

Transcription of Afghan Placenames to an English-style Romanization*

FREDERICK E. WILSON

The currently-used transliteration systems employed by American cartographers are inadequate for the rendition of Afghan place names. Although Dari (the official language of Kabul, the capital of Afghanistan) is written in a modified Arabic script and is linguistically related to Persian (Farsi), transliteration systems based on either Arabic or Persian are inherently inadequate for Dari. It is here suggested that a transcription system based on the Dari language be employed by American cartographers. In this article such a system is proposed. The advantages of this system are that the place names correspond in transcription to the actual Dari-Afghan pronunciation of them and that computerized transcription provides consistent Latin equivalents of Afghan place names.

The two major languages of Afghanistan are Pashto (Pushtu, Push-tun, Pakhtu, Pakhtun) and Dari (Tadzhik).¹ Dari, like Pashto, is linguistically related to Persian and other Iranian languages (ex. Kurdish, Beluchi, Bakhtiari, Tadzhik) but not to Arabic, although both Dari and Persian have borrowed heavily from Arabic. Dari additionally has

¹ The official name of the Afghan language changed from Farsi to Dari in 1968. Janet Mueller Bing: "Linguistic Rhythm and Grammatical Structure in Afghan Persian." *Linguistic Inquiry* 11 (1979), p. 437.

many Pashto borrowings. The sound systems of Dari, Persian and Arabic are distinct. For example, the word “teacher” in Dari is “maelim

معلیم”; in Persian is “mo‘allem مُعَلِّم”; and in Arabic is “ma‘allem مَعَلِّم”. The word “first” in Dari is “awal اول”; in Persian is “avval اَوَّل”; and in Arabic is “’awwal اَوَّل”. The word “fruit”

in Dari is “mewa میوه”; in Persian is “mive میوه”; and in Arabic

has changed to a different word entirely. Although Pashto is the language of the majority of Afghans, Dari is the language used in the area around the capital. As such, Kabuli Dari is employed on most maps of Afghanistan. A transcription system must therefore include both Dari and Pashto sounds.

The difficulties in transcription of Afghan place names are analogous to those involving the place names of other non-English speaking countries. A first analogy involves countries using Latin scripts. As a rule, when place names of a country using the Latin alphabet appear on English-language maps, they retain their original spellings. This approach has numerous advantages including standardization of place names. The chief disadvantage is that precise pronunciation of these place names is not achievable without knowledge of the rules of pronunciation in many languages. For example, the sound “j” in English is represented in various ways in other languages which use the Latin alphabet: “dzi” in Polish; “gi” in Italian; “xh” in Albanian; “dj” in French; and “dsch” in German. Conversely, a “j” is pronounced quite differently in English, German, Spanish, and French. The same problem occurs in languages using the Arabic script. For example, the sound “s” is represented in Arabic by “س” and in Persian, Pashto, or Dari by

“س”, “س”, or “سپڼک.” Conversely, a “و” represents a “w” in

Arabic and Dari, but a “v” in Persian. An Arabic speaker then would have approximately the same difficulty with Dari place name pronunciation as would an English speaker with French place names. In addition, place names of areas using the Arabic script never retain their original spellings in Latin transliteration due to the major transformation in script which occurs. As such, there is no advantage in maintain-

ing a single system of transliteration of all Arabic script languages. Therefore, separate transcription schemes for each Arabic script should be developed to make adequate pronunciation available to English readers.

A second analogy involves Cyrillic scripts. When place names of countries using Cyrillic alphabets are transliterated to English, the Russian form of the original language is often used. For example the Bulgarian town “*Плѐвн*” is transliterated as “Pljeven,” which is a Russian pronunciation. The correct Bulgarian transliteration is “Pleven.” Another example is from Moldavian, a language very similar to Romanian. Moldavian was written in the Latin script prior to Soviet annexation, but the modern representation of the Moldavian capital of Chisinau now comes through the Russian “*Кишинев*” to become “Kishinev.” Likewise, many Central Asian place names have been modified on maps since the scripts were changed from Arabic to Cyrillic. For example, Merv has become the Russified “Mary,” and Tashkend has become “Tashkent.” Similarly, many Persian place names appear in Latin using transliteration systems designed for Arabic (ex. Eşfahān for Esfahaan). Afghan place names appear using systems designed for Persian (ex. Shinvar for Shinwaar). This is analogous to creating Latin transliterations of Bulgarian and Cyrillified Uzbek or Moldavian place names using the rules for Russian pronunciation. Persian is not pronounced as is Arabic, nor is Dari pronounced exactly as is Persian.

