щь].

---


35
Malayalam

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

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

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 1970's a new orthography has been introduced and the use of many character combinations and ligatures has been simplified. Nevertheless, the user of the romanization table would have to recognize many ligatures not given in the original table. The system is mostly reversible but there exist some ambiguities in the romanization of vowels (independent vs. abbreviated characters) and consonants (ligatures vs. character sequences), particularly in the case of older orthography.

Romanization

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

I. Independent vowel characters

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>്a</td>
<td>3</td>
<td>ൈi</td>
<td>5</td>
<td>ൈu</td>
<td>7</td>
<td>്r</td>
</tr>
<tr>
<td>2</td>
<td>്ā</td>
<td>4</td>
<td>ോī</td>
<td>6</td>
<td>ോū</td>
<td>8</td>
<td>്ē</td>
</tr>
</tbody>
</table>

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

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 *</td>
<td>a</td>
<td>3 *</td>
<td>i</td>
<td>5 *</td>
<td>u&lt;sup&gt;A&lt;/sup&gt;</td>
<td>7 *</td>
<td>r&lt;sup&gt;C&lt;/sup&gt;</td>
</tr>
<tr>
<td>2 *</td>
<td>ā</td>
<td>4 *</td>
<td>ī</td>
<td>6 *</td>
<td>ū&lt;sup&gt;B&lt;/sup&gt;</td>
<td>8 *</td>
<td>ē</td>
</tr>
</tbody>
</table>

<sup>A</sup> Traditional script form: ് - ്q ku, q mu, etc. Exceptions in the older spelling: ്k ku (�ku, tuku), ്gu, ്chu, ്ju, ്nu, ്tu, ്nu (ʾna nu), q bu, q ru, q shu, q hu.

<sup>B</sup> Traditional script form: ് - q khi, q mū, etc. Exceptions in the older spelling: q kū, q gu, q chū, q jū, q nū, q tū, q nū, q bhū, q rū, q shū, q hū.

<sup>C</sup> Traditionally written in conjunction with the consonant character: ്k r.

<sup>D</sup> In the older spelling: ۓr.
III. Other symbols

1. m 2. h 3. v (A)

^ Indicates absence of the inherent short [a] vowel, e.g. ձ k, ռ y. Special forms of consonants with that symbol: ռռ h, ռռ n, ռ r, ռռ l, ռռ (a).

IV. Consonant characters

1 հ կ 2 ղ խ 3 ժ ժ 4 ժ ժ 5 ժ ժ

^ Special ligature: ռռ ի.

V. Ligatures

Adscript forms of some consonants

Example

- r - ջ rgga (in modern usage it is substituted by ռղ rgga)

* - r - ռղ kra (older style: ռ)

* - v - ռղ tya

* - v - ռղ sva

Other ligatures (the list is not complete)

kka kta kta ksha kha gga gda gddha gna gma gla nka

ntha nda nma tta thia tna tbha tma tsa thia ddha
ta nda nda nma nma nra ptia pma pna ppa pla psa

bda bdha bba bla mma mpa mma mla yka ykka yta yttta
ypa yma yma yva lka lka lpa lma lla vva shca shla shha

shka shtha shtha shna shpa shpha shma ska sta stha sna

spa spha sma sla ssa hna hma hla llia lka lcha lhta

lm* lpa*

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

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

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

<table>
<thead>
<tr>
<th>UN</th>
<th>State of Kerala</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td>zhe</td>
</tr>
<tr>
<td>sh</td>
<td>s</td>
</tr>
<tr>
<td>t</td>
<td>th</td>
</tr>
<tr>
<td>t</td>
<td>t or d</td>
</tr>
</tbody>
</table>

Marathi

The United Nations recommended system was approved in 1972 (II/11) and amended in 1977 (III/12), based on a report prepared by D. N. Sharma. The tables and their corrections were published in volume II of the conference reports21.

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

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

Romanization

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

1. Additional characters (absent in Hindi):
   छ | a
   ँ आ (abbreviated: के kā)

2. Different shapes of some characters and ligatures:

<table>
<thead>
<tr>
<th>Hindi</th>
<th>Marathi</th>
<th>Roman</th>
</tr>
</thead>
<tbody>
<tr>
<td>क</td>
<td>का</td>
<td>kha</td>
</tr>
<tr>
<td>ल</td>
<td>ला</td>
<td>la</td>
</tr>
<tr>
<td>कत</td>
<td>कता</td>
<td>kta</td>
</tr>
<tr>
<td>कक</td>
<td>kka</td>
<td></td>
</tr>
</tbody>
</table>

3. Ligatures may be formed with r as the first component: भ rya, भ rha.
4. Typical are ligatures ending in h: घ nh, घ nh, घ lh, घ vh.

5. The following independent characters are sometimes used: द, ओ औ, ऐ ए, ऐ ऐ.

**Mongolian**

*(in China)*

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

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

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

**Romanization**

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

**Notes**

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

---

Nepali

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

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

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

**Romanization**

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

1. In normal spelling र and र are not marked, instead dotless द द, द दha are used to denote these sounds.
2. Ligatures may be formed with र as the first component: र रha.

**Other systems of romanization**

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

<table>
<thead>
<tr>
<th>UN</th>
<th>BGN/PCGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>रि</td>
<td>रि</td>
</tr>
<tr>
<td>रि</td>
<td>रि</td>
</tr>
<tr>
<td>रि</td>
<td>रि</td>
</tr>
<tr>
<td>रि</td>
<td>रि</td>
</tr>
</tbody>
</table>

^ The second variant is used in word-medial position.

Oriya

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

---


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

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

**Romanization**

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

### I. Independent vowel characters

| 1 | Ḫ | a | 3 | Ḃ | i | 5 | ḇ | u | 7 | ꜁ | Ḩ | 8 | ə | e | 10 | ꜀ | o<sup>B</sup> |
| 2 | ꜁ | ā | 4 | ṭ | ĩ | 6 | ṭ | ū | 9 | ḍ | ai | 11 | ꜁ | au |

<sup>A</sup> Rarely used characters that are positioned between characters 7 and 8: ꜁ Ḩ, ə ṭ.<br><sup>B</sup> Special form: ꜁ va (not given in the original table).

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

<table>
<thead>
<tr>
<th>1 *</th>
<th>a</th>
<th>3 *</th>
<th>i&lt;sup&gt;A&lt;/sup&gt;</th>
<th>5</th>
<th>ṭ</th>
<th>7 *</th>
<th>Ḩ</th>
<th>8 *</th>
<th>e</th>
<th>10 *</th>
<th>o</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 *</td>
<td>ā</td>
<td>4 *</td>
<td>ĩ</td>
<td>6 *</td>
<td>ū</td>
<td>9 *</td>
<td>꜀</td>
<td>ai</td>
<td>11 *</td>
<td>Ḩ ( ꜁ ) au</td>
<td></td>
</tr>
</tbody>
</table>

<sup>A</sup> Variations: ꜁ thi, ꜁ dhi.

### III. Other symbols

| 1 * | m | 2 * | Ṁ | 3 | ꜀ | ḫ | 4 | * |<sup>(A)</sup> |

<sup>A</sup> Pronunciation without a vowel: ꜀ k.

### IV. Consonant characters

| 1 | ꜀ | ka | 6 | ꜀ | cha | 11 | ꜀ | ṭa | 16 | ꜀ | ta | 21 | ꜀ | pa | 26 | ꜀ | ya<sup>B</sup> | 30 | ꜀ | sha |
| 2 | ꜀ | kha | 7 | ꜀ | chha | 12 | ꜀ | ṭha | 17 | ꜀ | tha | 22 | ꜀ | pha | 27 | ꜀ | ra | 31 | ꜀ | śha |
| 3 | ꜀ | ga | 8 | ꜀ | ja | 13 | ꜀ | ḍa<sup>A</sup> | 18 | ꜀ | da | 23 | ꜀ | ba | 28 | ꜀ | la | 32 | ꜀ | sa |
| 4 | ꜀ | gha | 9 | ꜀ | jha | 14 | ꜀ | ḍha<sup>A</sup> | 19 | ꜀ | dha | 24 | ꜀ | bha | 29 | ꜀ | ḫa | 33 | ꜀ | ha |
| 5 | ꜀ | ṇa | 10 | ꜀ | ṇa | 15 | ꜀ | ṇa | 20 | ꜀ | na | 25 | ꜀ | ma | 34 | ꜀ | kṣha |

<sup>A</sup> Dotted variants of the characters: ꜀ ōa, ꜀ Ḥha.<br><sup>B</sup> Special form: ꜀ ya.
V. Subscript and adscript consonant characters
Character numbers correspond to those in Table IV.

1. k 6. ch 11. t 16. t 21. p 26. y
2. kh 7. chh 12. th 17. th 22. ph 27. r
3. g 8. j 13. d 18. d 23. b 28. l
5. h 20. n 25. m

VI. Ligatures (the list is not complete)
कका क्ता क्ला क्षा क्ष्वा क्सा क्वा न्का न्क्वा न्खा न्गा न्ग्वा च्वा
च्हच्हा ज्वा ज्व्वा ङ्वा ङ्ख्वा ङ्ख्व्वा ङ्च्वा ङ्च्व्वा ङ्ञ्वा ङ्ञ्व्वा ङ्ठ्वा
ङ्ङा ङ्ङ्वा ङ्ङ्व्वा ङ्ङ्व्व्वा ङ्ङ्ङा ङ्ङ्ङ्वा ङ्ङ्ङ्व्वा ङ्ङ्ङ्व्व्वा
म्बा म्ब्वा म्का म्पा म्प्वा म्प्व्वा म्प्व्व्वा म्प्व्व्व्वा
ष्वा ष्व्वा ष्व्व्वा ष्व्व्व्वा ष्व्व्व्व्वा ष्व्व्व्व्व्वा
Combinations with r as the first component are written ळ, e.g. ळ्रका.

