

**Twenty-third Session
Vienna, 28 March – 4 April 2006**

**Item 10 of the Provisional Agenda:
Activities relating to the Working Group on
Romanization Systems**

Transcription Procedure for Iranian Toponymic Items

Transcription Procedure for Iranian Toponymic Items

2-1- Foreword

Seeking for a reliable scientific method for registering toponymic items on basic 1/25,000 maps, the Iranian National Cartography Center (hereinafter, NCC), entered negotiations with Institute of Geography, affiliated to University of Tehran, in 1987. This led to the establishment of a joint committee, attended by specialists in geographical sciences, linguistics and history, as well as by representatives of Directorate of Cartography and Geographical Information Documentation Department, both affiliated to NCC. After one year of hard work, the body, titled "Committee for Romanization of Iranian Geographical Names", succeeded in preparing a notebook titled "Procedure for Transcribing Iranian Geographical Names", which was immediately approved and utilized by NCC.

The Transcription Working Group for Geographical Names, a newly established branch of the Specialist Committee for Romanization and Standardization of Geographical Names, has been authorized, in the new phase of the activities by NCC, to complete the job of scientific registration of Iranian toponymic items. In a series of sessions which followed, the point was emphasized by this author that the outcome would not be scientific unless three scripts of every item could be prepared:

- A *narrow transcription*, for inclusion in the database for narrow transcriptions of geographical names, to be utilized by interested researchers, especially, specialists in linguistics;

- A *broad transcription*, for inclusion in the database for broad transcriptions of geographical names, to be utilized, in actual registration of toponymic items on maps as well as in all other applications thereof;

- A *transliteration*, for inclusion in the database for transliterations of geographical names, to be utilized by researchers interested in their history, as well as any other diachronic research on them.

The proposal was approved and the working group undertook to prepare the needed scientific and practical procedures. The procedures are now available.

Among these three systems and procedures, the second one, the, so called, "broad transcription system and procedure", considers, as its goal, all practical needs at the national and international levels. The present notebook comprises a full justification for it.

2-2- Characteristics of Persian Writing System

The Persian writing system, like all of its other counterparts worldwide, can not be relied upon as a full-fledged candidate for ideal transcription. Several indispensable idiosyncrasies have been reported to be responsible for this:

- In Persian writing system, several letters are distinguishable which stand in correspondence with more than one speech-sound (e.g. چ^a represents both [i] and [y]; ح^b stands for both [e] and [h]; and و^a is distinguished as either [u], [o] or [v]).

- On the other hand, there are several groups of other letters of Persian alphabet with their members, all, referring to the same speech-sound (e.g. the four letters ژ^a , ز^a , ژ^a and ز^a , all stand in correspondence with [z]; and the three letters س^a , س^a and س^a with [s]).

- Half of the Persian vowels are not (normally, at least) represented in Persian writings. The speech-sounds [æ], [e] and [o], are usually guessed while reading scripts in Persian, although there are special diacritics for them. This, normally, results in ambiguities for non-native readers (e.g. the word گورد^a can be read as either [gærd] "powder", [gerd] "round" or [gord] "hero").

- Few of the obligatorily used letters in Persian historical spellings are not read (e.g. the letter و^a , normally standing in correspondence with either [u], [o] or [v], is not read in words like خاهر^a [xahar] "sister").

- Most of the Persian characters change their shape in different collocations (e.g. the first letter in گه^a "cave", the second letter تغیر^a "change", the final letter in گهرا^a "scream", and the final letter in باغ^a "garden" are different contextual forms of the same letter.

- Several subsets are distinguishable among Persian letters of alphabet which are minutely different from each other. Normally, the presence or absence of one dot or one small part, above or under a letter, can change it into another one (e.g. ب^a standing for [b], پ^a standing for [p], ت^a standing for [t], and س^a standing for [s]; or ک^a standing for [k] and گ^a standing for [g]).

To these one must add the two facts that the Persian writing system is not Latin-based and that it is written from right to left: these two characteristics are not in line with the strategy accepted at the international level to use romanized systems for transcriptions and transliterations.

The abovementioned considerations, among others, led the Transcription Working Group to devise and adopt a special transcription system for the toponymic items it had to deal with. The system, however, was not the only one. The whole literature covers, in addition, other works, each with its own shortcomings.

2-3- Literature

Registration of geographical names in Iran has a relatively long history. Nevertheless, very few of the works can be described as truly scientific and comprehensive. The rest, done either by lexicographers or cartographers, comprise short-term efforts to solve, in a way, the problem of more than one readings for given names; e.g. through devising a phonetic alphabet, by way of using homophones, or through finding words of the same rhythm. These, all, had their own deficiencies.

Among these, "Procedure for Transcribing Geographical Names", written by this author in 1987, and approved by Institute of Geography, affiliated to the University of Tehran, for use by NCC, was the only one in which, for the first time, the scientific method had been utilized for updating and optimizing the procedures proposed by its predecessors. The same system and procedure were approved in 2003, by the Transcription Working Group, to be

used as the system and procedure for preparing broad transcriptions of toponymic items be NCC. What follows is detailed description of this system.

