

INTRODUCTION

This collection contains all of the romanization systems and tables of correspondences¹ that are currently approved by the U.S. Board on Geographic Names (BGN) and the U.K. Permanent Committee on Geographical Names (PCGN). It therefore supersedes the *Transliteration Guide* of 1961; the *Romanization Guide* of 1964, 1967 and 1972; and the publications *Romanization Systems and Roman Script Spelling Conventions* of 1994 and 2008. Each romanization system and spelling convention presented is identified as being a BGN/PCGN system or a BGN/PCGN agreement², with the date of its joint adoption by the BGN and PCGN indicated in most cases (see Table 1).

Within the U.S. and U.K. Governments, 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. Geographical names that have been romanized (often alongside the original script) and names originally written in Roman script are made available for general use on the [GEOnet Names Server](#), an online 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 geographical 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 more thorough knowledge of both the language and its writing system is necessary. The latter category includes the systems for Arabic, Hebrew, Persian, and Pashto, 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 other 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 system³.

There are a number of principles that BGN and PCGN have agreed should be considered when devising a romanization system: it should be as reversible and parsimonious as possible, should include diacritical signs to a minimal degree, and should neither be a guide to pronunciation nor a language treatise. Some romanization systems, e.g. that for 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 systems, e.g.

¹ Tables of correspondence exist for those languages that are written in both Roman-script and another script, where the table can be applied to derive the alternative script. Romanization systems exist for those languages that are not commonly written in Roman-script; the romanization system is applied to the script in order to obtain a standard Roman-script form.

² The term *system* is used to denote transliteration systems developed by the BGN and/or the PCGN. The term *agreement* is used to denote transliteration systems developed by third parties, and adopted by the BGN and the PCGN.

³ 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: https://unstats.un.org/UNSD/geoinfo/UNGEGN/docs/2nd-uncsgn-docs/E_Conf61_4_Add1_e.pdf

that for Amharic, are not wholly 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 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 ensure 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 basic letter **z** and three **zs** with diacritical marks, i.e. **ż**, **z̈** and **z̉**. In addition to their use in Roman-letter equivalents in this collection, diacritical marks are used with Roman letters and with non-Roman characters in many languages of the world: for example, the cedilla is used with the letter **c** to form **ç** in French. Diacritical marks are just as important as the basic letters and characters of any orthography or romanization system, and, therefore, should not be omitted. However, in cases where the user is unable to reproduce correctly the appropriate letter-diacritic combination, omission of the diacritical marks is permissible. Modifying marks that occur internally in both Roman letters and non-Roman characters, e.g. the horizontal bar in the Croatian letter **đ** and the horizontal bar in the Kazakh Cyrillic character **ѣ**, are not generally considered to be diacritical marks but rather an integral part of the basic letter or character and, therefore, should always be retained.

For the correct display and transfer of the letters and characters in this collection, information on the relevant Unicode encoding has been included within each file. Information on other transliteration systems for toponyms is also maintained on the website of the United Nations Group of Experts on Geographical Names [Working Group on Romanization Systems](#).

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 collection. 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 usefully be pointed out that although the romanization systems and agreements contained in this collection have been approved by the BGN and the PCGN for application to geographical names, some or all of the systems may prove similarly applicable to personal names and to text and indeed have frequently been used for such purposes by organizations both within the U.S. and the U.K.

It is requested that users of this collection aid in its correction for future editions 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.