Other systems of romanization

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

Persian

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

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

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

---

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

**Romanization**

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

<table>
<thead>
<tr>
<th>1</th>
<th>1</th>
<th>A</th>
<th>9</th>
<th>•</th>
<th>17</th>
<th>ص</th>
<th>25</th>
<th>l</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>ب</td>
<td>b</td>
<td>10</td>
<td>•</td>
<td>18</td>
<td>ض</td>
<td>26</td>
<td>g</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>p</td>
<td>p</td>
<td>11</td>
<td>•</td>
<td>19</td>
<td>ط</td>
<td>27</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td>4</td>
<td>ت</td>
<td>t</td>
<td>12</td>
<td>•</td>
<td>20</td>
<td>ض</td>
<td>28</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ث</td>
<td>s</td>
<td>13</td>
<td>•</td>
<td>21</td>
<td>ع</td>
<td>29</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ج</td>
<td>j</td>
<td>14</td>
<td>•</td>
<td>22</td>
<td>غ</td>
<td>30</td>
<td>w</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>ج</td>
<td>ch</td>
<td>15</td>
<td>•</td>
<td>23</td>
<td>ف</td>
<td>31</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>ه</td>
<td>h</td>
<td>16</td>
<td>•</td>
<td>24</td>
<td>ق</td>
<td>32</td>
<td>y</td>
<td></td>
</tr>
</tbody>
</table>

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

**Vowel characters and diacritical marks (∗ stands for any consonant)**

| 1 | • | a | 5 | • | 6 | • | א, א | 10 | • | 11 | • | (C) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | • | e | 6 | • | 7 | • | א | 11 | • | (B) |
| 3 | • | o | 7 | • | 8 | • | א | 12 | • |

* A Word-initially.

B Not romanized; marks absence of the vowel.

CDoubling of the consonant letter.

D After a consonant (excl. -*ah*).

E After a vowel (see also note 2).

**Notes**

1. The adjectival ending of Arabic origin *-ye* in Persian is romanized *-yeh*. In romanizing the definite article the same rules of assimilation of consonants are applied as in Arabic, e.g. زين الدين Zeyn od Dīn.

2. The relational suffix (ezāfeh) *-e* is usually not expressed in Persian writing after a consonant. After final *l* or *r* it is written with ī, e.g. پای آب Pā-ye Āb. After final *ی* and *،* it is expressed by writing *hamzeh* over the character: دهانه ممبر Dahāneh-ye Mambār.

3. To point Persian vowels two systems are in use that are separated by a slash in the table. The first system is a Persian one while the other adheres to the Arabic tradition. In normal spelling vowel points are not used.

**Other systems of romanization**

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

<table>
<thead>
<tr>
<th>UN</th>
<th>ISO</th>
</tr>
</thead>
<tbody>
<tr>
<td>ā</td>
<td>ā</td>
</tr>
<tr>
<td>ch</td>
<td>ċ*</td>
</tr>
<tr>
<td>gh</td>
<td>ġ</td>
</tr>
<tr>
<td>h</td>
<td>ħ</td>
</tr>
<tr>
<td>sh</td>
<td>ʂ</td>
</tr>
<tr>
<td>zh</td>
<td>Ŗ*</td>
</tr>
<tr>
<td>t</td>
<td>š</td>
</tr>
<tr>
<td>u</td>
<td>ū</td>
</tr>
<tr>
<td>z</td>
<td>ź</td>
</tr>
</tbody>
</table>

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

**Punjabi**

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

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

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

**Romanization**

\(^{1}\) Th (u)\(^\text{A}\)

<table>
<thead>
<tr>
<th>1 Th</th>
<th>6 अ</th>
<th>11 च</th>
<th>16 ट</th>
<th>21 त</th>
<th>26 प</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 ह</td>
<td>7 आ</td>
<td>12 ह</td>
<td>17 थ</td>
<td>22 ठ</td>
<td>27 फ</td>
</tr>
<tr>
<td>3 द</td>
<td>8 ग</td>
<td>13 ज</td>
<td>18 ढ</td>
<td>23 द</td>
<td>28 ब</td>
</tr>
<tr>
<td>4 घ</td>
<td>9 घ</td>
<td>14 घ</td>
<td>19 ढ</td>
<td>24 घ</td>
<td>29 ः</td>
</tr>
<tr>
<td>5 ह</td>
<td>10 रा</td>
<td>15 रा</td>
<td>20 रा</td>
<td>25 रा</td>
<td>30 म</td>
</tr>
</tbody>
</table>

\(^{2}\) These characters are used in combination with abbreviated vowel characters: ऋ u, ॠ ू, ॠ a, ३ d̪̪̪̪̪̈

\(^{3}\) Dotted variants of the characters: ट ha, ठ kha, ढ ga, ॠ za, ॠ fa.

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

\[
\begin{array}{cccccccccc}
1 & a & 3 & ñ & i & 5 & u & 7 & e & 9 & o & 11 & m^a \\
2 & ä & 4 & ñ & i & 6 & ñ & 8 & ai & 10 & au & 12 & m^b \\
\end{array}
\]

^ Used if it is preceded by short vowels (a, i, u) and ñ, excluding the independent vowel character ñ (u, ñ), e.g. नेत samag, मणि singh, बुि bundā, बुि bund, मिना irjan.

^ Used in all other occasions, e.g. घट jhorākā.

^ Marks doubling of the following consonant: टिक्का nikkā, लाभ्य labhbhpā, कुत्ता kuttā, हिस्सा hissā, लाम्बा lammā.

Subscript consonant characters

\[
\begin{array}{cccc}
1 & -h & 2 & -v & 3 & -r & 4 & -y \\
\end{array}
\]

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

Other systems of romanization

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

Russian

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

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

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

---

Romanization

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

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

Other systems of romanization

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

e (e)  e, ye^A  й (j)  у  ч (č)  ch  э (ē)  e
ě (ě)  ě, ye^A  х (h)  kh  ш (š)  sh  ю (ju)  yu
ž (ž)  zh  ц (c)  ts  щ (śč)  shch  я (ja)  ya

^A Initially, after the vowel characters and after й, ъ, and ъ.

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

Ш (ŚČ)  Š  Ю (ju)  Ù  Я (ja)  ā

Serbo-Croatian Cyrillic

The United Nations recommended system was approved in 1977 (III/11), based on the Roman (Croatian) equivalent of the Serbo-Croatian Cyrillic alphabet that had long been employed in official gazetteers and maps of Yugoslavia. The table was published as an annex to the resolution28.

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

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

---

Romanization

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

Note. Cursive forms of some characters might be formed differently: Аа Бб Вв Гг Дд Ёж Ёе Жж Ёж
Зз Ии Йй Іи Іп Ьљ Ьь Мм Нн Ьн Оо Іп Рр Сс Тт Ьн Уу Фф Хх Цц Чч Цч Шш.

Tamil

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

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

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

Romanization

I. Independent vowel characters

1 ᡶ a 3 غضب i 5 ᡶ u 7 е e 9 诰 ai 11 ᡬ o
2 ᡦ ā 4 ᡞ ī 6 ᡷ ū 8 ը e 10 ᡶ ō 12 诰 au

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

1 ∗ a 3 ܡ iB 5 ܙ uD 7 ܓ e 9 ܓ aiF 11 ܓ o 13 ܙ hG
2 ∗ ā 4 ∗ īC 6 ∗ ūD 8 ġ e 10 ġ o 12 ġ au 16 ∗ (H)

A Exceptions in the older and Classical literature: ᡨ nā, ᡧ rā, ᡨ pā.
B Variation: тā, ти.

etc.
C Variation: ka ti.
D Variations: ku ni nu, chu, ti nu, tu, nu, tu, nu, pu, mu, ju, ru, lu, etc.
E Variations: ku, ni, chi, ti nu, tu, ntu, nü, ti, pü, mü, etc.
F Exceptions in the older and Classical literature: lai, lai, nai, nai.
G Absent in the original table.
H Pronunciation without a vowel: k.

III. Consonant characters

1 ka la ta na pa ya ja ha
2 na na na na ma ra la sa ha
3 cha ta sa la ra ssa ra
4 fi na va na sa sa

A Characters borrowed from the Grantha script.
B Special syllable character: shri.

Other systems of romanization

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

Telugu

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

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

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

Romanization

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

---

I. Independent vowel characters

1 ə a 4 ₍ i 7 ṭu g 9 ə e 12 ə o
2 a̯ a 5 ə u 8 ṭu g 10 ə e 13 ə o
3 ə i 6 ə u 11 ə ai 14 ə au

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

1 a* a 4 i* i 7 u* u 8 e* e 11 o* o 14 m* m 17 k* k
2 a 5 u* u 9 e* e 12 o* o 15 m* m 16 h
3 * i 6 u* u 10 * ai 13 * au 16 * h

A Variations: ə chi, ə ni, ə fi, ə li, etc.
B Variations: ə ju, ə pu, ə vu, etc.
C Variations: ə mā, ə mā yā.
D Variations: ə mā, ə mā yo.
E Signified historic nasal sound, now obsolete.
F End-of-syllable mark (i.e., a consonant without a vowel): ᵆ ə Rādgāpur.

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

kā kā ki kā ku kā kā ke kā kā kā kā kā kā kā kā

ka kā ki kā ku kā kā ke kā kā kā kā kā kā kā kā

III. Consonant characters

1 k kā 6 ch chā 11 3 ta 16 e ta 21 ṭa pa 26 ya 31 ṭa
2 k kha 7 3 chh chha 12 e tha 17 ṭa tha 22 3 pha 27 3 ra 32 3 sa
3 g ga 8 3 ja 13 3 da 18 3 da 23 3 ba 28 3 la 33 3 ha
4 ḍ gha 9 ḍ ja 14 3 dha 19 3 dha 24 3 bha 29 3 va 34 3 la
5 ḍ 3 na 10 3 na 15 3 na 20 3 na 25 3 ma 30 3 sha

A Variant: ə.