2-4- Methodological Justifications

In the Foreword above (cf. 2-1), the idea of registering the geographical names in three ways (in the form of, so called, narrow transcriptions, broad transcriptions and transliterations) to guarantee a scientific approach was introduced. There is, however, a methodological justification for this approach, too. Today's Iranian community consists of a collection of different ethnic groups, each with its own language, dialect or, at least, accent. Numerous communities of Kurds, Turks, Lors, Turkmen, Baluches, Arabs and Armenians, among others, with their own different languages, live within the borders. In addition to these, there are numerous other communities each with a different dialect of Persian; Gilaks, Mazandarannies, Tats, Taleshes, Zabolies, and Sorkhe'ies being but a few examples. Obviously enough, no full rendering of even the most rudimentary pronunciation characteristics of these can be established within a single transcription system. In addition, no full-fledged broad transcriptions of their utterances would be obtainable unless descriptions of their phonological systems precede.

It has been for this same reason that narrow transcriptions of toponymic items have always been emphasized by this author. In addition to what was mentioned above in relation to the scientific value of a narrow transcriptions database, this methodological obligation must be emphasized. For the purposes of this procedure, the International Phonetic Alphabet (IPA), together with all its linguistically significant strategies, was selected as the system for preparing narrow transcriptions of Iranian toponymic items (See Annex 3).

Narrow transcriptions, however, will have their limitations for practical purposes: their detailed nature will, no doubt, make fluent readings impossible; the extremely unfamiliar symbols and diacritics used in preparing them shall be problem-making for fluent mutual understandings, especially at the international level; and the obligation to stick to the details shall be extremely time-consuming at the time of their preparation. It was for these same reasons, as well as other similar considerations, that the committee agreed upon enforcing a normalization process through which the narrow transcriptions produced in the first stage introduced above would be normalized into, so called, counterpart broad transcriptions. These broad transcriptions were agreed upon to be prepared in as much a simple and straightforward manner as possible.

To this end, the 1987 transcription system, proposed by this author, was reinvestigated and approved. The system is, basically, Standard-Persian-based, by which it is meant that the narrow transcriptions introduced above have to undergo a, so called, "standardization process", through which they are rewritten in as much a reliable style as possible to fulfill the mentioned practical purposes while trying to present an honest "standardized" pronunciation for each. The system shall, however, remain open for minute further modifications as specific cases may require, of course, after being reinvestigated and approved by the Transcription Working Group. In the next step, the transliteration process shall be activated, through which the broad transcriptions will be converted into romanized translations of their spellings in Standard Persian writing system. These, no doubt, shall make various historical investigations

of them possible. For the time being, the procedures for the steps introduced above are available in the Transcription Working Group.

In what follows, the approved Broad Transcription Alphabet, as well as their counterparts from IPA, for narrow transcription, and their equivalent letters of Persian writing system shall be introduced together with examples. The related Procedure will follow in the end.

2-5- NCC Broad Transcription Alphabet and Procedure

2-5-1- Alphabet

The following tables present the symbols necessary for transcribing Persian consonants (tables 1 and 2) and Persian vowels (tables 3 and 4) broadly.

Table 1

NCC Consonant Symbols for Persian

(The symbols, which are equivalent to Standard Persian consonantal phonemes, have been presented between two slashes, according to the related standard practice)

No.	NCC Symbol / Standard Consonant	Consonant Persian	Technical Articulator Description
1	/p/		voiceless bilabial stop
2	/b/		voiced bilabial stop
3	/t/		voiceless dentil-alveolar stop
4	/d/		voiced dentil-alveolar stop
5	/k/		voiceless palatal stop
6	/g/		voiced palatal stop
7	/q/		voiced uvular stop
8	/ʔ/		voiceless glottal stop
9	/f/		voiceless labiodentals fricative
10	/v/		voiced labiodentals fricative
11	/s/		voiceless dentil-alveolar fricative
12	/z/		voiced dentil-alveolar fricative
13	/ʃ/		voiceless alveopalatal fricative
14	/ʒ/		voiced alveopalatal fricative
15	/x/		voiceless uvular fricative
16	/h/		voiceless glottal fricative
17	/c/		voiceless alveopalatal affricate
18	/j/		voiced alveopalatal affricate
19	/r/		voiced alveolar thrill
20	/m/		voiced bilabial nasal
21	/n/		voiced dentil-alveolar nasal
22	/l/		voiced alveolar lateral approximant
23	/y/		voiced palatal approximant

Table 2

NCC Consonant Symbols for Persian

(Other columns contain IPA equivalents, equivalent letters/symbols in Persian writing system, and examples in all three. IPA transcriptions have been enclosed between square brackets and NCC equivalents between two slashes)