A third analogy involves places with “standard” names. For example, the Latin representation “Moscow” comes from a Polish name, not from the Russian “*Москва*.” “Kiev” is from the Russian name of the city, not the Ukrainian. Similarly, “Warsaw” is a German word and “Prague” is French. In addition, some places have no “standard” name: the Yugoslav town of Skopje is “Skopje” in Macedonian, “Skoplje” in Serbian, “Shkuf” in Albanian, and “Üsküb” in Turkish, four major languages of Yugoslavia. Similarly, “Kabul,” “Kandahar,” and “Mazari Sharif” are colonial British names of Afghan places which are not precisely related to either their spelling or their pronunciation.

To date, no standard transcription system for Afghan place names exists, although use of the FARSI BGN/PCGN 1958 system is becoming more widespread.² Many systems are confusing because they are

² U.S. Board on Geographic Names and Permanent Committee on Geographic Names for British Official Use. U.S. Army Topographic Command. Geographic Names Division. *Afghanistan Official Standard Names Gazetteer*. (Washington, D.C.: GPO, 1971), p. i.

designed for letter-by-letter transcription of Arabic letters to Latin script rather than of Dari sounds to Latin script. Letter-by-letter transliteration is undesirable because of the unnecessary distinctions it draws among allophones—letters which all have the same sound in Dari—and because several Arabic letters and signs, such as “ع” and “ف”, have quite different functions in Dari.

The Latin alphabet has been used for languages as dissimilar as English, modern Turkish, and Malay. Similarly, the Arabic alphabet has been used for writing many unrelated languages—Arabic, Turkish, Persian, Central Asian and Caucasian languages up to 1929, and Urdu. There is no reason other than precedent to prefer Arabic to Latin script for Dari. In fact, the Latin alphabet of English serves very well to represent the sounds of Dari, and without need for diacritical marks.

First, a discussion of existing systems investigates their inadequacies, then a new system is proposed to correct these disadvantages. The existing systems have been divided into four categories for discussion.

Table I lists the atlases, maps, and transliteration schemes studied with their categories. These sources were drawn from the collections of the University of Oregon and the University of Nebraska Atlas of Afghanistan map collection. In all, 24 maps and atlases were found, and all were included in this study. These sources appear fairly representative in their treatment of Afghan place names when compared to written sources also employing transcribed Afghan place names. Except for the FARSI BGN/PCGN 1958 system (number 27 in Table I), none of the transliteration systems was designed for cartographic applications.

Category 1. In this category, Afghan place names are written using historically-evolved “standard” names, which may not match Afghan pronunciations. This is not actually a transcription system, since no consistently-applied rules are used to change names from Arabic to Latin script. The familiarity of the result is an advantage of this system. People are used to seeing “کابل” written as “Kabul.”

Category 2. Afghan place names are written using an official system of transliteration, such as the FARSI BGN/PCGN 1958 system or the PASHTU BGN/PCGN 1968 system, adopted by the United States and Great Britain. They produce accurate Latin equivalents of place names on a letter-by-letter basis. In these systems, many diacritical marks and unfamiliar letter values are required. Maps using Category 2 methods are far less readable than are those using Category 1 methods, since

knowledge of an involved representation of sounds and recognition of unfamiliar diacritical marks around the letters is required. And, the transliteration system used is rarely indicated on the map. Groups of allophones in Dari are unnecessarily written using unique symbols in most Category 2 maps. For example, the Dari letters ز, ذ, ظ, and ن can all be represented by the Latin letter "z." Likewise, ط and ت can be represented by "t," and س and س by "s." But in Category 2 maps, eight diacritically-marked Latin letters rather than three uninflected letters are usually used for these Dari letters. Since Dari pronunciation is different from Persian, maps produced using sources such as the *Afghanistan Official Standard Names Gazetteer* give Persian pronunciations, not Afghan.³

Category 3. Afghan place names are transcribed, not transliterated, into Latin forms. Transliteration is a one-to-one replacement of Dari letters by Latin ones, and transcription is a one-to-one replacement of Dari sounds (usually phonemes) by English-style ones. In Category 3, diacritical marks are used on the maps. Again, maps using these marks are less easily understood because no key is given for interpreting them and the marks may be difficult to read after map reduction.