IV. Subscript consonant characters

1 k k 3 n 5 y 7 1
2 t t 4 m 6 r 8 5

Most other subscript consonant characters are recognizable modifications of the main character, e.g. ə chcha, ə tta, ə ppa.

Other systems of romanization

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

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

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

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

Romanization

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

1 ก (k) 8 ธ ch (t) 14 ฉ d (t) 20 ฑ d (t) 26 บ b (p) 34 ย y 42 ｽ s (t)
2 ข kh (k) 9 ธ ch (t) 15 ฦ t (t) 21 ฅ t (t) 27 ฏ p (p) 35 ร r (r) 43 ส s (t)
3 ฃ kh (k) 10 ธ ch (t) 16 ฏ th (t) 22 ฅ th (t) 28 ฑ th (t) 36 ฎ ruBC 44 ｽ s (t)
4 ฅ kh (k) 11 ษ s (t) 17 ฑ th,d(t) 23 ฅ th (t)A 29 ฑ f (p) 37 ฎ ruB 45 ห hD
5 ฅ kh (k) 12 ว ch (t) 18 ฏ th (t) 24 ฅ th (t) 30 ฑ th (t) 38 ฅ l (n) 46 ใ l (n)
6 ฅ kh (k) 13 ฏ y (n) 19 ฅ n (n) 25 ฅ n (n) 31 ฅ f (p) 39 ฅ lJ luB 47 ฅ oE
7 ฅ ng(ng)

A Combination ฅษ is romanized ｓ if so pronounced.
B These Thai characters are traditionally considered as vowel characters; for their revised romanizations see below.
C Also romanized ริ and รี.
D See note 7.
E See note 8.

**Vocalic nuclei** (• stands for any consonant character)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>ə</em></td>
<td><em>ə</em></td>
<td><em>ə</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>e</td>
<td>a</td>
<td>e</td>
<td>a</td>
<td>e</td>
<td>i</td>
<td>u</td>
<td>o</td>
<td>u</td>
<td>o</td>
<td>a</td>
<td>e</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>ə*</td>
<td><em>ə</em></td>
<td><em>ə</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
</tr>
<tr>
<td></td>
<td>ia</td>
<td>ia</td>
<td>ua</td>
<td>ua</td>
<td>ua</td>
<td>ua</td>
<td>ua</td>
<td>ai</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>oe</td>
<td>oe</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td><em>ə</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
<td><em>ɪ</em></td>
</tr>
<tr>
<td></td>
<td>ao</td>
<td>ui</td>
<td>oi</td>
<td>oi</td>
<td>oei</td>
<td>uai*</td>
<td>uai</td>
<td>iu*</td>
<td>eo</td>
<td>eo</td>
<td>aeo</td>
<td>aeo</td>
<td>ieo*</td>
</tr>
</tbody>
</table>

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

**Notes**

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

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

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

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

---

**Notes** 1-2 have been taken from the original paper (1967), 3-4 from the revised version (1997), and 5-10 have been adapted from the BGN/PCGN 1970 romanization system, to cover the most essential features of the Thai script.
5. A Thai consonant character with no accompanying vowel mark or diacritical mark may be romanized with the vowel letter a or o: ปทом pathom, โพธ phop.

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

7. อะ at the beginning of a syllable and serving as a tone mark should not be romanized: อะแวก wo, อะแวกลูป lup; otherwise, อะ should be romanized อะแวก huai, อะแวก wo. ะ

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

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

10. The vowel shortening mark ″ and the tone marks ′, ″″, and ″ should not be romanized: ง กอ, งิห่อ yiho, กระตอ kratop.

**Other systems of romanization**

The 1997 revised version\(^{33}\) of the official Thai romanization system gives different Roman equivalents to the following Thai characters:

\begin{align*}
\text{ง} & \text{ eve, ri, roe} & \text{ิ} & \text{ lue} & \text{ี} & \text{ue} & \text{ื} & \text{uea} & \text{ื} & \text{uei} & \text{ื} & \text{ia}
\end{align*}

\begin{align*}
\text{ง} & \text{ eve} & \text{ิ} & \text{ lue} & \text{ี} & \text{ue} & \text{ื} & \text{uea} & \text{ื} & \text{i o}
\end{align*}

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

---


Tibetan

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

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

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

**Romanization**

**Consonant characters**

<table>
<thead>
<tr>
<th>1</th>
<th>ṝ</th>
<th>ga</th>
<th>5</th>
<th>ṝ</th>
<th>jā</th>
<th>9</th>
<th>ṝ</th>
<th>da</th>
<th>13</th>
<th>ṝ</th>
<th>ba</th>
<th>17</th>
<th>ṝ</th>
<th>za</th>
<th>21</th>
<th>ṝ</th>
<th>xa</th>
<th>26</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>ṝ</td>
<td>ka</td>
<td>6</td>
<td>ṝ</td>
<td>qa</td>
<td>10</td>
<td>ṝ</td>
<td>ta</td>
<td>14</td>
<td>ṝ</td>
<td>pa</td>
<td>18</td>
<td>ṝ</td>
<td>ca</td>
<td>22</td>
<td>ṝ</td>
<td>sa</td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td>ṝ</td>
<td>ka,ga(^{\wedge})</td>
<td>7</td>
<td>ṝ</td>
<td>qa,ja(^{\wedge})</td>
<td>11</td>
<td>ṝ</td>
<td>ta,da(^{\wedge})</td>
<td>15</td>
<td>ṝ</td>
<td>pa,ba(^{\wedge})</td>
<td>19</td>
<td>ṝ</td>
<td>ca,za(^{\wedge})</td>
<td>23</td>
<td>ṝ</td>
<td>a</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>ṝ</td>
<td>nga</td>
<td>8</td>
<td>ṝ</td>
<td>nya</td>
<td>12</td>
<td>ṝ</td>
<td>na</td>
<td>16</td>
<td>ṝ</td>
<td>ma</td>
<td>20</td>
<td>ṝ</td>
<td>wa</td>
<td>24</td>
<td>ṝ</td>
<td>ya</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>ṝ</td>
<td>ra</td>
<td>30</td>
<td>ṝ</td>
<td>na</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{\wedge}\) If the character is not accompanied by a prefixed character or a superscripted consonant, it is read as aspirated and the first romanization equivalent is used; otherwise it is read as nonaspirated and the second equivalent is used.

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

<table>
<thead>
<tr>
<th>1</th>
<th>ṝ</th>
<th>gya</th>
<th>3</th>
<th>ṝ</th>
<th>kya, gya(^{\wedge})</th>
<th>5</th>
<th>ṝ</th>
<th>cha</th>
<th>7</th>
<th>ṝ</th>
<th>sha</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>ṝ</td>
<td>kya</td>
<td>4</td>
<td>ṝ</td>
<td>zha</td>
<td>6</td>
<td>ṝ</td>
<td>cha, zha(^{\wedge})</td>
<td>8</td>
<td>ṝ</td>
<td>lha</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{\wedge}\) If the combination is not accompanied by a prefixed character or a superscripted consonant, it is read as aspirated and the first romanization equivalent is used; otherwise it is read as nonaspirated and the second equivalent is used.

---

Vowel characters (η stands for any consonant character)

η a ə i ə̲ u ə̲ e ə̲ o

Other symbols
- syllable separation character, e.g. ་ ། Nag Qu.

Syllable endings

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

Notes
1. Prefixed consonant characters η(k), ζ(t), ξ(p), η(m), ξ(a), and superscripted consonant characters η(r), ζ(I), ξ(s) remain mostly unpronounced and are not reflected in the romanization.
2. Subscript characters ι(w), ι(y), ι(r), ι(I) are pronounced only in certain combinations (see also the consonant combinations at the beginning of syllables).
3. A number of homonyms exist in the Tibetan language and they are not all listed here.
4. The characters η and η are treated as without initials when transcribed with Pinyin.
5. When the prefixed characters η and η or the superscripted consonant η give the preceding...
syllable an additional nasal or make the vowel in it nasalized, it is spelt according to the actual pronunciation. The nasalized vowels (a), (o), (u) are represented by the finals an, on, un in the Pinyin system.

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

Uighur

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

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

Romanization

| 1 | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v |
| 25 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 |

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

---

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

**Urdu**

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

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

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

**Romanization**

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

1 a, ā, ā' 11 d 21 z 31 m
2 b 12 ẓ 22 t 32 n
3 p 13 ẓ 23 z 33 w, w
3 p 13 ẓ 23 z 33 w, w
4 t 14 r 24 g, gh 34 h, ḥ
5 ẓ 15 r 25 gh 35 h
6 s 16 z 26 f 36 y
7 j 17 ẓ 27 q 37 e, ai
8 ch 18 s 28 k
9 h 19 sh 29 g
10 kh 20 s 30 l

A If preceded by short a, it is romanized ‘ā, e.g. مَعَمَل Māmul.
B When j is imperceptible, e.g. in a few words of Persian origin when preceded by خ (kh).
C Word-finally after a short vowel.
D Marks aspiration of consonants.

---


The character  is used only word-finally.

Note. The Arabic  has two distinct shapes in Urdu:  (w) and  (w). The latter denotes aspiration after consonants b, ch, d, d̆, j, k, p, r̆, t, t̆.

<table>
<thead>
<tr>
<th>Vowels</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>٢</td>
<td>٣</td>
<td>٤</td>
<td>٥</td>
</tr>
<tr>
<td>٦</td>
<td>٧</td>
<td>٨</td>
<td>٩</td>
<td>١٠</td>
</tr>
<tr>
<td>١١</td>
<td>١٢</td>
<td>١٣</td>
<td>١٤</td>
<td>١٥</td>
</tr>
<tr>
<td>١٦</td>
<td>(C)</td>
<td>(B)</td>
<td>(hyphen)</td>
<td>(hyphen)</td>
</tr>
</tbody>
</table>

The character  is used word-finally.

