

Approved: FNC 373
12 December 2012

Foreign Names Committee
Geographic Names Standardization Policy for
Afghanistan
Version 2.3

United States Board on Geographic Names – Foreign Names Committee



December 2012

1. Introduction

This geographic names standardization policy has been prepared as an aid to those geographic names experts who are the working staff for the United States Board on Geographic Names (BGN) in the standardization of the geographic names of Afghanistan for United States Government use. The BGN and its staff work to effect consistent treatment of geographic name spellings in U.S. Government databases, publications, maps, and charts, and are intended to satisfy the statutory requirements levied upon the BGN in Public Law USC 80-242 to develop principles, policies and procedures for geographic names standardization, and to promulgate decisions with respect to the principles of geographic nomenclature and orthography.

2. Languages and Language Policy

a. Demographics

The July 2012 estimated population of Afghanistan is 30,419,928. The Pashtuns are the largest single ethnic group constituting approximately 42% of the population. They inhabit primarily the southern and eastern parts of the country. The Tajiks are the second largest group, constituting approximately 27% of the total population. They are concentrated in the Kabul, Badakhshan and Herat provinces. The Hazaras and Uzbeks each make up approximately 9% of the total population. While several other ethnic groups also exist in Afghanistan, none of them currently account for more than 4% of the overall population. Islam is the official religion with 80% Sunni and 19% Shia. Other religious groups comprise just 1% of the population.¹

b. Languages

The number of languages spoken by Afghans has been reported as anywhere from 32 to 47. Approximately fifty percent of Afghanistan's inhabitants speak Dari (a dialect of modern Persian), while approximately thirty-five percent speak Pashto. Related dialects of Dari are also the primary language of the Tajik and Hazara. Dari, rather than Pashto, serves as the means of communication between speakers of different languages in Afghanistan and, therefore, functions as the lingua franca and the working language of the government.

Afghanistan's 2004 Constitution (unofficial Afghan Government translation - <http://afghanembassy.com/afghanistan/afghanConstitution.asp>) makes the following statements concerning language policy:

Chapter 1, Article 16 [Languages]

(1) From among the languages of Pashto, Dari, Uzbeki, Turkmani, Baluchi, Pashai, Nuristani, Pamiri (alsana), Arab and other languages spoken in the country, Pashto and Dari are the official languages of the state.

¹ Iran. CIA World Factbook Online. < <https://www.cia.gov/library/publications/the-world-factbook/geos/af.html> > .24May2012

(2) The Turkic languages (Uzbaki and Turkmen), Baluchi, Pashai, Nuristani and Pamiri (alsana) are -- in addition to Pashto and Dari -- the third official language in areas where the majority speaks them. The practical modalities for implementation of this provision shall be specified by law.

(3) The state adopts and implements effective plans for strengthening, and developing all languages of Afghanistan.

(4) Publications and radio and television broadcasting are allowed in all languages spoken in the country.

(Adopted by Grand Council on: 4 Jan 2004)

c. Geographic Names Standardization

Although the 2004 constitution of the Islamic Republic of Afghanistan recognizes two, and in some areas three, official languages (Dari, Pashto, and other locally-official languages), in geographic naming practices, Afghanistan appears to follow a principle of univocity, i.e. one and only one name per feature, except in a few cases. The BGN therefore approves one standard name for each feature in Afghanistan², based on a preponderance of evidence from official sources.

There is no legal geographic naming authority in Afghanistan under the framework of its 2004 Constitution. However, the leaders of three Afghan government agencies, the Afghan Geodesy and Cartography Head Office (AGCHO), the Central Statistics Office (CSO), and the Independent Directorate of Local Governance (IDLG) worked together to establish the Board on Geographic Names for Afghanistan (BGN-A). Based on a 2007 Afghan Presidential decree, AGCHO has sole responsibility for all Geodesy, Cartography and Cadastral activities in Afghanistan. Given this authority, the General President of AGCHO was established as the permanent chairperson of the BGN-A. The general presidents of CSO and IDLG make up the other two members of the BGN-A's executive committee. The BGN recognizes the BGN-A as Afghanistan's national names authority.

3. Toponymic Policies

a. Native Script

For many features in Afghanistan, no official Afghan mapping in the Perso-Arabic or Pashto-modified Arabic script is available. Native script names for such features are generally derived and converted from roman-script sources.

² The PCLI is the exception to this rule. It has an approved Dari name and an approved Pashto name to go along with the approved Conventional Name (see section 4a). The Foreign Names Committee (FNC) approved this change at the meeting number 357 held in October of 2008.

The limited sources that are produced in Afghanistan are generally produced in Perso-Arabic or Pashto-modified Arabic script. No official geographic sources are published in any other minority language of Afghanistan. Approved toponyms for Afghanistan are written in a modified version of Arabic script. They are written from right to left and have no distinction between capital and lower-case letters. Each letter appears in up to four forms: initial, medial, final, and isolated. Some letters join either on the right or left side with adjacent characters. While both Dari and Pashto script share all of these aforementioned characteristics, the Pashto alphabet contains several additional letters not found in the Dari alphabet.

The Dari alphabet consists of thirty-two letters, shown in Table 1 below. Numerals are written left-to-right, even though Dari and Pashto are otherwise written right-to-left, Table 2. The Pashto alphabet consists of forty letters, Table 3.