Category 4. Afghan place names are transcribed to another language based on how they are pronounced by educated Afghans. No unfamiliar diacriticals are used. A system like this should be easy to develop for English because non-Pashto Dari has only four sounds English lacks: the

consonants خ (χ in the International Phonetic Alphabet), غ (ʁ), ر (r), and ق (q). But, there are no maps in this category in English.

However, maps of Afghanistan printed in Russian and German use alphabets and letter-sound relationships familiar to general audiences

³ *Afghanistan Official Standard Names Gazetteer*. p. v-xvi.

in those languages. This category offers an immediately-readable, consistently-produced set of Dari place names in another language's script, and is apparently the method of preference in the U.S.S.R. and in West Germany.

For general audiences, the most useful transcription system would represent educated Kabuli Afghan pronunciation using spellings which can be pronounced almost as they are in English, and having objective rules for transliterating Dari and Pashto to English-style place names. Such a system has been developed and has been automated to facilitate the conversions.

To develop this system, the author interviewed Dr. Ishmael Burhan, a native speaker of Kabuli Dari and Pashto and professor of Kabuli Dari at the University of Nebraska at Omaha. An earlier excellent, informal transliteration scheme by Burhan provided the initial stimulus for this work. The author tape-recorded Burhan's pronunciation of each fully inflected place name given on a Dari-language map of Afghanistan.⁴ Empirical rules of letter combination were worked out and included in a short, easily modified FORTRAN V computer program in the form of a 46 X 2 array (see Table III).

To use the system, one transliterates a fully inflected Dari or Pashto word into a computer code (Table II). The computer code, when entered into the program, produces an English-style transcription of the Afghan place name (Tables III and IV). These transcriptions use no diacriticals, and they follow absolute rules of pronunciation (Table V). Additional advantages of this system are that the pronunciations are those of a Kabuli Afghan, not a Tehrani Persian speaker, and that the system is automated for rapid, consistent results.

There also exist limitations to the system proposed. Once transcribed to English, the place name cannot be retranscribed back to Dari. The process is one-way. Also, this system appears very different from the FARSI BGN/PCGN 1958 system, artificially making Persian appear unrelated to Dari. Perhaps a systematic approach for all Iranic-language transcriptions is now needed.

⁴ *Afghanistan*. (Kabul: Afghanistan Government Printing Office, 1958.) (In Dari)

TABLE I
ATLASES AND MAPS OF AFGHANISTAN
AND TRANSLITERATION SYSTEMS INVESTIGATED

Atlases and maps:

1. *Economical Map of Afghanistan*⁵ Category 1
2. *Archaeological Map of Afghanistan*⁶ Category 1
3. *Kabul*⁷ Category 1
4. *Physical Map of Afghanistan*⁸ Category 3
5. *Map of Afghanistan*⁹ Category 1
6. *Tourist Map of Afghanistan*¹⁰ Category 3
7. *Allgemeiner Handatlas*¹¹ (in German) Category 4
8. *The Times Atlas of the World: Mid-Century Edition*¹² Category 1
9. *Kabul Phases of Urban Development*¹³ Category 3
10. *Encyclopaedia Britannica World Atlas*¹⁴ Category 1

⁵ Afghan Cartographic and Cadastral Survey Institute. *Economical Map of Afghanistan*. (Kabul: ACCSI, 1967)

⁶ Afghan Cartographic and Cadastral Survey Institute. *Archaeological Map of Afghanistan*. (Kabul: ACCSI, 1970)

⁷ Afghan Cartographic and Cadastral Survey Institute. *Kabul*. (Kabul: ACCSI, 1967.)

⁸ Afghan Cartographic and Cadastral Survey Institute. *Physical Map of Afghanistan*. (Kabul: ACCSI, 1975)

⁹ Afghan Tourist Organization. *Map of Afghanistan*. (Kabul: Education Press, Franklin Book Program, 1965)

¹⁰ Afghan Tourist Organization. *Tourist Map of Afghanistan*. (Kabul: ACCSI, 1975)

¹¹ Ernst Ambrosius. *Andrees Allgemeiner Handatlas*. (Bielefeld und Leipzig: Verlag Von Velhagen und Klasing, 1930) p. 155.

¹² John Bartholomew, ed. *The Times Atlas of the World: Mid-Century Edition*. (Boston: Houghton-Mifflin, 1955) v. 2, pl. 31.

¹³ *Kabul Phases of Urban Development*. (Bonn: Bonn Univ. Press, 1967)

¹⁴ G. Donald Hudson. *Encyclopaedia Britannica World Atlas*. (Chicago: Encyclopaedia Britannica, 1949), pp. 48-49.