B. Not romanized; marks absence of the vowel.

C. Doubling of the consonant letter.

**Other systems of romanization**

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

<table>
<thead>
<tr>
<th>UN</th>
<th>Hunterian</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ah (w)</td>
<td>a (\approx a^A)</td>
</tr>
<tr>
<td>ā</td>
<td>ā, a^A</td>
</tr>
<tr>
<td>d̆</td>
<td>d̆</td>
</tr>
<tr>
<td>dh</td>
<td>dh</td>
</tr>
<tr>
<td>gh</td>
<td>gh</td>
</tr>
<tr>
<td>h̆</td>
<td>h̆</td>
</tr>
<tr>
<td>t̆</td>
<td>t̆</td>
</tr>
<tr>
<td>n̆</td>
<td>n̆</td>
</tr>
<tr>
<td>v̆</td>
<td>v̆</td>
</tr>
<tr>
<td>ŭ</td>
<td>ŭ</td>
</tr>
<tr>
<td>y̆</td>
<td>y̆</td>
</tr>
</tbody>
</table>

A. Word-finally.

B. Exceptionally, e.g. before i.
OTHER LANGUAGES/SCRIPTS

Armenian

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

Systems of romanization

The BGN/PCGN 1981 System gives the following romanization:

1. U \( \mu \) a
2. f \( \phi \) b
3. q \( \eta \) g
4. \( \eta \) d
5. t \( \phi \) e, ye\(^A\)
6. Q q z
7. L l t
8. L l y
9. \( \rho \) p t'
10. \( \phi \) zh
11. b \( \mu \) i
12. L l l
13. h h kh
14. U d ts
15. U u k
16. h h
17. 2 a dz
18. \( \eta \) n gh
19. A k ch
20. U u m
21. 3 j y
22. U u n
23. 3 \( \gamma \) sh
24. \( \eta \) n o, vo\(^B\)
25. 2 \( \gamma \) ch'
26. \( \eta \) u p
27. 2 g j
28. \( \eta \) n rr
29. U u s
30. U u v
31. S u p t
32. f n r
33. g ts'
34. \( \eta \) m u
35. Q p'
36. 2 p k'
37. b t u ev, yev\(^C\)
38. o o o
39. 3 $ f
40. \( \phi \) m u

\(^A\) ye is romanized initially and after the vowel characters \( u, b, t, \Omega, h, n, m, \) and o.

\(^B\) vo is romanized initially except in the word \( n \) which should be romanized ov.

\(^C\) yev is romanized initially, in isolation, and after the vowel characters \( u, b, t, \Omega, h, n, m, \) and o.

Note. In newer sources the character 37 is used in the lower case form only; in the upper case it is written as a combination of characters 5 and 30 (b t l), cf. b t l \( \mu \) YEREVAN, b t l \( \mu \) Yerevan.

Also there is some indication that in 1997 the alphabet was modified so that the character b l w was restored as an independent character. According to the same source the end of the alphabet would read as follows (starting from the point of difference with the above table):

34. 3 b l w
35. 3 \( \phi \) p'
36. 3 p k'
37. 3 o o o
38. 3 o o o
39. 3 $ f
40. \( \Omega \) n u

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

b (e/ye) e
t (e) e
t (zh) z
t (kh) x
t (e) e
t (zh) z
t (kh) x
t (e) e
t (zh) z
t (kh) x
Burmese

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

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

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

*System of romanization*

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

**Consonant characters**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>
| ka | ka | ga | ga | nga | **\(^{A}\)**
| 6 | 7 | 8 | 9 | 10 |
| sa | sa | za | za | nya | **\(^{B}\)**
| 11 | 12 | 13 | 14 | 15 |
| ta | da | da | da | na | **\(^{C}\)**
| 16 | 17 | 18 | 19 | 20 |
| ta | da | da | da | ma | **\(^{D}\)**
| 19 | 20 | 21 | 22 | 23 |
| pa | pa | pa | ya | ba | **\(^{E}\)**
| 24 | 25 | 26 | 27 | 28 |
| ya | la | la | wa | wa |
| 28 | 29 | 30 | 31 | 32 |
| ha | la | la | wa | ya |

\(^{A}\) See note 1.

\(^{B}\) Syllable-final variant character: \(ʒ\).

\(^{C}\) \(ŋ\) is romanized *sha*.

\(^{D}\) Syllable-final variant character: **\(^{\circ}\)**.

\(^{E}\) See note 3.

**Independent vowel characters** (see also note 4)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>i</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>e</td>
<td>e</td>
<td>aw</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>an</td>
<td>à</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
</tr>
</tbody>
</table>

**Other symbols** (\(\circ\) stands for any consonant character)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>ʃ</td>
<td>ʃ</td>
<td>ʃ</td>
<td>ʃ</td>
<td>ʃ</td>
<td>ʃ</td>
<td>ʃ</td>
<td>ʃ</td>
<td>ʃ</td>
</tr>
</tbody>
</table>

\(^{A}\) Exceptional romanization: ʃ  *cha* (see also note 1).
**Syllable endings** (∞ stands for any initial consonant character)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>et</td>
<td>eik</td>
<td>òk</td>
<td>ut</td>
<td>aik</td>
<td>auk</td>
<td>et</td>
<td>in</td>
<td>ein</td>
<td>aing</td>
<td>ang</td>
<td>at</td>
<td>it</td>
<td>âk</td>
<td>i, in, e'</td>
<td>at</td>
</tr>
<tr>
<td>33</td>
<td>32</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td>40</td>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td>47</td>
</tr>
<tr>
<td>ein</td>
<td>an</td>
<td>ein</td>
<td>un</td>
<td>wut</td>
<td>wut</td>
<td>wun</td>
<td>eik</td>
<td>in</td>
<td>at</td>
<td>eik</td>
<td>òk</td>
<td>ut</td>
<td>wut</td>
<td>an</td>
<td>eik</td>
</tr>
<tr>
<td>49</td>
<td>51</td>
<td>53</td>
<td>55</td>
<td>56</td>
<td>58</td>
<td>59</td>
<td>60</td>
<td>61</td>
<td>62</td>
<td>63</td>
<td>64</td>
<td>65</td>
<td>66</td>
<td>67</td>
<td>68</td>
</tr>
<tr>
<td>un</td>
<td>è</td>
<td>an</td>
<td>eik</td>
<td>un</td>
<td>ein</td>
<td>wun</td>
<td>an</td>
<td>at</td>
<td>at</td>
<td>eik</td>
<td>ök</td>
<td>un</td>
<td>wun</td>
<td>an</td>
<td>un</td>
</tr>
</tbody>
</table>

---

**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: ꜳa ꜳadama.
3. ꜳ 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: ꜳa ꜳa ꜳu, ꜳadadad Thibin-aing.
4. The independent vowel characters should be romanized with a hyphen at the beginning of medial and final syllables: ꜳadadad kye-e.
5. In consonant clusters the first component is romanized according to the main table, not according to the table of syllable endings: ꜳa ꜳmissa
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: ꜳadadad in-gut (but ꜳadadad shwengan).
7. Tone marks, and : should not be romanized.

Byelorussian

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

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

Systems of romanization

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

<table>
<thead>
<tr>
<th>1</th>
<th>А а а</th>
<th>10</th>
<th>І і і</th>
<th>19</th>
<th>С с s</th>
<th>28</th>
<th>'</th>
<th>(not romanized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Б б б</td>
<td>11</td>
<td>Й й й</td>
<td>20</td>
<td>Г т т</td>
<td>29</td>
<td>БЫ y</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>В в v</td>
<td>12</td>
<td>К к k</td>
<td>21</td>
<td>Г Я y</td>
<td>30</td>
<td>Бь '</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Г г h</td>
<td>13</td>
<td>Л л l</td>
<td>22</td>
<td>І я i</td>
<td>31</td>
<td>Е e</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Д д d</td>
<td>14</td>
<td>М м m</td>
<td>23</td>
<td>Ф ф f</td>
<td>32</td>
<td>Ю ю ju, ju</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Е е je, ie</td>
<td>15</td>
<td>Н н n</td>
<td>24</td>
<td>Ц ц ch</td>
<td>33</td>
<td>Я я ja, ja</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Ё Ё jоA, io</td>
<td>16</td>
<td>О о о</td>
<td>25</td>
<td>Ч ч е</td>
<td>34</td>
<td>Бь бь Э э Юю Яя</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Ж ж ž</td>
<td>17</td>
<td>П п р</td>
<td>26</td>
<td>Ш ш ш</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>З з z</td>
<td>18</td>
<td>Р р r</td>
<td>27</td>
<td>Щ щ ш</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>І і і</td>
<td>19</td>
<td>С с s</td>
<td>28</td>
<td>І я i</td>
<td>33</td>
<td>Я я ja, ja</td>
<td></td>
</tr>
</tbody>
</table>

A Initially, after vowels, apostrophe, Ь and Ь.
B After consonants.

Note. Cursive forms of some characters might be formed differently: Аа Бб Вв Гг Іг Ее Ёё Жж Зз Іі Ій Кк Лл Мм Нн Оо Пп Рр Сс Тт Уу Іў Фф Хх Цц Чч Шш Ьь Ьь Ээ Юю Яя.