Table 1: Dictionary Order of Dari Alphabet

Letter	Letter Name	Letter	Letter Name	Letter	Letter Name
آ	Alif-mad	ذ	zāl	غ	Ghayn
ا	Alif	ر	Reh	ف	Feh
ب	Beh	ز	Zeh	ق	Qāf
پ	Peh	ژ	Zehh	ک	Kāf
ت	Teh	س	Sīn	گ	Gāf
ث	Seh	ش	Shīn	ل	Lām
ج	Jīm	ص	Ṣwāt	م	Mīm
چ	Cheh	ض	Zwāt	ن	Nūn
ح	Ḥej	ط	Ṭoy	و	Wāw
خ	Kheh	ظ	Zoy	ه	Heh
د	Dāl	ع	‘Ayn	ی	Yā

Table 2: Dari/Pashto Numerals

0 : ۰	1 : ۱	2 : ۲	3 : ۳	4 : ۴, ۴	5 : ۵, ۵	6 : ۶, ۶	7 : ۷	8 : ۸	9 : ۹
-------	-------	-------	-------	----------	----------	----------	-------	-------	-------

Table 3: Dictionary Order of Pashto Alphabet

Letter	Letter Name	Letter	Letter Name	Letter	Letter Name
آ	Alif-maddah	ډ	Dāl	ع	'Ayn
ا	Alif	ذ	Zāl	غ	Ghayn
ب	Beh	ر	Reh	ف	Feh
پ	Peh	ړ	R̄eh	ق	Qaf
ت	Teh	ز	Zeh	ک	Kāf
ټ	T̄eh	ژ	Zehh	گ	Gāf
ث	Seh	ږ	Z̄ehh	ل	Lām
ج	Jīm	س	Sīn	م	Mīm
چ	Cheh	ش	Shīn	ن	Nūn
ځ	Dzeh	ښ	Sh̄īn	ڼ	N̄ūn
خ	Tseh	ص	Ṣād	و	Vāv
ح	Ḥeh	ض	Z̄ād	ه	Heh
خ	Kheh	ط	Ṭā	ی	Yeh
د	Dāl	ظ	Z̄ā		

b. Romanization

Geographic names in Afghanistan are to be transliterated using the 2007 BGN/PCGN National Romanization System for Afghanistan. BGN and PCGN agreed that Afghanistan required its own transliteration system, separate from the 1958 BGN/PCGN Romanization System for Persian, which had been used previously. This system was developed in response to the unique naming conventions specific to Afghanistan and its sources. It is designed to address romanization policies for toponyms rooted not solely in the Dari and Pashto languages, but rather as combined language Afghan toponyms as they are often found throughout the country³. For convenience, the transliteration system is included as Appendix A of this document.

³ Analysis of toponyms in Afghanistan should determine whether the origin of the geographic name is Dari or Pashto. However, this analysis often yields results demonstrating that there is no specific language base for Afghan toponyms. Toponyms often combine language specific terms with grammar elements explicit to another language, yielding an Afghan toponym which encompasses all elements rather than one with specificity to Dari or Pashto.

c. Diacritics in Romanization

The diacritic marks, shown in Table 4, are permitted in Romanized standard name forms in Afghanistan.

Table 4: BGN Approved Diacritic Marks for Afghan Toponyms

Character Name	Character	Unicode	Character Name	Character	Unicode
Uppercase a with acute accent	Á	00C1	Lowercase a with acute accent	á	00E1
Uppercase a with macron	Ā	0100	Lowercase a with macron	ā	0101
Uppercase d with sub-macron	Ḑ	1E0E	Lowercase d with sub-macron	ḑ	1E0F
Uppercase e with circumflex	Ê	00CA	Lowercase e with circumflex	ê	00EA
Uppercase e with macron	Ē	0112	Lowercase e with macron	ē	0113
Uppercase h with sub-macron	Ḣ	0048+0332	Lowercase h with sub-macron	ḣ	0068+0332
Uppercase h with cedilla	Ĥ	1E28	Lowercase h with cedilla	ĥ	1E29
Uppercase i with macron	Ī	012A	Lowercase i with macron	ī	012B
Uppercase n with sub-macron	Ṇ	1E48	Lowercase n with sub-macron	ṇ	1E49
Uppercase r with sub-macron	Ṛ	1E5E	Lowercase r with sub-macron	ṛ	1E5F
Uppercase s with macron	Ŝ	0053+0304	Lowercase s with macron	ŝ	0073+0304
Uppercase s with sub-macron	Ṣ	0053+0332	Lowercase s with sub-macron	ṣ	0073+0332
Uppercase s with cedilla	Ş	015E	Lowercase s with cedilla	ş	015F
Uppercase t with sub-macron	Ṫ	1E6E	Lowercase t with sub-macron	ṽ	1E6F
Uppercase t with cedilla	Ṭ	0162	Lowercase t with cedilla	ṭ	0163
Uppercase u with macron	Ū	016A	Lowercase u with macron	ū	016B
Uppercase z with macron	Ẑ	005A+0304	Lowercase z with macron	ẑ	007A+0304
Uppercase z with sub-macron	Ẓ	1E94	Lowercase z with sub-macron	ẓ	1E95

Uppercase z with cedilla	Ź	005A+0327	Lowercase z with cedilla	ź	007A+0327
-----------------------------	---	-----------	-----------------------------	---	-----------

The apostrophe (') and inverted apostrophe (´) are special symbols, and are permitted in the romanized forms of geographic names. Uppercase letters in standardized name forms retain diacritics. Diacritic marks other than the ones listed in Table 4 above are not permitted in official names for Afghanistan.