11. *National Geographic Atlas of the World*¹⁵ Category 3
12. *Pergamon World Atlas*¹⁶ Category 3
13. *General Atlas of Afghanistan*¹⁷ (in Farsi, English, French) Category 3
14. *Afghanistan Economic Map*¹⁸ Category 1
15. *The Times Atlas of the World: Comprehensive Edition*¹⁹ Category 1
16. *Afghanistan*²⁰ (AMS, 1942) Category 3
17. *Afghanistan*²¹ (AMS, 1969) Category 2
18. *Afghanistan*²² (CIA, 1971) Category 2
19. *Afghanistan*²³ (CIA, 1972) Category 2
20. *Southwest Asia: Preliminary Edition (Iran, Afghanistan, Pakistan)*²⁴ Category 2
21. *Afghanistan: Telegraph and Telephone Communications*²⁵ Category 2
22. *Roads of Afghanistan*²⁶ Category 2

¹⁵ Melvin M. Payne, ed. *National Geographic Atlas of the World*. 4th ed. (Washington, D.C.: National Geographic Society, 1975) p. 134.

¹⁶ Poland. Wojsko Polskie. Służba Topograficzna. *Pergamon World Atlas*. (Oxford: Pergamon Press, 1968) pp. 261-62.

¹⁷ Abbas Sahab, ed. *General Atlas of Afghanistan*. (Tehran: Sahab Geographic and Drafting Institute, 1970) In Farsi, French, English.

¹⁸ *Afghanistan Economic Map*. (Tehran: Sahab Geographic and Drafting Institute, 197?).

¹⁹ *The Times Atlas of the World: Comprehensive Edition, 1980*. (New York: Times books, 1980) pl. 31.

²⁰ U.S. Army Map Service. *Afghanistan*. (Washington, D.C.: US AMS, 1942).

²¹ U.S. Army Map Service. *Afghanistan*. (Washington, D.C.: US AMS, 1969).

²² U.S. Central Intelligence Agency. *Afghanistan*. (Washington, D.C.: US CIA, 1971).

²³ U.S. Central Intelligence Agency. *Afghanistan*. (Washington, D.C.: US CIA, 1972).

²⁴ U.S. Central Intelligence Agency. *Southwest Asia: Preliminary Edition (Iran, Afghanistan, Pakistan)*. (Washington, D.C.: US CIA, 1980).

²⁵ U.S. Office of Strategic Services. *Afghanistan: Telegraph and Telephone Communications*. (Washington, D.C.: US OSS, 1942).

²⁶ U.S. Office of Strategic Services. *Roads of Afghanistan*. (Washington, D.C.: US OSS, 1942).

23. *Afganistan*²⁷ (in Russian) Category 4
 24. *Afganistan*²⁸ (in Russian) Category 4

Transliteration systems:

25. Allworth, 1971²⁹ Category 2
 26. Chavarria-Aguilar, 1962³⁰ Category 2
 27. *Afghanistan Official Standard Names Gazetteer*³¹ Category 2
 28. *Persian Romanization*³² Category 2
 29. *Pushto Romanization*³³ Category 2

²⁷ U.S.S.R. Glavnoe Upravlenie Geodezii i Kartografii pri Sovete Ministrov SSSR. *Afganistan*. Serial B-2227. (Moscow: GUGK, 1978) In Russian.

²⁸ U.S.S.R. Glavnoe Upravlenie Geodezii i Kartografii pri Sovete Ministrov SSSR. *Afganistan*. Serial B-2227. (Moscow: GUGK, 1982) In Russian.

²⁹ Edward Allworth. *Nationalities of the Soviet East: Publications and Writing Systems. A Bibliographical Directory and Transliteration Tables for Iranian and Turkic-Language Publications, 1818-1945*. (New York: Columbia Univ. Press, 1971).

³⁰ Oscar Luis Chavarria-Aguilar. *Pashto Basic Course*. (Ann Arbor: Univ. of Michigan Press, 1962).

³¹ *Afghanistan Official Standard Names Gazetteer*. pp. v-xvi.

³² U.S. Library of Congress—Processing Department. Cataloging Service. *Persian Romanization*. Bull. 92. (Washington, D.C.: GPO, 1970).

³³ U.S. Library of Congress—Processing Department. Cataloging Service. *Pushto Romanization*. Bull. 93. (Washington, D.C.: GPO, 1970).