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

<table>
<thead>
<tr>
<th>1</th>
<th>е (je/je)</th>
<th>ye</th>
<th>1</th>
<th>й (j)</th>
<th>y</th>
<th>2</th>
<th>ц (с)</th>
<th>ts</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>е (je/je)</td>
<td>yo</td>
<td>1</td>
<td>й (i)</td>
<td>w</td>
<td>2</td>
<td>ч (c)</td>
<td>ch</td>
</tr>
<tr>
<td>3</td>
<td>ж (ж)</td>
<td>zh</td>
<td>1</td>
<td>х (ch)</td>
<td>kh</td>
<td>2</td>
<td>ш (ш)</td>
<td>sh</td>
</tr>
<tr>
<td>4</td>
<td>Я я ja, ja</td>
<td>1</td>
<td>&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

59 Published in "National'ny Reestr Pravovyh Aktov Respubliki Belarus" (National Registry of Legal Acts, Republic of Belarus) No. 3, 11 January 2001
Dzongkha

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

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

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

**System of romanization**

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

**Consonant characters**

| 1 न् | ka       | 9 त | ta       | 17 त्स | tsa     | 25 ऩ | ra       |
| 2 खा | kha      | 10 थ | tha      | 18 थ्स | tsha    | 26 ल | la ('la) |
| 3 ग'ा (ga) | 11 ध | d'a (da) | 19 द्स | dza     | 27 ल | sha      |
| 4 ङ | nga ('nga) | 12 ण | na ('na) | 20 व | wa      | 28 ण | sa       |
| 5 ऋ | ca       | 13 प | pa       | 21 श | zh'a (zha) | 29 उ | ha       |
| 6 ट | cha      | 14 फ | pha      | 22 झ | z'a (za) | 30 भ | 'a       |
| 7 ढ | j'a (ja) | 15 ब'ा (ba) | 23 ं | a       |
| 8 ञ | nya ('nya) | 16 म | ma ('ma) | 24 य | ya       |

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

**Syllable-initial consonant combinations** (the list is not complete)

Where there are multiple romanizations, the romanization in parentheses is to be generally used if the character is accompanied by a prefixed character or a superscripted consonant.

---

⁴⁰ Adapted from *Dzongkha* by George van Driem with the collaboration of Karma Tshering of Gaselô. Research School CNWS, School of Asian, African, and Amerindian Studies. Leiden, The Netherlands 1998.
| 1 | k, c<sup>A</sup> | 6 | j’ (j), bj’ (bj) | 11 | tr | 16 | dr |
| 2 | kh, ch<sup>A</sup> | 7 | ny | 12 | thr | 17 | hra |
| 3 | g’ (g), j’ (j)<sup>A</sup> | 8 | tr | 13 | dr’ (dr) | 18 | wā |
| 4 | c, pc | 9 | thr | 14 | tr | 19 | d |
| 5 | ch, pch | 10 | dr’ (dr) | 15 | thr | 20 | lh |

<sup>A</sup> Palatal variants j, j’, c, ch are generally used before a (ā), o (ō) and u (ū).

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

1 | a | 2 | i | 3 | u | 4 | e | 5 | o

**Other symbols**
- syllable separation character

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

| 1 | a | 20 | ტჰჰჰ ი, i, ik | 39 | ტჰჰჰ um | 58 | ტჰჰჰ em |
| 2 | ჰჰ ა, a, ak | 21 | ტჰჱ ი, i | 40 | ტჰჰ უ, უ | 59 | ტჰჰ ḍ |
| 3 | ჰჰ ა, a, ak | 22 | ტჰჱ ი, i | 41 | ტჰჰ უ, უ | 60 | ტჰჰ ე, en |
| 4 | ა, ang | 23 | ტჰჱ ი, ing | 42 | ტჰჰ უ, უ | 61 | ტჰჰ o, o |
| 5 | ა, ang | 24 | ტჰჱ ი, ing | 43 | ტჰჰ უ, უ | 62 | ტჰჰ o, ok |
| 6 | ჰჰ ap | 25 | ტჰჳ ip | 44 | ტჰჰ უ, უ | 63 | ტჰჰ ḍ, o, ok |
| 7 | ჰჰ ap | 26 | ტჰჳ ip | 45 | ტჰჰ უ, უ | 64 | ტჰჰ o, ong |
| 8 | ჰჰ am | 27 | ტჰჳ im | 46 | ტჰჰ e, i | 65 | ტჰჰ o, ong |
| 9 | ჰჰ am | 28 | ტჰჳ im | 47 | ტჰჰ ḍ | 66 | ტჰჰ ḍ, op |
| 10 | ა | 29 | ტჰჳ i | 48 | ტჰჰ iu, iu | 67 | ტჰჰ ḍ, op |
| 11 | ა, ḍ | 30 | ტჰჳ in | 49 | ტჰჰ ḍ, ek | 68 | ტჰჰ om |
| 12 | ა | 31 | ტჰჳ უ, უ, ū | 50 | ტჰჰ ḍ, ek | 69 | ტჰჰ om |
| 13 | ე, ḍ | 32 | ტჰჳ უ, უ, ū, uk | 51 | ტჰჳ ḍ | 70 | ტჰჳ o |
| 14 | ა, ḍ | 33 | ტჰჳ უ, უ, ū, uk | 52 | ტჰჳ ḍ | 71 | ტჰჳ o |
| 15 | ე, en | 34 | ტჰჳ უ, ū, ung | 53 | ტჰჳ e, eng | 72 | ტჰჳ o |
| 16 | ტჳ i | 35 | ტჳ უ, ū, ung, u, ung | 54 | ტჳ e, eng | 73 | ტჳ o, ō |
| 17 | ტჳ ი, i | 36 | ū | 55 | ტჳ ep | 74 | ტჳ ḍ, on |
| 18 | ტჳ iu | 37 | ū | 56 | ტჳ ep | 75 | ტჳ ḍ, ep |
| 19 | ტჳ ი, i, ik | 38 | ū | 57 | ტჳ um | 76 | ტჳ em |

The choice of variants depends on pronunciation. There are often exceptions to the equivalents above. The table was compiled on the basis of the rules expressed in the handbook of Dzongkha.
and with regard to the examples given. In addition there exist more complex syllables with consonants ’am, ‘an or ‘aw at the end, e.g. აკმ აპ, ახშ იპ, ახშ დოპ, აგმ ადპ, აგშ იპ, აგშ დპ, აგშ ამ, ახშ ა, etc. These are not covered here.

Georgian

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

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

Systems of romanization

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

<table>
<thead>
<tr>
<th>No.</th>
<th>Character</th>
<th>Character</th>
<th>Character</th>
<th>Character</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a</td>
<td>ი</td>
<td>კ</td>
<td>თ</td>
<td>დ</td>
</tr>
<tr>
<td>2</td>
<td>b</td>
<td>გ</td>
<td>მ</td>
<td>უ</td>
<td>ც</td>
</tr>
<tr>
<td>3</td>
<td>გ</td>
<td>დ</td>
<td>ნ</td>
<td>კ</td>
<td>ხ</td>
</tr>
<tr>
<td>4</td>
<td>ჲ</td>
<td>ჭ</td>
<td>შ</td>
<td>ჰ</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ჲ</td>
<td>ჷ</td>
<td>ჸ</td>
<td>ჴ</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ჲ</td>
<td>ჴ</td>
<td>ჵ</td>
<td>ჶ</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>z</td>
<td>ჲ</td>
<td>შ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>ტ</td>
<td>ჭ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>ი</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

- ა (t) თ’ დ (t’) თ ქ (q’) ქ წ (ts’) წ
- დ (k’) კ გ (p) პ’ ხ (ch) ხ’ წ (ch’) წ
- ჰ (p’) პ ჲ (gh) გ ძ (dz) ძ ი (j)
- ჯ (zh) ჲ ჴ (q’) ჴ წ (ts’) წ ჱ (ch’) ხ

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

- ა (t) თ’ გ (p) პ’ ხ (ch) ხ’ ხ (kh) ხ
- დ (k’) კ ჲ (k) კ’ წ (ts) წ’
- ჰ (p’) პ ჲ (gh) გ ძ (dz) ძ ი (j)
- ჳ (zh) ჲ ჴ (q’) ჴ წ (ts’) წ ჱ (ch’) ხ

64
Japanese

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

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 (Heboshiki), continues to be used in Japan in certain spheres where it is estimated that a change would seriously affect customs and considerations of international relations, e.g. in passports. In international cartographic products the Modified Hepburn System remains the most used system.

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

Systems of romanization

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

<table>
<thead>
<tr>
<th>Combined forms:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>A アイウエオ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>オ (オオ5)</td>
</tr>
<tr>
<td>B あいうえお</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>オう (おお)</td>
</tr>
<tr>
<td>C かきくけこ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>し</td>
</tr>
<tr>
<td>D がかりゅくょく</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>さ</td>
</tr>
<tr>
<td>1 カキクケコキュキュウキョウ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>だ</td>
</tr>
<tr>
<td>B かきくけこ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>し</td>
</tr>
<tr>
<td>C ka ki ku ke ko</td>
<td>kya</td>
<td>kyu</td>
<td>kyū</td>
<td>kyo</td>
<td>kyō</td>
<td>kō</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D ke (ga4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>kyū</td>
</tr>
<tr>
<td>2 サシスセソシャシュショショウソウ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>そ</td>
</tr>
<tr>
<td>3 サシスセソシャシュショショウソウ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>そ</td>
</tr>
<tr>
<td>B たちつてとちゅちゅうちゅうちょっと</td>
<td>sya</td>
<td>syu</td>
<td>syū</td>
<td>syo</td>
<td>syō</td>
<td>sō</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C タチツテトチャチュチョチョウトウ</td>
<td>cha</td>
<td>chu</td>
<td>chū</td>
<td>cho</td>
<td>chō</td>
<td>tō</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D CHITSU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>nyū</td>
</tr>
<tr>
<td>4 タチツテトチャチュチョチョウトウ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>nyō</td>
</tr>
<tr>
<td>5 タニヌネノニャニュニュョニョノウ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>nyō</td>
</tr>
<tr>
<td>B たちつてとちゅちゅうちょっと</td>
<td>nyū</td>
<td>nyō</td>
<td>nō</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C NA NI NU NE NO NYA NYU NYO NYO NŌ</td>
<td>nyū</td>
<td>nyō</td>
<td>nō</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6 A ハ ヒフ ヘ ホ ヒャ ヒュ ヒョ ヒョ ホ
B は ひ へ ほ ひゃ ひゅ ひょ ひョ ほ
C ha hi hu he ho hya hyu hyo hyo hō
D ha (wa') fu he (e') hyū hyō hō (d')