US Board on Geographic Names: [geonames @ nga.mil](mailto:geonames@nga.mil)

UK Permanent Committee on Geographical Names: [info @ pcgn.org.uk](mailto:info@pcgn.org.uk)

Table 1. Provenance and status of Romanization Systems contained in this collection

System	Type	Date adopted by BGN/PCGN	Originator
Abkhaz	System	2011	BGN/PCGN
Adyghe	System	2012	BGN/PCGN
Afghanistan	System	2007	BGN/PCGN
Amharic	System	1967	BGN/PCGN
Arabic	System	1956	BGN/PCGN
Armenian	System	1981	BGN/PCGN
Avar	System	2011	BGN/PCGN
Azerbaijani	Table of Correspondences	1993	Govt of Azerbaijan, 1991
Baluchi	System	2008	BGN/PCGN
Bashkir	Table of Correspondences	2007	Bashkir practice
Belarusian	System	1979	BGN/PCGN
Bulgarian	Agreement	2013	Govt of Bulgaria, 2009
Burmese	Agreement	1970	Govt of Burma, 1907
Chechen	Table of Correspondences	2008	Chechen practice
Chinese Pinyin	Agreement	1979	Govt of China, 1958
Chuvash	System	2011	BGN/PCGN
Dzongkha	Agreement	2010	Govt of Bhutan, 1997
Faroese	Roman Script Spelling Convention	1968	Faroese orthography
Georgian	Agreement	2009	Govt of Georgia, 2002
German	Roman Script Spelling Convention	2000	German orthography

System	Type	Date adopted by BGN/PCGN	Originator
Greek ELOT 743	Agreement	1996	Hellenic Organization for Standardization, 1987
Hebrew	Agreement	2018	Hebrew Academy, with BGN/PCGN modifications
Icelandic	Roman Script Spelling Convention	1968	Icelandic orthography
Inuktitut	Table of Correspondences	2013	Inuit Cultural Institute, 1976
Japanese Kana Modified Hepburn	Agreement	1976	J.C. Hepburn, 1867
Kabardian	System	2011	BGN/PCGN
Karachay-Balkar	Table of Correspondences	2008	Karachay-Balkar practice
Kazakh	System	1979	BGN/PCGN
Khmer (Cambodian)	Agreement	1972	Service Géographique Khmère (SGK) 1959, with BGN/PCGN modifications, 1972
Korean Ministry of Culture and Tourism (for Republic of Korea)	Agreement	2011	Ministry of Culture and Tourism, ROK, 2000
Korean McCune-Reischauer (for Democratic People's Republic of Korea)	Agreement	1945	Transactions of the Royal Asiatic Society, XXIX, 1939
Kurdish	System	2007	BGN/PCGN
Kyrgyz	System	1979	BGN/PCGN
Lao	Agreement	1966	Lao Commission Nationale de Toponymie (CNT), 1965
Macedonian	Agreement	2013	Govt of Macedonia
Maldivian	Agreement	1988	Govt of Maldives, 1987
Moldovan	Table of Correspondences	2002	Govt of Moldova, 1990
Mongolian	System	1964	BGN/PCGN
Nepali	Agreement	2011	Nepal Survey Department

System	Type	Date adopted by BGN/PCGN	Originator
Northern Sami	Roman Script Spelling Convention	1984	Northern Sami orthography
Ossetian	System	2009	BGN/PCGN
Pashto	System	1968	BGN/PCGN
Persian (Dari and Farsi)	System	1958	BGN/PCGN
Russian	System	1947	BGN/PCGN
Rusyn	System	2016	BGN/PCGN
Serbian	Table of Correspondences	2005	Govt of Serbia
Shan	System	2011	BGN/PCGN
Syriac	System	2011	BGN/PCGN
Tajik	System	1994	BGN/PCGN
Tatar	Table of Correspondences	2007	Tatar practice
Thai	Agreement	2002	Royal Institute of Thailand, 2000
Tigrinya	System	2007	BGN/PCGN
Turkmen	Table of Correspondences	2000	Govt of Turkmenistan, 1993
Udmurt	System	2011	BGN/PCGN
Ukrainian	Agreement	2019	Govt of Ukraine, 2010
Urdu	System	2007	BGN/PCGN
Uzbek	Table of Correspondences	2000	Govt of Uzbekistan, 1995
Yakut	System	2012	BGN/PCGN