d. Generic Terms

Appendix B of this document contains a glossary of generic terminology for geographical names in Afghanistan.

e. Temporary Districts

Due to issues with security and stability in some areas of Afghanistan, the Afghan government has, in the past, and may continue to establish temporary districts. However, there is no specific designation for those districts and their names are not currently collected by the BGN or stored in the Geographic Names Database (GNDB).

f. Abbreviations, Numbers and Hyphens

The BGN does not use abbreviations in the approved form of any name in Afghanistan. In cases where a source lists the abbreviated form of a generic or name, the abbreviated word is completely spelled out. Likewise, when a geographic name contains a number, the number is spelled out in the BGN-approved name. Hyphens are used on occasion when they are found on the source and are grammatically appropriate. Hyphens are also used for Afghan names containing *izāfah* as outlined in the 2007 BGN/PCGN National Romanization System for Afghanistan (Appendix A).

4. Political Geography Policy

a. Country Name and Capital

Country Name:

Conventional Short Form	Afghanistan
Conventional Long Form	Islamic Republic of Afghanistan
Native Short Form – Romanized	Afghānistān
Native Long Form – Romanized [Dari]	Jamhūrī-ye Islāmī-ye Afghānistān
Native Long Form – Romanized [Pashto]	Dê Afghānistān Islāmī Jumhūriyat
Native Short Form – Native Script	افغانستان
Native Long Form – Native Script [Dari]	جمهوری اسلامی افغانستان
Native Long Form – Native Script [Pashto]	د افغانستان اسلامی جمهوریت

Capital:

Conventional Approved
Native – Romanized
Native – Native Script

Kabul
Kābul
کابل

b. First-Order Administrative Divisions

The BGN recognizes ولایت – Wilāyat, or “province”, as the first-order administrative division (ADM1) in Afghanistan.

First-order Administrative Divisions

	<i>Name</i>	<i>Seat</i>	<i>FIPS 10-4</i>	<i>ISO 3166-2</i>
1.	Badakhshān بدخشان	Faīzābād فیض آباد	AF01	AF-BDS
2.	Bādghīs بادغیس	Qal‘ah-ye Now قلعه نو	AF02	AF-BDG
3.	Baghlān بغلان	Pul-e Khumrī پل خمری	AF03	AF-BGL
4.	Balkh بلخ	Mazār-e Sharīf مزار شریف	AF30	AF-BAL
5.	Bāmyān بامیان	Bāmyān بامیان	AF05	AF-BAM
6.	Dāykundī دایکندی	Nīlī نیلی	AF41	AF-DAY
7.	Farāh فراه	Farāh فراه	AF06	AF-FRA
8.	Fāryāb فاریاب	Maīmanah میمنه	AF07	AF-FYB
9.	Ghaznī غزنی	Ghaznī غزنی	AF08	AF-GHA
10.	Ghōr غور	Chaghcharān چغچران	AF09	AF-GHO
11.	Helmand هلمند	Lashkar Gāh لشکر گاه	AF10	AF-HEL
12.	Herāt هرات	Herāt هرات	AF11	AF-HER
13.	Jowzjān جوزجان	Shibirghān شبرغان	AF31	AF-JOW
14.	Kābul کابل	Kābul کابل	AF13	AF-KAB
15.	Kandahār کندهار	Kandahār کندهار	AF23	AF-KAN
16.	Kāpīsā کاپیسا	Maḥmūd-e Rāqī محمود راقی	AF14	AF-KAP
17.	Khōst خوست	Khōst خوست	AF37	AF-KHO

18.	Kunar کنړ	Asadābād اسد آباد	AF34	AF-KNR
19.	Kunduz کنډز	Kunduz کنډز	AF24	AF-KDZ
20.	Laghmān لغمان	Mehtar Lām مهتر لام	AF35	AF-LAG
21.	Lōgar لوگر	Pul-e ‘Alam پل علم	AF17	AF-LOG
22.	Nangarhār ننگرهار	Jalālābād جلال آباد	AF18	AF-NAN
23.	Nīmrōz نیمروز	Zaranj زرنج	AF19	AF-NIM
24.	Nūristān نورستان	Pārūn پارون	AF38	AF-NUR
25.	Paktīkā پکتیکا	Sharan شرن	AF29	AF-PKA
26.	Paktiyā پکتیا	Gardēz گردېز	AF36	AF-PIA
27.	Panjshayr ⁴ پنجشیر	Bāzārak بازارک	AF42	AF-PAN
28.	Parwān پروان	Chārīkār چاریکار	AF40	AF-PAR
29.	Samangān سمنگان	Aībak ایبک	AF32	AF-SAM
30.	Sar-e Pul سر پل	Sar-e Pul سر پل	AF33	AF-SAR
31.	Takhār تخار	Tāluqān تالقان	AF26	AF-TAK
32.	Uruzgān ارزگان	Tarīn Kōṭ ترین کوت	AF39	AF-URU
33.	Wardak وردک	Maīdān Shahr میدان شهر	AF27	AF-WAR
34.	Zābul زابل	Qalāt قلات	AF28	AF-ZAB

c. Unique Geopolitical Situations

Note: For the latest country-specific boundary dispute information, visit the U.S. Department of State’s Boundaries and Sovereignty Encyclopedia, or B.A.S.E., at <http://base.us-state.osis.gov>.

1. Disputed Names

The BGN-Foreign Names Staff is not aware of any names disputed between Afghanistan and its neighbors.