7 A マ ミム メ モ ミャ ミュ ミョ ミョ モ
B ま み め も みゃ みゅ みょ みョ も
C ma mi mu me mo mya myu myo myo mō
D

8 A ヤ ユ ヨ ヨ ヨ
B や ゆ よ よ
C ya yu yo
D

9 A ラ リル レ ロ リャ リュ リョ リョ ロ
B ら り り ろ りャ りュ りョ りョ ろ
C ra ri ru re ro rya ryu ryo ryō rō
D ryū ryō rō

10 A ワ ロ ロ ロ
B わ を
C wa
D

11 A ン B ん C n C n² D

12 A ガ カ ゲ ゲ コ ギャ ギュ ギョ ギョ ゴ
B が か げ こ ギャ ギュ ギョ ギョ ゴ
C ga gi gu ge go gya gyu gyo go
D gyū gyō go

13 A ザ ジ ス サ ジャ ジュ ジョ ジョ ソ
B ざ じ す す じゃ ジュ じョ ジョ ソ
C za zi zu ze zo zya zyu zyo zyo zō
D ji ja ju jū jo jō zō

14 A ダ チ デ ド ド
B だ ぢ で ど ど
C da di du de do
D ji zu dō

15 A パ ビ ブ ベ ポ ピュ ピョ ポ ポ ポ
B ぱ ぴ ぷ ぺ ぽ ピュ ピョ ポ ポ ポ
C ba bi bu be bo bya byu byo byō bō
D byū byō bō

16 A ハ ビ ブ ベ ポ ビュ ビョ ポ ポ ポ
B ば び ぷ ぺ ぽ ビュ ビョ ポ ポ ポ
C pa pi pu pe po pya pyu pyo pyō pō
D pyū pyō pō

Notes
1. A small-script tu/tsu form (つ / つ) is inserted between kana symbols to indicate a double consonant (kk, ss, tt, tt, tsv, tsv, pp in Kunrei-siki; kk, ss, ssh, tt, tts, tch, pp in Hepburn).
2. ショ is romanized always n in Kunrei-siki; when it is necessary to separate the sound n from
the vowel or y to follow, the apostrophe is added after the n. In modified Hepburn the character is romanized n’ before y or a vowel letter, n in all other cases; earlier also m was used before b, p, or m.

3. Long vowels are expressed in Kunrei-siki by placing a circumflex (’) over a vowel; in Hepburn long vowels are indicated by a macron (¨).

4. The romanization in parentheses (in modified Hepburn) is used only in those cases where the kana symbol is known to be pronounced in the manner indicated.

5. The combination in parentheses is used to denote the word meaning 'big, great'.

Kazakh

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

Systems of romanization

The BGN/PCGN 1979 System gives the following romanization:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>А</td>
<td>Ә</td>
<td>Б</td>
<td>В</td>
<td>Г</td>
<td>Е</td>
<td>Д</td>
<td>Е</td>
<td>Е</td>
<td>Ж</td>
<td>З</td>
</tr>
<tr>
<td>a</td>
<td>ә</td>
<td>б</td>
<td>в</td>
<td>г</td>
<td>е</td>
<td>д</td>
<td>е</td>
<td>е</td>
<td>ж</td>
<td>з</td>
</tr>
<tr>
<td>ї</td>
<td>й</td>
<td>к</td>
<td>q</td>
<td>к</td>
<td>х</td>
<td>х</td>
<td>х</td>
<td>х</td>
<td>ц</td>
<td>ц</td>
</tr>
<tr>
<td>й</td>
<td>й</td>
<td>у</td>
<td>у</td>
<td>у</td>
<td>у</td>
<td>у</td>
<td>у</td>
<td>у</td>
<td>ts</td>
<td>ts</td>
</tr>
<tr>
<td>й</td>
<td>й</td>
<td>т</td>
<td>t</td>
<td>т</td>
<td>т</td>
<td>т</td>
<td>т</td>
<td>т</td>
<td>ch</td>
<td>ch</td>
</tr>
<tr>
<td>ї</td>
<td>й</td>
<td>ш</td>
<td>sh</td>
<td>ш</td>
<td>ш</td>
<td>ш</td>
<td>ш</td>
<td>ш</td>
<td>tsch</td>
<td>tsch</td>
</tr>
</tbody>
</table>

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

о (i) ә а k (q) қ x (kh) h i (i) i
g (gh) ғ ғ h (ng) ғ h (h) ғ ә (ә) ә
y (yo) ө ө ү (ts) ү c ю (yu) ү
c (zh) Җ Җ y (u) ү у ш (sh) щ
л (l) i y (u) ү y (u) ү у ш (sh) щ
ї (y) j y (u) ү ў (shch) љ

67
Kirghiz

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

**Systems of romanization**

The BGN/PCGN 1979 System gives the following romanization:

<table>
<thead>
<tr>
<th>1</th>
<th>А а a</th>
<th>10</th>
<th>И и i</th>
<th>19</th>
<th>П п p</th>
<th>28</th>
<th>Ч ч ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Б б b</td>
<td>11</td>
<td>Й й y</td>
<td>20</td>
<td>Р р r</td>
<td>29</td>
<td>Ш ш sh</td>
</tr>
<tr>
<td>3</td>
<td>В в v</td>
<td>12</td>
<td>К к k</td>
<td>21</td>
<td>С с s</td>
<td>30</td>
<td>Шч шч shch</td>
</tr>
<tr>
<td>4</td>
<td>Г г g</td>
<td>13</td>
<td>Л л l</td>
<td>22</td>
<td>Т т t</td>
<td>31</td>
<td>Ьь Ьь</td>
</tr>
<tr>
<td>5</td>
<td>Д д d</td>
<td>14</td>
<td>М м m</td>
<td>23</td>
<td>У у u</td>
<td>32</td>
<td>Ъъ Ѵъ</td>
</tr>
<tr>
<td>6</td>
<td>Е е e</td>
<td>15</td>
<td>Н н n</td>
<td>24</td>
<td>Уу уу</td>
<td>33</td>
<td>Ьь Ьь</td>
</tr>
<tr>
<td>7</td>
<td>Ё ё yo</td>
<td>16</td>
<td>Нн нн ng</td>
<td>25</td>
<td>Ф ф f</td>
<td>34</td>
<td>Э э e</td>
</tr>
<tr>
<td>8</td>
<td>Ж ж j</td>
<td>17</td>
<td>О о o</td>
<td>26</td>
<td>Хх kh</td>
<td>35</td>
<td>Ю ю yu</td>
</tr>
<tr>
<td>9</td>
<td>З з z</td>
<td>18</td>
<td>О o o</td>
<td>27</td>
<td>Ц ц ts</td>
<td>36</td>
<td>Я я ya</td>
</tr>
</tbody>
</table>

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

ё (yo)  е  ө (ö)  ò  ч (ch)  ѓ  ѣ (yu)  ú  
ж (j)  ž  ɣ (ü)  ü  ў (sh)  š  я (ya)  ā  
ў (у)  j  х (kh)  h  ў (shch)  š  
у (ng)  ŋ  є (ts)  c  э (e)  ё  

Korean

No romanization system for Korean has been approved at the United Nations conferences on the standardization of geographical names, although systems for the romanization of Korean have been presented at several sessions of the United Nations Group of Experts on Geographical Names (UNEGGN). 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 UNEGNN in 1994. In the Republic of Korea the Ministry of Culture and Tourism adopted in July 2000 a new system

---

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

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

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

In linguistics also the so-called Yale system of romanization is widely accepted\textsuperscript{45}.

Korean uses an alphabetic script in which the characters are grouped graphically together into complex syllable blocks. So, the geographical name P'anmunjom is written 

\begin{verbatim}
PKPM, not PKLO-LPKM.
\end{verbatim}

\textbf{Systems of romanization}

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

\begin{tabular}{|c|c|c|c|c|}
\hline
   & 1  & 2  & 3  & 4  & 5  \\
\hline
1 & k/g & k   & g, k$^*$ & k   & k  \\
2 & kh/k & kh  & k   & k'  & kh \\
3 & kk/gg & kk  & kk  & kk  & kk \\
4 & t/d & t   & d, t$^*$ & t   & t  \\
5 & th/t & th  & t   & t'  & th \\
6 & tt/dd & tt  & tt  & tt  & tt \\
7 & p/b & p   & b, p$^*$ & p   & p  \\
8 & ph/p & ph  & p   & p'  & ph \\
9 & pp/bb & pp  & pp  & pp  & pp \\
10 & c/j & j   & j   & ch  & c  \\
11 & ch/c & ch  & ch  & ch' & ch \\
12 & cc/jj & jj  & jj  & tch & cc \\
\hline
\end{tabular}


\textsuperscript{43} \textit{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.

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

A See notes 2, 3 and 4.
B Word-initially.
C At the end of a syllable.
D At the beginning of a syllable.
E 0 is not romanized syllable-initially, syllable-finally it is romanized ng.
F The first romanization equivalent is used when occurring before a vowel, the second equivalent is used when followed by another consonant or if it forms the final sound of a word.
G 2 is romanized as r when followed by a vowel, and as l when followed by a consonant or when appearing at the end of a word. 22 is romanized as ll.

Notes to ISO TR 11941: 1996
1. Where there are pairs of romanizations, the first corresponds to Method I and the second to Method II.
2. To avoid ambiguity the apostrophe is used to denote the beginning of the next syllable in a polysyllabic word.
3. By Method I the apostrophe is needed in the following cases:
   a. if the consonant 0 is at the initial position of a non-initial syllable of a polysyllabic word: ḳषे kkoch’i, ḳषे kang’e, ḳषे anc’ara;
b. if the five double consonants appear at the initial positions of the non-initial syllables of polysyllabic words: 야끼 a'kka, 흰색 hyin'teok;

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

4. By Method II the apostrophe is needed in the following cases:
   a. if the consonant ㅇ is at the initial position of a non-initial syllable of a polysyllabic word: 광gcc'i, 광g'ang'e, 양바람 anj'ara;
   b. if the five double consonants appear at the initial positions of the non-initial syllables of polysyllabic words: 야끼 a'gga, 흰색 hyin'ddeog.

