

TABLE OF CONTENTS

I. Introduction

II. Approved Romanization Systems and Agreements

Amharic

Arabic

Armenian

Azeri

Bulgarian

Burmese

Byelorussian

Chinese Characters

Georgian

Greek

Hebrew

Japanese Kana

Kazakh Cyrillic

Khmer (Cambodian)

Kirghiz Cyrillic

Korean

Lao

Macedonian

Maldivian

Moldovan

Mongolian Cyrillic

Nepali

Pashto

Persian (Farsi and Dari)

Russian

Serbian Cyrillic

Tajik Cyrillic

Thai

Turkmen

Ukrainian

Uzbek

III. Roman-script Spelling Conventions

Faroese

German

Icelandic

North Lappish

IV. Appendices

A. Unicode Character Equivalents

B. Optimizing Software and Operating Systems to Display BGN-approved geographic names

Table 1. Provenance and Status of Romanization Systems Contained in this Publication

Transliteration System	Class	Date Approved	Originator
BGN/PCGN Amharic 1967 System	System	1967	
BGN/PCGN Arabic 1956 System	System	1956	
BGN/PCGN Armenian 1981 System	System	1981	
Azeri	Roman Alphabet Spelling Convention	2002	Azeri Government 1991
BGN/PCGN Bulgarian 1952 System	System	1952	
BGN/PCGN Burmese 1970 Agreement	Agreement	1970	Burmese Government 1907 System
BGN/PCGN Byelorussian 1979 System	System	1979	
Chinese Pinyin System	Agreement	1979	Xinhua Zidian dictionary. Commercial Press, Beijing 1982.
Chinese Wade-Giles System	Agreement	1979	
BGN/PCGN Faroese Spelling Convention	Roman Script Spelling Convention	1968	
BGN/PCGN 1981 Georgian System	System	1981	
BGN/PCGN German Spelling Convention	Roman Script Spelling Convention	1986	
Greek ELOT 743 System	Agreement	1996	Greek Organization for Standardization
BGN/PCGN Hebrew 1962 System	Agreement	1962	Hebrew Academy System
Japanese Kana Modified Hepburn System	Agreement	1930	
BGN/PCGN Icelandic Spelling Convention	Roman Script Spelling Convention	1968	

Transliteration System	Class	Date Approved	Originator
BGN/PCGN Kazakh Cyrillic 1979 System	System	1979	
BGN/PCGN Khmer (Cambodian) 1972 System	Agreement	1972	Service Géographique Khmère (SGK) 1959 System
BGN/PCGN Kirghiz Cyrillic 1979 System	System	1979	
Korean McCune-Reischauer System	Agreement	1943	Transactions of the Royal Asiatic Society, XXIX (1939)
BGN/PCGN Lao System	Agreement	1966	Lao Commission Nationale de Toponymie (CNT) System
BGN/PCGN Macedonian 1981 System	Agreement	1981	
BGN/PCGN Maldivian 1988 System	Agreement	1988	Maldivian Government
Moldovan	Roman Alphabet Spelling Convention	2002	
BGN/PCGN Mongolian Cyrillic 1964 System	System	1964	
BGN/PCGN Nepali 1964 System	System	1964	
BGN/PCGN North Lappish Spelling Convention	Roman Script Spelling Convention	1984	
BGN/PCGN Pashto 1968 System	System	1968	
BGN/PCGN Persian 1958 System	System	1958	
BGN/PCGN Russian 1947 System	System	1947	
BGN/PCGN Serbian Cyrillic 2005 System	Agreement	2005	
BGN/PCGN Tajik Cyrillic 1994 System	System	1994	

Transliteration System	Class	Date Approved	Originator
BGN/PCGN Thai 2002 System	Agreement	2002	Royal Institute of Thailand, 2000 Version
Turkmen Alphabet	Roman Alphabet Spelling Convention	2000	Government of Turkmenistan
BGN/PCGN Ukrainian 1965 System	System	1965	
Uzbek Alphabet	Roman Alphabet Spelling Convention	2000	Government of Uzbekistan

Note: The term ‘system’ is used to denote transliteration systems developed by the BGN or the PCGN. The term ‘agreement’ is used to denote transliteration systems developed by third parties, and adopted by the BGN and PCGN.

INTRODUCTION

This publication contains all of the romanization systems and Roman-script spelling conventions that are currently approved by the U.S. Board on Geographic Names (BGN) and its British counterpart, the Permanent Committee on Geographic Names for British Official Use (PCGN). It therefore supersedes the Transliteration Guide of 1961, the Romanization Guide of 1964, 1967, 1972, and the publication Romanization Systems and Roman Script Spelling Conventions of 1994. Each romanization system and spelling convention presented is identified as being a BGN/PCGN system or a BGN/PCGN agreement, with the date of its joint adoption by the BGN and PCGN indicated in most cases.