⁴ This is the Approved name while the Conventional name for this province is Panjshir

2. Disputed Territories

There are differences in interpretation in the placement of Afghanistan's boundary with Pakistan, commonly referred to as the "Durand Line". While the Durand Line is shown on most maps as the western international border of Pakistan, Afghanistan does not recognize it as an official border claiming that it splits ethnic Pashtun areas that should rightfully belong to Afghanistan. However, the Durand Line is considered a "definite, historical de jure boundary", meaning that it is the internationally accepted delimitation. There are no known naming issues that result from the Durand Line dispute.

5. Source Material

Geographic names within Afghanistan have not been standardized by the Afghan authorities. Therefore, there is no single source of names that could be considered official. Instead, a preponderance of evidence, based primarily on Romanized source and backed up by Dari and Pashto dictionaries, is used to determine the most authoritative spelling of each name.

a. Authoritative Maps and Map Series

Fairchild aerial surveys (FAS) topographic line maps are the preferred authorities for topographic map sources. AGCHO has also produced some larger-scale maps that are considered authoritative.

FAS 1:50,000, 1:100,000 and 1:250,000 Series

These maps were produced by the American firm Fairchild during the early 1960s, based on aerial surveys done under contract to the Afghan Government. The names on these maps are displayed in the Yaqubi transliteration system, and require application of the 2007 BGN/PCGN National Romanization System for Afghanistan in Appendix B to derive the approved form.

b. Non-Authoritative Maps and Map Series

While not considered authoritative, the following map sources have been used in name extraction over Afghanistan and are valuable sources when used in concert with authoritative sources.

National Atlas of the Democratic Republic of Afghanistan

This is an atlas of Afghanistan that was co-produced by AGCHO and Poland. It was published in 1984. The names contained in the atlas are considered unverified.

AGCHO Topographic Line Maps

These maps have been produced at scales of 1:25,000, 1:50,000 and 1:100,000, as well as at smaller scales by AGCHO and its predecessors since the late 1970s. AGCHO-produced maps are not currently considered authoritative sources because BGN products are used as the primary source of names data on their maps vice traditional toponymic collection techniques.

Soviet General Staff Maps

These maps were produced at various scales from the 1960s up until the late 1980s during the Soviet occupation of Afghanistan. The names contained in this series of maps are Cyrillic transcriptions.

Iran National Geographic Organization, Various Map Series
1:50,000, 1:100,000 and 1:250,000 scales

Iran National Cartography Center
1:50,000 scale

Survey of Pakistan Topographic Line Maps
1:50,000 and 1:250,000 scales

c. Gazetteers Containing Non-Authoritative Names

There are currently no gazetteer sources that the BGN would consider authoritative name sources. However, there are several gazetteers, specifically AGCHO's villages list from 2006 and CSO's provisional gazetteer, that are considered useful working aids when used in concert with more authoritative sources.

AGCHO 2006 Villages List

This is a gazetteer of names of populated places containing approximately 40,000 entries, reviewed by AGCHO. The names on this list are in Dari or Pashto script and English but are not spatially referenced.

CSO 1975 Provisional Gazetteer of Afghanistan

This is a gazetteer of populated places separated by province. The names on this list are in Dari script and English but not spatially referenced.

Slovar' Geograficheskikh Nazvaniy Afganistana

Names in this gazetteer, produced during the Soviet occupation, are presented in Dari and Pashto script, as well as in Cyrillic and Yaqubi transliterations.

Appendix A. 2007 BGN/PCGN National Romanization System for Afghanistan

BGN/PCGN NATIONAL ROMANIZATION SYSTEM FOR AFGHANISTAN

BGN/PCGN 2007 System

This romanization system agreed by BGN and PCGN in November 2007, accommodates the linguistic complexity of Afghanistan as manifest in its geographical names. The following tabulation shows the original Perso-Arabic script with accompanying Unicode value (columns 1a and b), the Yaqubi romanization⁵ (column 2), the BGN/PCGN romanization with accompanying Unicode value (columns 3a and b), an English phonetic example (column 4), and an example toponym (columns 5a and b).

The tables function as both a romanization system for Afghanistan (i.e. with access to the original script, these tables can be applied to get a standardized Roman result - moving from columns 1 to 3) and as a means of converting the available Yaqubi Roman-script spellings, as appear on the Fairchild Aerial Surveys map series, to standard BGN/PCGN spellings (moving from columns 2 to 3).

⁵ The “Yaqubi” romanization system was developed in 1959 by Muzaffarud Din Yaqubi. It is a native official system designed to reflect Afghan names, both Dari and Pashto, and both pronunciation and genuine linguistic truth.