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

Lao

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

System of romanization

The BGN/PCGN 1966 System gives the following romanization:\textsuperscript{46}:

Consonant characters

\begin{tabular}{cccccc}
1 & น (k) & 5 & ง ch & 9 & โ (t) \\
2 & ข kh & 6 & s & 10 & ต t \\
3 & ค kh & 7 & x & 11 & ฑ th \\
4 & ฅ ng (ng) & 8 & ก gn (y)\textsuperscript{A} & 12 & ฑ th \\
 & 13 & น (n) & 18 & ph \\
 & 14 & b (p) & 17 & f \\
 & 15 & p & 24 & o, o\textsuperscript{C} \\
 & 16 & ph & 19 & f \\
 & 21 & y & 20 & m \\
 & 22 & r & 25 & h\textsuperscript{D} \\
 & 26 & (k) \\
\end{tabular}

\textsuperscript{A} Character variation ง is used at the end of a syllable (ริยา khay), after แ (นิว gnap) and between consonants (โทก siang).

\textsuperscript{B} Subscript variation:_'., e.g. กนภ lam.

\textsuperscript{46} For full text and explanations see the original table in Romanization Systems and Roman-script Spelling Conventions. Prepared by the U.S. Board on Geographic Names Foreign Names Committee Staff. Published by the Defense Mapping Agency, 1994, pp. 63-67.
Lao characters are romanized in two ways depending on whether they denote syllable-initials or syllable-finals. The romanization of syllable-finals is indicated in parentheses. Some other consonant characters which are not accompanied by a vowel character form a full syllable when used within a word: ໒ kkha, ໓ kkha, ໔ tcha, ໕ tsa, ໖ tsa, ໗ tta, ໘ ttha, ໙ ttha, ໚ ppa, ໛ ppha, ໜ pfa.

Vocalic nuclei (• stands for any consonant character)

| 1   | a  | 12 | •  | é  | 23 | •  | eu | 34 | •  | oua |
| 2   | a  | 13 | •  | è  | 24 | •  | eu | 35 | •  | oua |
| 3   | a  | 14 | •• | é  | 25 | •  | eu | 36 | •  | oua |
| 4   | i  | 15 | •  | è  | 26 | •• | eu | 37 | •  | oua |
| 5   | i  | 16 | •• | o  | 27 | •  | ia | 38 | •  | ai  |
| 6   | u  | 17 | •• | o  | 28 | •  | ia | 39 | •  | ai  |
| 7   | u  | 18 | •  | o  | 29 | •  | ia | 40 | •  | ao  |
| 8   | ou | 19 | •• | o  | 30 | •  | ia | 41 | •  | am  |
| 9   | ou | 20 | •  | o  | 31 | •• | ou | 42 | •  | oua |
| 10  | •• | é  | 21 | •  | o  | 32 | •• | ou | 43 | •  | oua |
| 11  | •• | é  | 22 | •• | o  | 33 | •• | ou | 44 | •  | oua |

Reduplicating mark

• (e.g. ວໝໝໝ xèoxèo)

Notes

1. Tone marks ໒, ໔, ໖, ໘ should not be romanized.

2. At the beginning of a syllable, the character ໒ unaccompanied by a vowel or tone mark and occurring immediately before ໒ gn, ໒ n, ໒ m, ໒ r, ໒ l, or ໒ v should generally not be romanized: ຂ໒໒໒໒໒໒ Lounghphrabang. Often ligatures ຂ໒, ຂ໒, ຂ໒ are used in such cases.

3. At the beginning of a word ໒ should not be romanized: ໒ ເ k. At the beginning of a syllable within a word the character should beromanized by a hyphen: ທ໒໒໒໒ xa-ём, ທ໒໒໒໒ Ban Kheng-l.

4. The character ໒ at the beginning of a syllable should be romanized v: ທ໒໒໒ thanva. As the second character of a combination, ໒ should be romanized o. At the end of a syllable the character is romanized o generally (໒໒ nao, ທ໒໒ kêu, ທ໒໒ diao) but after i it is romanized ou (ຸ໒ khou, ທ໒໒ tiou).

5. In consonant combinations where the second character is ໒ r, ໒ l or ໒ v, it should be
romanized character by character, even if the second character may represent no pronounced sound: បែបវត្ត្រាណ Ban Vatphraxi, អ៊ុងខត្ត Khlitsatian, សោយ soay, គ័ន khoeng.

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

7. In some words of foreign origin the variant ṭ r occurs: ṭīr beur.

Maldivian

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

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

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

System of romanization

The official romanization (1987) is as follows.

Consonants

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

A Romanized n ' if used without any vowel or auxiliary sign: ʦːnː kan'du.

B Not romanized but see note 1.

C Character ʈ is romanized iy, e.g. ʈːʁː aiypulhu.
Consonants with diacritical marks (used mainly in words of Arabic origin, corresponding Arabic characters and their romanizations are shown in parentheses)

25 ꝰ (ṣ th)  29 ꝰ (ṣ zh)  33 ꝰ (ṣ t)  37 ꝰ (ṣ q)
26 ꝰ (ṣ h)  30 ꝰ (ṣ sh)  34 ꝰ (ṣ z)  38 ꝰ (ṣ w)
27 ꝰ (ṣ kh)  31 ꝰ (ṣ s)  35 ꝰ (ṣ ')
28 ꝰ (ṣ dh)  32 ꝰ (ṣ l)  36 ꝰ (ṣ gh)

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

Vowels (* stands for any consonant character)

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

Diacritical mark

* (not romanized; marks absence of the vowel)

Notes

1. The character ꝰ is not romanized. If it bears a vowel character, that vowel character alone is romanized, e.g. ꝰ ꝰ irugal.
2. Characters ꝰ and ꝰ word-finally are romanized h, e.g. ꝰ ꝰ rah. In other cases, when ꝰ and ꝰ are used, the romanized value of the following consonant character is doubled, e.g. ꝰ ꝰ ꝰ ꝰ batteh.

Mongolian
(Cyrillic)

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

The BGN/PCGN 1964 System gives the following romanization:

1. A  a
2. B  b
3. В  v
4. Г  g
5. Д  d
6. Е  yо
7. Ё  yo
8. Ж  j
9. З  dz
10. И  i
11. Й  y
12. К  k
13. Л  l
14. М  m
15. Н  n
16. О  o
17. Ò  ô
18. П  p
19. Р  r
20. С  c
21. Т  т
22. У  у
23. Я  y
24. Ф  f
25. Х  h
26. Ц  ts
27. Ч  ch
28. ш  sh
29. шч  shch
30. Ь  ’
31. Ъ  ’
32. Й  ‘
33. Е  е
34. Ю  yu, yü
35. Я  ya
36. ш  sh
37. шч  shch
38. Ь  ’
39. Ъ  ’
40. Й  ‘

A The character b occurs initially and after а, м and б.
B The character v occurs medially in Mongolian words and initially in foreign words.
C The character й occurs only as the second element of diphthongs.
D The characters k, ы, ф and ю occur only in foreign words.
E The character б occurs only in Russian words.
F The character й occurs only in suffixes to words containing the vowel character a, o or y.
G In monosyllables, the character ю is romanized yu or yю depending on pronunciation; in polysyllables, it is romanized yu when followed by a, о, or u, but yю when followed by i, e, o, or у.

Pashto

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

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

System of romanization

The BGN/PCGN 1968 System gives the following romanization:

1. A  a
2. В  b
3. П  p
4. Т  t
5. й  t
6. С  s
7. ж  sz
8. жч  shch
9. ts
10. dz
11. ж  zh
12. жч  shch
13. ж  zh
14. жч  shch
15. ж  zh
16. жч  shch
17. ж  zh
18. жч  shch
19. ж  zh
20. жч  shch
21. ж  zh
22. жч  shch
23. ж  zh
24. жч  shch
25. ж  zh
26. жч  shch
27. ж  zh
28. жч  shch
29. ж  zh
30. жч  shch
31. ж  zh
32. жч  shch
33. ж  zh
34. жч  shch
35. ж  zh
36. жч  shch
37. ж  zh
38. жч  shch
39. ж  zh
40. жч  shch

A Not romanized, but see romanizations accompanying alip (ı) in the table for vowels.
Vowel characters and diacritical marks (*) stands for any consonant

1. a 6. ay 11. ē 16. ū
2. ā 7. ey 12. ēy 17. ā
3. i 8. ey 13. u 18. ē

4. ī 9. ē 14. ū 19. ē (A)
5. ĕ 10. ew 15. ow

* Doubling of consonant letter.

## Sinhalese

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

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

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

### System of romanization

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

#### I. Independent vowel characters

1. a 5. i 9. Ɂ 13. ai\(\textsuperscript{A}\)
2. ā 6. ĭ 10. Ɂa 14. ē
3. ē 7. u 11. ē 15. o
4. ē 8. ĕ 12. e 16. au

\(\textsuperscript{A}\) The character Ɂ was romanized as ăi and the character Ē as ai in Sharma’s table, but there was no romanization for the character ē 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.

---

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

1.  a  5.  i  9.  a  r  13.  ai  m
2.  ā  6.  ī  10.  ā  ō  14.  ō  ō  h
3.  ā  17.  ā  u  11.  ē  e  15.  ē  o  18.  ē  o  (D)
4.  ā  19.  ā  u  12.  ā  e  16.  ā  u  Au

A Exceptions: ā re, ā ra.
B Exceptions: ḗ ku, ḗ gu, ḗ tu, ḗ bhū, ā ru, ḗ shu.
C Exceptions: ḗ ku, ḗ gu, ḗ tu, ḗ bhū, ā ru, ḗ shu.
D Marks absence of the inherent vowel. Special forms: ḗ kh, ḗ ni, ḗ ch, ḗ t, ḗ d, ḗ dh, ḗ b, ḗ m, ḗ v.