TABLE II
ENCODING TABLE TO CHANGE DARI OR PASHTO TO COMPUTER INPUT

Dari Letter ^a	Computer Code	Dari Letter ^a	Computer Code
ا	A1	ف	F
آ	A2	ق	Q
ب	I	ک	K
پ	U	گ	G
ت	B	ل	L
ث	P	م	M
ج	T	ن	N
چ	S1	و	W
ح	J	ه, ه, د	H
خ	C	ی	Y
د	H	ی	I
ذ	X	dash above letter	A2
ر	D	dash below letter	I
ز	Z1	و above letter	U
س	R	(sukuun) .	nothing
ش	Z1	لا	LA1
ص	Z2	لا	LA2A1
ض	S1	ت	T
ط	S2	س	S3
ظ	S1	د	D
ع	Z1	ز	Z3
غ	T	س	S2

(Continued)

TABLE II CONTINUED

Dari Letter ^a	Computer Code	Dari Letter ^a	Computer Code
ظ	Z1	ن	N
ع	A3	ک	KA1
غ	G2	ر	R

Sample Afghan place names and computer inputs:

Place name	Computer input
بدخشان	BA2DA2XS2A2A1N
پکتیا	PA2KTIYA2A1

^a Afghan letters shown only in their isolated forms in this sample Table.

TABLE III
TABLE OF CORRESPONDENCES BETWEEN COMPUTER CODES
AND OUTPUT ROMANIZED DARI LETTER GROUPS^a

Computer Code	Roman Letters	Computer Code	Roman Letters
A1	AA	L	L
A2	A	M	M
I	I	N	N
U	U	W8	O
B	B	A2A1	AA
P	P	UW	OO
T	T	W	OO
S1	S	Y	EE
J	J	YA2A1	YAA
C	CH	WA2A1	WAA
H	H	IY	EE
X	KH	IYA2A1	IYAA
D	D	A2Y	AY
Z	Z	UWY	UWEE
R	R	WA2H	WAH

(Continued)

TABLE III CONTINUED

Computer Code	Roman Letters	Computer Code	Roman Letters
Z1	Z	initial Y	Y
Z2	ZH	A2A3A1	AA
S2	SH	A3A1	AA
A3	A	A1Y1	AAZI
G2	GH	S3	DZ
F	F	Z3	Z
Q	Q	WA1	AA
K	K	G	G

^a In the actual program, these letter groups reside in a search array.

TABLE IV
SAMPLE INPUT-COMPUTER CODES
AND OUTPUT ROMANIZED DARI PLACE NAMES

Computer Code	Place name	Computer Code	Place name
KA1BUL	KAABUL	QA2NDA2HA2A1R	QANDAHAAR
PA2KTIYA2A1	PAKTIYAA	QA2LA2A1T	QALAAAT
G2A2ZNY	GHAZNEE	XA1NA1BA1D	KHAANAABAAD
NA2NGA2RHA2A1R	NANGARHAAR	DUWS2IY	DOOSHEE
BA2DA2XS2A2A1N	BADAKHSHAAN	S2YXA1LY	SHEEKHAALEE
MA2ZA2A1RIS2A2RYF	MAZAARISHAREEF	S1URXWPA2A1RS1A2A1	SURKHOOPAARSAA
A2FG2A2A1NIS1TA2A1N	AFGHAANISTAAN	PA1RWA2A1N	PAARWAAN
KIS2IM	KISHIM	XUWS1T	KHOOST
JURM	JURM	LHUWGA2RD	LHOOGARD
CA2A1HIA1B	CHAAHIAAB	PA2G2MA2A1N	PAGHMAAN
RUS1TA1Q	RUSTAAQ	A2CIYN	ACHEEN
TA2A1LUQA2A1N	TAALUQAAN	S2INWA2A1R	SHINWAAR
QUNDUWZ	QUNDOOZ	DA2RA2YPA2YC	DARAYPAYCH

TABLE V
PRONUNCIATION GUIDE FOR ROMANIZED AFGHAN PLACE NAMES

Most letters have their English values. The exceptions are:

R is always trilled as in Spanish "pero"

Q is always an uvular "k" as in "Iraq"

GH is always an uvular "r" as in French "Paris"

KH is always an aspirated "kh" as in Scottish "loch"

G is always hard as in "go"

The vowels and vowel combinations are always pronounced as follows:

A as in "calm"

I as in "bit"

U as in "but"

AY as in "pay"

AA as in "paw"

EE as in "see"

O as in "go"

OO as in "food"

Diphthongs are always pronounced as follows:

CH as in "church"

SH as in "shine"

ZH as the "s" in "pleasure"