Consonant Characters

	Script	Unicode value	Yaqubi	BGN/PCGN	BGN/PCGN Unicode value	U.S. English Phonetic	Example	
1	ب	0628	b	b	0062	/b/ as in <i>bob</i>	بَغْلان	Baghlān
2	پ	067E	p	p	0070	/p/ as in <i>pep</i>	پوٲگی	Pōtakay
3	ت	062A	t	t	0074	/t/ as in <i>tall</i>	شیرین تَغاب	Shīrīn Tagāb
4	ٲ	067C	t	ٲ	1E6F	retroflexive /t/	کوت	Kōٲ
5	س	062B	s	s	0073+0304	/s/ as in <i>sand</i>	سَابر	Sābir
6	ج	062C	j	j	006A	/j/ as in <i>jazz</i>	جَلال آباد	Jalālābād
7	چ	0686	č	ch	0063/0068	/ch/ as in <i>chat</i>	چارِگار	Chārīkār
8	ح	062D	h	h	1E29	/h/ as in <i>hot</i>	حَضْرَتِ سُلطان	Ḥazrat-e Sultān
9	خ	062E	kh	kh	006B/0068	/ch/ as in German <i>ach</i>	خوست	Khōst
10	خ ⁶	0681	j	dz	0064+007A	/dz/ as in <i>adze</i>	خَدْران	Dzadrān
11	ځ	0685	c	ts	0074/0073	/ts/ as in <i>gnats</i>	ځوکی	Tsowkēy
12	د	062F	d	d	0064	/d/ as in <i>dog</i>	سپین بولدک	Spīn Bōldak

	Script	Unicode value	Yaqubi	BGN/PCGN	BGN/PCGN Unicode value	U.S. English Phonetic	Example	
13	ډ	0689	d	ډ	1E0F	retroflexive /d/	ډنډ وَ پٲان	Dand wa Patān
14	ز	0630	z	z	007A+0304	/z/ as in <i>zero</i>	گنډرگاه نور	Gužargāh- e Nūr
15	ر	0631	r	r	0072	/r/ as in <i>run</i>	گنډهار	Kandahār
16	ړ	0693	r	ړ	1E5F	retroflexive /r/	انډر	Andar
17	ز	0632	z	z	007A	/z/ as in <i>zero</i>	گنډز	Kunduz

⁶ A variant form ځ is seen infrequently; it does not have a single Unicode encoding.

18	ژ	0698	ž	zh	007A+0068	/zh/ as in <i>azure</i>	میر آسلم ژرندہ	Mīr Aslam Zhrandah
19	ز	0696	ž	zh	007A+0332 0068+0332	retroflexive /zh/	زیرہ	Zhīrah
20	س	0633	s	s	0073	/s/ as in <i>sand</i>	سَمَنگان	Samangān
21	ش	0634	š	sh	0073+0068	/sh/ as in <i>ship</i>	مَزار شَرِیف	Mazār-e Sharīf
22	ش	069A	š	sh	0073+0332 0068+0332	retroflexive /sh/	کِبَنَتہ کلا	Kshêtah Kalā
23	ص	0635	ṣ	ṣ	015F	/s/ as in <i>sand</i>	قَیصار	Qayṣār
	Script	Unicode value	Yaqubi	BGN/PCGN	BGN/PCGN Unicode value	U.S. English Phonetic	Example	
24	ض	0636	ḏ	z	1E95	/z/ as in <i>zero</i>	فَیض آباد	Faīzābād
25	ط	0637	ṭ	ṭ	0163	/t/ as in <i>tall</i>	حَضْرَتِ سُلْطَان	Ḥazrat-e Sulṭān
26	ظ	0638	ẓ	z	007A+0327	/z/ as in <i>zero</i>	ظَاهِر کلا	Zāhir Kalā
27	ع	0639	'	'	2018	no sound	پُلِ عَلم	Pul-e 'Alam
28	غ	063A	gh	gh	0067+0068	no English equivalent / like /r/ in French <i>mer</i>	غَزَنی	Ghaznī
29	ف	0641	f	f	0066	/f/ as in <i>fat</i>	مَزار شَرِیف	Mazār-e Sharīf
30	ق	0642	q	q	0071	no English equivalent / closest sound is guttural /k/	قَیصار	Qayṣār
31	ک	06A9	k	k	006B	/k/ as in <i>kitten</i>	کَنَدَهَار	Kandahār
32	گ	06AF	g	g	0067	/g/ as in <i>gain</i>	گَرْدِز	Gardēz
33	ل	0644	l	l	006C	/l/ as in <i>land</i>	کَابُل	Kābul
34	م	0645	m	m	006D	/m/ as in <i>mint</i>	مَیْمَنَه	Maīmanah
35	ن	0646	n ⁷	n	006E	/n/ as in <i>night</i>	خَان آباد	Khānābād
36	ن	06BC	n ⁸	ṅ	1E49	retroflexive /n/	مَانی	Māṅhēy

⁷ The Fairchild Aerial Survey Maps display *n* for ن instead of *ṅ*.

⁸ The Fairchild Aerial Survey Maps display *ṅ* for ن instead of *n*.

	Script	Unicode value	Yaqubi	BGN/PCGN	BGN/PCGN Unicode value	U.S. English Phonetic	Example	
37	و	0648	w	w	0077	/w/ as in <i>wait</i>	واخان	Wākhān
38	ه	0647	h	h	0068	/h/ as in <i>hot</i>	هرات	Herāt
39	ی	0649	y	y	0079	/y/ as in <i>yard</i>	ينگی قلعه	Yangī Qal'ah