III. Consonant characters

2.  ḗ kha 7.  ḗ chha 12.  ḗ tha 17.  ḗ tha 22.  ḗ pha 27.  ḗ ra 32.  ḗ sha
4.  ḗ gha 9.  ḗ jha 14.  ḗ dha 19.  ḗ dha 24.  ḗ bha 29.  ḗ va 34.  ḗ ha
5.  ḗ thā 10.  ḗ thā 15.  ḗ na 20.  ḗ na 25.  ḗ ma 30.  (Fa) 35.  ḗ ja

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

IV. Consonant ligatures

Examples

- r-  ḗ rka, ḗ rta, ḗ rma, ḗ rva, ḗ rsha.
- r-  ḗ kra, ḗ gra, ḗ gra, ḗ tra, ḗ dra, ḗ tra, ḗ dra, ḗ dhra, ḗ pra, ḗ bra, ḗ bhra, ḗ shra.
- r-  ḗ kya, ḗ tya, ḗ wya.

Prenasalized sounds: ḗ ñga, ḗ ñja, ḗ ñda, ḗ nda, ḗ mba.
Other ligatures: ḗ kh, ḗ jha, ḗ dva, ḗ nda.

Tajik

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

Systems of romanization

The BGN/PCGN 1994 System gives the following romanization:

<table>
<thead>
<tr>
<th>A</th>
<th>a</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>B</td>
<td>b</td>
</tr>
<tr>
<td>V</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>Г</td>
<td>g</td>
<td></td>
</tr>
<tr>
<td>Ф</td>
<td>gh</td>
<td></td>
</tr>
<tr>
<td>Д</td>
<td>d</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>e</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Е</th>
<th>е</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ж</td>
<td>zh</td>
</tr>
<tr>
<td>И</td>
<td>i</td>
</tr>
<tr>
<td>И</td>
<td>i</td>
</tr>
<tr>
<td>О</td>
<td>o</td>
</tr>
<tr>
<td>П</td>
<td>p</td>
</tr>
<tr>
<td>Р</td>
<td>r</td>
</tr>
</tbody>
</table>

No information was available on the BGN/PCGN 1994 System.
The above table reflects the new Tajik alphabetical sequence as decreed by the Tajik government in 1998. Some of the characters were then abolished but should they appear in names they are romanized as follows:

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

- ғ (gh) ғ ғ (y) j ғ (ü) ü ғ (‘) 
- ё (yo) ё қ (q) k қ (ch) қ қ (ё) қ
- ж (zh) ж ҳ (kh) h ҳ (j) ҳ ҳ (yu) ҳ
- и (i) і қ (h) І қ (sh) $ я (ya) ҳ

Tigrinya

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

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

System of romanization

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

Combinations with W:

<table>
<thead>
<tr>
<th>1(D)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6(A)</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>ғ</td>
<td>ғ</td>
<td>ғ</td>
<td>ғ</td>
<td>ғ</td>
<td>ғ</td>
<td>ғ</td>
</tr>
<tr>
<td>hа</td>
<td>hu</td>
<td>hі</td>
<td>ha</td>
<td>hе</td>
<td>h, hi</td>
<td>ho</td>
</tr>
<tr>
<td>л</td>
<td>л</td>
<td>л</td>
<td>л</td>
<td>л</td>
<td>л</td>
<td>л</td>
</tr>
<tr>
<td>le</td>
<td>lu</td>
<td>лі</td>
<td>la</td>
<td>lе</td>
<td>l, li</td>
<td>lo</td>
</tr>
<tr>
<td>б</td>
<td>б</td>
<td>б</td>
<td>б</td>
<td>б</td>
<td>б</td>
<td>б</td>
</tr>
<tr>
<td>hа</td>
<td>hu</td>
<td>hі</td>
<td>ha</td>
<td>hе</td>
<td>h, hi</td>
<td>ho</td>
</tr>
<tr>
<td>а</td>
<td>а</td>
<td>а</td>
<td>а</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>me</td>
<td>mu</td>
<td>mі</td>
<td>ma</td>
<td>mе</td>
<td>m, mi</td>
<td>mo</td>
</tr>
<tr>
<td>mwa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

For documentation purposes the following consonants may be romanized with diacritical marks: ħ (row 3), ś (row 5), k’ (row 10, specific to Tigrinya), ḥ (row 14), h (row 19), ts’ (row 32).

The vowel characters in row 17 should be romanized ā, ī, ě, ī, and o initially and ‘ă, ‘u, ‘ĭ, ’a, ’ě, ’i, and ’o in all other positions.

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

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

**Ukrainian**

At sessions of the United Nations Group of Experts on Geographical Names there have been presented several documents on the romanization of Ukrainian geographical names. The first scheme for romanizing Ukrainian names was adopted in Ukraine in 199349.

The current national system of romanization was approved by the Minister of Justice (also Chairman of the Committee on Law Terminology) in 199650. 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 (i) are not differentiated in romanization. Also, for practical use further simplifications are allowed which would make the romanization imprecise.

**Systems of romanization**

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

<table>
<thead>
<tr>
<th>1</th>
<th>А а</th>
<th>10</th>
<th>З з</th>
<th>19</th>
<th>О о</th>
<th>28</th>
<th>Ч ч</th>
<th>ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Б б</td>
<td>11</td>
<td>И и</td>
<td>20</td>
<td>П п</td>
<td>29</td>
<td>Ш ш</td>
<td>sh</td>
</tr>
<tr>
<td>3</td>
<td>В в</td>
<td>12</td>
<td>І і</td>
<td>21</td>
<td>Р р</td>
<td>30</td>
<td>Щ щ</td>
<td>sch</td>
</tr>
<tr>
<td>4</td>
<td>Г г</td>
<td>13</td>
<td>Ії ії, yi^a</td>
<td>22</td>
<td>С с</td>
<td>31</td>
<td>Ю ю</td>
<td>iu, yu^b</td>
</tr>
<tr>
<td>5</td>
<td>Гг</td>
<td>14</td>
<td>Ій ій, y^a</td>
<td>23</td>
<td>Т т</td>
<td>32</td>
<td>Я я</td>
<td>ia, ya^b</td>
</tr>
<tr>
<td>6</td>
<td>Д д</td>
<td>15</td>
<td>К к</td>
<td>24</td>
<td>У у</td>
<td>33</td>
<td>Ьь</td>
<td>'</td>
</tr>
</tbody>
</table>
| 7 | Е е | 16 | Л л | 25 | Ф ф | 34 | ' | '
| 8 | Є е | 17 | М м | 26 | Х х | kh |
| 9 | Ж ж | 18 | Н н | 27 | Ц ц | ts |

^a gh is used in the romanization of ș or (zgh).

^b The second variant is used at the beginning of a word.

Note. Cursive forms of some characters might be formed differently: Аа Бб Вв Гг Іі Ії Іїй Кк Лл Мм Нн Оо Пп Рр Сс Тт Уу Фф Хх Цц Чч Щщ Щщ Ію Ія Ьь'.

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

- р (b/gh) h
- і (i/yi) yi
- ш (sch) shch
- я (ia/ya) ya
- е (ie/ye) ye
- й (i/y) y
- ю (iu/yu) yu
ANNEX:
LANGUAGES THAT HAVE RECENTLY ADOPTED
ROMAN ALPHABETS

Azerbaijani

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

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

<table>
<thead>
<tr>
<th>Cyrillic</th>
<th>Roman</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>z</td>
</tr>
<tr>
<td>b</td>
<td>t</td>
</tr>
<tr>
<td>c</td>
<td>ć</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>ə</td>
<td>ū</td>
</tr>
<tr>
<td>ж</td>
<td>ž</td>
</tr>
</tbody>
</table>

Turkmen

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

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

<table>
<thead>
<tr>
<th>Cyrillic</th>
<th>Roman</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>i</td>
</tr>
<tr>
<td>b</td>
<td>ž</td>
</tr>
<tr>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>ə</td>
<td>ū</td>
</tr>
<tr>
<td>ж</td>
<td>ž</td>
</tr>
</tbody>
</table>

A The letter sequence ūe is used initially, after vowel characters (a, e, ē, ī, o, ō, u, ū, y, ы, э, ю, я) and ĭ.

B Earlier variations of the letters that may still be encountered are as follows (1999 = 1993): Ññ = Nñ, Șṣ = Sṣ Ė́y = ӄį, Žž = Ľľ.

C The Cyrillic ș was used in loanwords only, it has been replaced mainly by s, e.g. Франция Fransîja.

D This letter sequence is rarely encountered.
**Uzbek**

In May 1995 a new Roman alphabet was officially decreed for the Uzbek language, amending the alphabet adopted earlier, in 1993. The Roman alphabet is now in parallel use with the Cyrillic alphabet. Full implementation of it has been postponed to the year 2005\(^5\).

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

<table>
<thead>
<tr>
<th>Cyrillic</th>
<th>Roman</th>
</tr>
</thead>
<tbody>
<tr>
<td>А а а</td>
<td>И и i</td>
</tr>
<tr>
<td>Б б b</td>
<td>Й й y</td>
</tr>
<tr>
<td>В в v</td>
<td>К к k</td>
</tr>
<tr>
<td>Г г g</td>
<td>Л л l</td>
</tr>
<tr>
<td>Д д d</td>
<td>М м m</td>
</tr>
<tr>
<td>Е е e</td>
<td>Н н n</td>
</tr>
<tr>
<td>Ж ж j</td>
<td>П п p</td>
</tr>
<tr>
<td>З з z</td>
<td>Р р r</td>
</tr>
</tbody>
</table>