Within the U.S. Government, BGN/PCGN romanization systems and agreements are used primarily for the purpose of establishing standardized Roman-script spellings of those foreign geographical names that are written in non-Roman scripts or in Roman alphabets that contain special letters. Geographical names that have been romanized and names originally written in Roman script are made available for general use on the Geonet Names Server, an on-line service of the National Geospatial-Intelligence Agency. This database, which covers virtually every foreign country in the world, provides information as to the name, type, and location of every geographic feature listed, as well as variant spellings of names for finding purposes.

In most cases, familiarity with the writing system of a given language is all that is needed in order to apply the appropriate BGN/PCGN romanization system or agreement correctly. In some cases, however, a thorough knowledge of both the language and its writing system are an absolute requirement. The latter category includes the systems for Arabic, Hebrew, Persian, and Pushto, i.e. systems for languages in which vowels are not ordinarily represented in the script. The BGN/PCGN romanization systems for those languages and for the languages represented in this publication generally contain elements of transliteration – the process of recording the graphic symbols of one writing system in terms of the corresponding graphic Symbols of a second writing system – and of transcription – the process of recording the phonological and/or morphological elements of a language in terms of a specific writing system.*

Some BGN/PCGN romanization systems, e.g. those for Bulgarian and Georgian, exhibit a high degree of reversibility; i.e., the Roman letters that serve as the equivalents of the non-Roman characters of the source script may be converted to the original characters almost unambiguously. Other BGN/PCGN systems and agreements, e.g., those for Amharic and Korean, are not easily reversible. The BGN/PCGN romanization system for Thai provides an extreme example of a non-reversible system; in that system, the Roman letter t is used to represent a total of nineteen different Thai characters in syllable-final position. The Thai romanization system, therefore, can be said to be undifferentiated, since it contains several instances of a single Roman letter or letter combination serving as the equivalent of more than one Thai character.

The Roman letters and letter combinations that are shown as equivalents of the non-Roman characters in the BGN/PCGN romanization systems generally reflect the letters and letter combinations that are used in English orthography. In many cases, however, the number of Roman-letter equivalents needed for a particular system exceeds the number of appropriate letters and letter combinations available in English orthography. As a result, several Roman letters may be shown with diacritical marks in order to provide the necessary differentiation of graphic symbols and insure proper reversibility. In the Persian alphabet, for example, there are four different characters that are pronounced like the letter z in English. In order to differentiate the romanizations of those four characters, the BGN/PCGN system for Persian utilizes the ordinary letter z and three z's with diacritical marks, i.e., \underline{z} , \bar{z} , and \dot{z} . In addition to their use in Roman-letter equivalents in this publication, diacritical marks are used with Roman letters and with non-Roman characters in many languages of the world; e.g. the cedilla (,) in $\tilde{\text{ñ}}$ in Russian. Diacritical marks are just as important as the basic letters and characters of any orthography or romanization system, and, therefore, should never be omitted. Modifying marks that occur internally in both Roman letters and non-Roman characters, e.g., the horizontal bar in the Croatian letter \bar{d} and the vertical bar in the Azerbaijani Cyrillic character $\bar{\text{Ч}}$, are not Generally considered to be diacritical marks but are just as significant and, therefore, should always be retained.

It should be noted that for clarity of presentation and ease of reference the terms character and letter have been used in a mutually exclusive way throughout this publication. The term character has been used to refer to a graphic symbol used only in a non-Roman-script writing system, thereby restricting the term letter to a graphic symbol used only in a Roman-script writing system or in a set of romanization equivalents.

Finally, it may well be pointed out that although the romanization systems and agreements contained in this publication have been approved by the BGN and the PCGN for application to geographic names, the systems are equally applicable to personal names and to text and have been used for such purposes for many years by organizations both inside and outside the U.S. Government.

Please refer to the appendices for Unicode values for letters used in BGN/PCGN romanization systems, as well as for hints on optimizing computer software and operating systems to allow display of Unicode characters and letters. Please refer to the BGN website for updates to this guide and additional information: <http://geonames.usgs.gov/bgn.html>, as well as to the PCGN website: <http://www.pcn.org.uk>. Information on other transliteration systems for toponyms is also maintained on the website of the United Nations Group of Experts on Geographic Names: <http://www.eki.ee/wgrs/>.

* *These definitions were agreed upon in 1971 by the U.N Working Group on a Single Romanization System for Each Non-Roman Writing System and were included in that group's report in the U.N. document, Second United Nations Conference on the Standardization of Geographical Names, London, 10-13 May 1972, vol. II, p. 115.*

It is requested that users of this publication aid in its correction for future printings by reporting errors and changes to either of the addresses listed below. References to or copies of the sources upon which the proposed changes are based should be included.

BGN Executive Secretary, Foreign Names Committee

National Geospatial-Intelligence Agency

Political Geography Division

4600 Sangamore Road

Bethesda, Maryland 20816

U.S.A.

UK Permanent Committee on Geographic Names

C/O The Royal Geographical Society

1 Kensington Gore

London SW7 2AR

U.K.