Vowel, Diphthong and Diacritical Characters

1	unwritten	064E	a, â	a	0061	short /a/ as in <i>fat, man</i>	جلال آباد	Jalālābād
2	unwritten	0650	e, i	e, i ⁹	0065+0069	short /i/ or /e/ as in <i>sit</i> or <i>ten</i>	هرات، پل حصار	Herāt, Pul-e Ḥiṣār
3	unwritten	064F	o, u	o, u ¹⁰	006F, 0075	short /u/ or /o/ as in <i>put</i> or <i>no</i>	مُرغاب، کابل	Murghāb, Kābul
4	unwritten	0659	ə, ě	ê	00EA	Pashto vowel /ə/ as in <i>up, jug</i>	گردون	Gêrdōn
	Script	Unicode value	Yaqubi	BGN/PCGN	BGN/PCGN Unicode value	U.S. English Phonetic	Example	
5	آ ¹¹	0622	ā	ā	0101	long /aa/ as in <i>car, fall</i>	آب بند	Āb Band
6	و	0648	ō	ō	014D	long /o/ as in <i>boat, toe</i>	سپين بولدک	Spīn Bōldak
7	و	0648	u, ū	ū	016B	long /oo/ as in <i>moon, rule</i>	بالا بلوک	Bālā Bulūk
8	و	0648	aw or āw	ow	006F+0077	diph. /ow/ as in <i>now, town</i>	جوزجان	Jowzjān
9	ی	06CC (064A)	i or ī	ī	012B	long /ee/ as in <i>meet, see</i>	غزنی، سپين	Ghaznī, Spīn

⁹ Both *e* and *i* are available to romanize this short vowel, depending on local usage and/or root language. In cases where the sound is uncertain, *i* is the default romanization in BGN/PCGN standardization procedures.

¹⁰ Both *o* and *u* are available to romanize this short vowel, depending on local usage and/or root language. In cases where the sound is uncertain, *u* is the default romanization in BGN/PCGN standardization procedures.

¹¹ An *alif* with *mad* (^ˆ) is written only in the initial position by BGN/PCGN standardization procedures, in keeping with Persian language family standards of use of the Arabic alphabet. The same letter written in a medial or final position is written ^ˆ.

10	ی	06CC	ay or ē	ay or ē ¹²	0113, 0061+0079	diph. /ay/ as in <i>date</i> , <i>maid</i>	میوند، ریگ	Maywand, Rēg
11	ې	06D0	ē or ay	ē	0113	Pashto vowel /ā/ as in <i>rain</i> , <i>tail</i>	گردېز	Gardēz
12	ی	06CC	ay or āy	ay or aī ¹³	0113, 0061+0079, 0061+012B	diph. /ī/ as in <i>my</i> , <i>eye</i>	میدان شهر	Maīdān Shahr
	Script	Unicode value	Yaqubi	BGN/PCGN	BGN/PCGN Unicode value	U.S. English Phonetic	Example	
13	آ	06CC	ā	á	00E1	long /aa/ as in <i>car</i> , <i>fall</i>	دندو سُفلی	Dand-e Suflá
14	ی	06CD	e y or ay	êy	00EA+0079	Pashto vowel /əy/ word ending	خوکی	Tsowkêy
15	unwritten	0621	'	'	2019	as in <i>day</i> or <i>yea</i>	هوائی ډگر	Hawāī Dagar
	unwritten	0674	-i-, e, ī	-e, -ye	012B	as in <i>day</i> or <i>yea</i> (glottal stop)	مزار شریف	Mazār-e Sharīf

¹² The character ی should be romanized *ay* or *ē* according to its root language or local pronunciation. In case of uncertainty a reference source (such as the Fairchild Aerial Surveys map series, or a BGN/PCGN approved policy document/list of recommended spellings) should be consulted.

¹³ Both the combination *ay* and *aī* are available to romanize this character according to its root language or local pronunciation. In cases where the sound is uncertain *ay* is the default romanization in BGN/PCGN standardization procedures

Other Diacritical Marks and Language Conventions

1	ای	0627, 06CC	ay or āy	āy	0101	/ī/ as in <i>my</i> , <i>eye</i>	دایکندی	Dāykundī
2	یا	06CC, 0627	ya or yā	īā, iyā or yā ¹⁴	0101, 012B	/īa/ as in <i>Alexandria</i>	زیارت	Ziārat
					OR		غوریان	Ghōriyān
					OR		میا	Myā
	Script	Unicode value	Yaqubi	BGN/PCGN	BGN/PCGN Unicode value	U.S English Phonetic	Example	
3	ا ¹⁵	0627				(no sound)	أرژگان	Uruzgān
4	و	0648	w	w	0648	(no sound)	خواجہ	Khwājah
5	ئ	0626		'	2019	(glottal stop)	درواز پائین	Darwāz-e Pāīn
6	ہ	06C0	. . .h-e			/heh/ as in <i>hay</i>	کوہ صافی	Kōh-e Ṣāfī
7	س	0651	(Consonant doubler)			(no sound)	محمد آغا	Muḥammad Āghā
8	الله ^{see} Note 5	FDF2		Allāh			عبدالله شیلہ	'Abdullāh Shēlah

¹⁴ The character combination ٰ should be romanized *īā*, *iyā* or *yā* according to its root language or local pronunciation. In case of uncertainty a reference source (such as the Fairchild Aerial Surveys map series, or a BGN/PCGN approved policy document or list of recommended spellings) should be consulted.

¹⁵ This convention applies only to the initial *alif* which has no sound of its own, serving as a seat for an unwritten short vowel. This is in keeping with Persian language family standards of use of the Arabic alphabet whereby words that begin or end with an unwritten short vowel must utilize a silent consonant to carry the sound in written form.

NOTES

- 1 Unicode values listed in this table are required for all BGN/PCGN production, to ensure standardization and to minimize confusion from competing representations of a given letter. Notably, the Persian ک (Unicode value 06A9) is prescribed rather than the Arabic ك (Unicode value 0643 or FEDA or FED9), the Persian گ (Unicode value 06AF) is prescribed rather than گ (Unicode value 06AB) or گ (Unicode value 06B0) or ك (Unicode value 0643 or FEDA or FED9), and the Pashto letter خ (Unicode value 0681) is prescribed rather than the *heh* with a dot above and a dot below (Unicode value unknown). For the letter ی in its many variations, care must be exercised to follow this romanization guide's prescriptions to eliminate confusion for search engines and software. BGN/PCGN does not use the letter ی (Unicode value FEEF) in any Afghan word.
- 2 Occasionally the letter sequences كه, زه, سه, and گه occur without intervening vowels. They may be romanized *k-h*, *z-h*, *s-h*, and *g-h* in order to differentiate these romanizations from the digraphs *kh*, *zh*, *sh*, and *gh*, which are used to represent the letters خ, ژ, ش, and غ. Additionally, the Pashto letters خ and خ, routinely romanized *ts* and *dz*, may be alternatively romanized *tš* and *dž* when for special reasons it is desired that confusion be avoided with the character sequences تس (*ts*) and دن (*dz*), respectively.
- 3 The vagaries of written Afghan languages, as pertains to spacing and word division, are addressed as follows. Spaces may be added to or subtracted from Afghan words written in Arabic script, for the purposes of standardization. This is particularly relevant when the words are hand-written, are rendered “artistically”, or express other such non-standard flourishes, as long as the sense of the toponym, word, or phrase is not compromised. Analysts proceed in this regard conservatively.

Romanized toponyms are typically divided into constituent words (spaces and other grammatical rules applied) when those words can stand independently, for purposes of standardization and minimization of confusion, particularly in situations where Afghan writers are inconsistent in their application of spacing and word breaks. When the Afghan word or suffix is only used in combination with other nouns or adjectives, then it should be appended to the preceding word in its romanization. This includes (but is not limited to) *-ābād*, *-zaī*, *-zādah*, *-tū*, *-wand*, *-gaī*, *-kaī*, *-pūr*, *-dēsh*, *-lar*, *-lī*, *-lū* and *ullāh*, as, for example, seen in *Raḥmatābād* (رحمت آباد) and *Raḥmatullāh* (رحمت الله), but *Raḥmat Khēl* (رحمتخیل) and *Raḥmat Shahr* (رحمتشهر).
- 4 The one-letter words و (Pashto) and و (Dari) are romanized *dê* and *wa*, respectively.
- 5 The word الله, meaning God, should always be romanized *Allāh*, except as specified in note 3. Note that the Unicode value FDF2 spells *Allāh*, but omits the *alif* in some common fonts, including Times New Roman. If in doubt, try in Arial Unicode MS to verify. Also note that the “dagger alif” (◌ِ) above the second ل (lām) in the word الله, is not written but should be romanized *ā*, like a full-size alif.
- 6 In names of Arabic origin, the *l* of the definite article *al* is assimilated before the ‘sun letters’ *t*, *ṣ*, *d*, *z*, *r*, *z*, *s*, *sh*, *ṣ*, *ẓ*, *ṣ*, and *n*. In its romanization, the article should be separated from the name it precedes and should not be capitalized except at the beginning of a name, e.g. جبل السراج → *Jabal as Sarāj*

- 7 Specifically Arabic names, particularly those of Islamic significance, often include a consonant which would ordinarily be written with a *tashdīd* (*shaddah*) in Arabic, which may be romanized by writing the equivalent consonant twice, e.g., مُحَمَّد → *Muḥammad*. This rule is particularly in force if the word is commonly romanized with a doubled consonant by Afghan naming authorities. For example, *ḥājī* is rarely seen in Afghan-produced romanizations with a doubled consonant, but *‘Abbās* is. The doubled y consonant is almost always seen thusly, as in *Sayyid* or *Qayyūm*.
- 8 In Afghan names which contain an *izāfah*, it should be romanized as *-e* or *-ye* according to common pronunciations, but generally, *-e* is used if the preceding word ends with a consonant other than silent *heh*, and *-ye* if the preceding word ends with a vowel sound. Scholarly sources indicate that *heh* is silent in *darah* and *qal‘ah* (thus *darah-ye*, *qal‘ah-ye*), but lightly spoken in *kōh* and *chāh* (thus *kōh-e*, *chāh-e*).
- 9 The character sequence خو, where followed by ۱ or ۱ should be Romanized *khwā* or *khwī*, where the *w* is not pronounced, or only weakly so, as in *khwājah*.
- 10 Plural nouns ending in *-hā* or *-ān* should always be romanized as a single word, regardless of whether a space appears in a Perso-Arabic script source.
- 11 Although Arabic script is written from right to left, numerical expressions, e.g. ١٩٦٨ → 1968 are written from left to right.

٠	١	٢	٣	٤, ٤	٥, ٥	٦, ٦	٧	٨	٩
0	1	2	3	4	5	6	7	8	9

Appendix B. Glossary of Generic Terms

Generic	Feature Designation Name	Feature Designation Code
Āb	STM	stream
Algaḍ	STM	stream
Bāgh	GDN	garden(s)
Bāgh	OCH	orchard(s)
Band	DAM	dam
Band	HLLS	hills
Band	MTS	mountains
Band	RDGE	ridge(s)
Bāzār	MKT	market
Chāh	WLL	well
Chahārrāhī	RDJCT	road junction
Chashmah	SPNG	spring(s)
Chīnah	SPNG	spring(s)
Chīnah	WLL	well
Dahān	VAL	valley
Dahānah	STMM	stream mouth(s)
Ḍanḍ	LK	lake
Darah	CNYN	canyon
Darah	GRGE	gorge(s)
Darah	RVN	ravine(s)
Darah	STM	stream
Darah	VAL	valley
Daryā	STM	stream
Dasht	DSRT	desert
Dasht	PLN	plain(s)
Daṣht	DSRT	desert
Daṣht	PLN	plain(s)
Daṣhtah	DSRT	desert
Daṣhtah	PLN	plain(s)
Dastgāh	MFG	factory
Gardan	PASS	pass
Gardanaḥ	PASS	pass
Ghar	HLL	hill
Ghar	MT	mountain
Ghārah	HLL	hill
Ghārah	MT	mountain
Ghūnd	HLL	hill
Ghūnd	MT	mountain

Ghundêy	HLL	hill
Ghundêy	MT	mountain
Ghundī	HLL	hill
Ghūndī	MT	mountain
Gīrō	SLP	slope(s)
Hadīrah	CMTY	cemetery
Ḥiṣār	FT	fort
Ḥowz	LK	lake
Ḥowzak	PND	pond
Jādah	ST	street
Jamhūrī	PCLI	independent political entity
Jangal	FRST	forest(s)
Jar	STM	stream
Jar	VAL	valley
Jazīrah	ISL	island
Jōy	CNL	canal
Jōy	DTCH	ditch
Jōy	STM	stream
Kamp	CMP	camp(s)
Kandah	STM	stream
Kandow	GRGE	gorge(s)
Kandow	PASS	pass
Kārēz	CNL	canal
Kārēz	STM	stream
Kawīr	DSRT	desert
Khānah	HSE	house(s)
Kharābah	PPLQ	abandoned populated place
Kharābah	RUIN	ruin(s)
Khēl	TRB	tribal area
Khūlah	STMM	stream mouth(s)
Khwaṛ	STM	stream
Khwaṛah	STM	stream
Kōh	HLL	hill
Kōh	MT	mountain
Kōhhā	HLLS	hills
Kōhhā	MTS	mountains
Kōl	LK	lake
Kōl	PND	pond
Kōtal	PASS	pass
Ma'dan	MN	mine(s)
Madrasah	SCH	school
Maḥal	AREA	area
Maḥal	LCTY	locality

Maḥbas	PRN	prison
Maīdān	AREA	area
Maīdān	FLD	field(s)
Maktab	SCH	school
Māndêh	STM	stream
Māndêh	VAL	valley
Māndêh	WAD	wadi
Manṭaqah	AREA	area
Manṭaqah	LCTY	locality
Maqbarah	SHRN	shrine
Masjid	MSQE	mosque
Mazār	SHRN	shrine
Mazār	TMB	tomb(s)
Nahr	CNL	canal
Nahr	STM	stream
Narāī	PASS	pass
Nāw	STM	stream
Nāw	VAL	valley
Nāwah	STM	stream
Nāwah	VAL	valley
Nāwêr	LKI	intermittent lake
Nāwêr	PND	pond
Pārk	PRK	park
Pāygāh	MILB	military base
Paytow	SLP	slope(s)
Pitāw	SLP	slope(s)
Pul	BDG	bridge
Pusht	HLL	hill
Pusht	RSV	reservoir(s)
Pushtah	HLL	hill
Pushtah	MT	mountain
Pushtah	MTS	mountains
Qabr	GRVE	grave
Qabr	TMB	tomb(s)
Qabristān	CMTY	cemetery
Qal'ah	CSTL	castle
Qal'ah	FT	fort
Qanāt	CNLSB	underground irrigation canal(s)
Qōl	STM	stream
Qōl	VAL	valley
Rāh	PASS	pass
Rāh	RD	road
Rāh	TRL	trail

Rēg	DUNE	dune(s)
Rēg	SAND	sand area
Rōd	STM	stream
Rōd	WAD	wadi
Sang	MT	mountain
Sang	RK	rock
Sar	HLL	hill
Sar	MT	mountain
Sar	PK	peak
Sar	RDGE	ridge(s)
Shēlah	STM	stream
Silsilah	HLLS	hills
Silsilah	MTS	mountains
Tahānah	PP	police post
Tahānah	PSTB	border post
Tahānah	PSTP	patrol post
Tāḥah	PP	police post
Tāḥah	PSTB	border post
Tāḥah	PSTP	patrol post
Tang	GRGE	gorge(s)
Tang	PASS	pass
Tangay	GRGE	gorge(s)
Tangay	PASS	pass
Tangay	STM	stream
Tangay	VAL	valley
Tangī	GRGE	gorge(s)
Tangī	PASS	pass
Tangī	STM	stream
Tangī	VAL	valley
Tapah	HLL	hill
Tapah	MT	mountain
Tōnal	TNL	tunnel
Tōnal	TNLC	canal tunnel
Tōp	HLL	hill
Tōp	MT	mountain
Tōp	PK	peak
Tōy	STM	stream
Tōy	VAL	valley
Tsāh	WLL	well
Wādī	STM	stream
Wādī	WAD	wadi
Wālêh	CNLA	aqueduct
Wālêh	DTCHI	irrigation ditch

Wilāyat	ADM1	first-order administrative division
Wuluswālī	ADM2	second-order administrative division
Yaylāq	CMP	camp(s)
Zhawêr	STM	stream
Zīarat	SHRN	shrine
Zīarat	TMB	tomb(s)

Afghanistan ADM1 Boundaries