No	NCC Consonant Symbol	IPA Equivalent	Equivalent Letter(s)/Symbols in Persian Writing System	Examples			
				Narrow IPA Transcription	Broad NCC Transcription	Persian Spelling	Meaning
1	/p/	ʔ ʔ ʔ	ʔ	ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ	/pamir/	ʔ ʔ ʔ ʔ	Proper Name
2	/b/	ʔ ʔ ʔ	ʔ	ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ	/babol/	ʔ ʔ ʔ ʔ	Proper Name
3	/t/	ʔ ʔ ʔ	ʔ ʔ	+ ʔ	/takab/ /tabarestan/	ʔ ʔ ʔ ʔ ʔ ʔ ʔ	Proper Name
4	/d/	ʔ ʔ ʔ	ʔ	ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ	/darab/	ʔ ʔ ʔ ʔ	Proper Name
5	/k/	ʔ ʔ ʔ / ʔ ʔ ʔ	ʔ	ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ	/kavir/	ʔ ʔ ʔ ʔ	"desert"
6	/g/	ʔ ʔ / ʔ ʔ ʔ	—	ʔ i ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ	/gajere/	ʔ ʔ ʔ ʔ	Proper Name
7	/q/	ʔ ʔ ʔ	ʔ ʔ	ʔ ʔ ʔ ʔ	/taleqan/ /damqan/	ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ	Proper Name
8	/ʔ/	ʔ ʔ	ʔ ʔ ʔ ʔ ʔ ʔ	ʔ ʔ ʔ ʔ + ʔ ʔ ʔ ʔ ʔ ʔ ʔ	/ʔliyabad/	ʔ ʔ ʔ ʔ	Proper Name
9	/f/	ʔ ʔ ʔ	ʔ	ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ ʔ	/famur/	ʔ ʔ ʔ ʔ	Proper
10	/v/	ʔ ʔ ʔ	ʔ	ʔ ʔ ʔ ʔ ʔ ʔ ʔ	/val/	ʔ ʔ	"whale"
11	/s/	ʔ ʔ ʔ	ʔ ʔ a ʔ	ʔ	/sarab/	ʔ ʔ ʔ ʔ	Proper

				+	/sanabad/	Om??	Name
				+	/sabra/	???	
12	/z/	???	? ?? ???	?	/zab/	? ??	Proper Name
					/zerrat/	? ??	"corn"
					/zarqam/	?????	Proper Name
13	/š/	???	e	?	/šemšak/	????	Proper Name
14	/ž/	???	?		/žir/	????	"alarm "
15	/x/	???	?	? ? ? ? ? ? ? ? ? ? ?	/xormiz/	??????	Proper Name
16	/h/	???	? ??	+	/harat/	? ???	Proper Name
				+	/hajibabad/	Om?? ? S	
17	/c/	????	,		/calus/	a ???	Proper Name
18	/j/	????	?	?	/jorjan/	? ???	Proper Name
19	/r/	???	?	+	/rudašt/	? ? Ø?	Proper Name
20	/m/	???	?		/marv/	???	Proper Name
21	/n/	???	?		/nur/	???	Proper Name
22	/l/	???	?		/lar/	??	Proper Name
23	/y/	???	?	?	/yervan/	? ????	Proper Name

Table 3

NCC Vowel Symbols for Persian

(The symbols, which are equivalent to Standard Persian vocalic phonemes, have been presented between two slashes, according to the related standard practice)

No.	NCC Consonant Symbol /	Consonant Persian	Technical Articulatory Description
	Standard Consonant		
1	/i/		high front unrounded vowel
2	/e/		mid front unrounded vowel
3	/a/		low front unrounded vowel
4	/u/		high back rounded vowel
5	/o/		mid back rounded vowel
6	/ɑ/		low back unrounded vowel
7	/ey/		mid front closing diphthong
8	/ou/		mid back closing diphthong

Table 4

NCC Vowel Symbols for Persian

(Other columns contain IPA equivalents, equivalent letters/symbols in Persian writing system, and examples in all three. IPA transcriptions have been enclosed between square brackets and NCC equivalents between two slashes)

No.	NCC Consonant Symbol	IPA Equivalent	Equivalent Letter(s)/Symbol(s) in Persian Writing System	Examples			
				Narrow IPA Transcription	Broad NCC Transcription	Persian Spelling	Meaning
1	/i/	???	???	?????? ??????	/xormiz/	??????	Proper Name
2	/e/	???	???	? ???? ?????	/gajere/	????	Proper Name
3	/a/	???	??	?????? ??	/marv/	??	Proper Name
4	/u/	???	????	?????? ?	/nur/	??	Proper Name

5	/o/	???	? ???	?? ? ? ? ? ? ? ? ?	/babol/	???	Proper Name
6	/a/	???	??	?? ? ? ? ? ? ? ? ?	/darab/	? ??O	Proper Name
7	/ey/	?? ??	?-	?? ? ? ? ? ? ?	/meyl/	??	"inclinati on"
8	/ou/	?? ??	??	?? ? ? ? ? ? ? ? ? ? ?	/gouhar/	????	"jewel"

2-5-2- Procedure

2-5-2-1- Objective: The present Procedure has been prepared to materialize the goal of converting the narrow transcriptions of toponymic items kept in NCC's Narrow Transcriptions Database (cf. 2-1 & 2-4) into their broad NCC equivalents (cf. 2-1, 2-4, & 2-5-1), to complete NCC's Broad Transcriptions Database (cf. 2-1), which is going to be used in various applications, and in the process of preparing transliterations out of them.

2-5-2-2- Definitions:

2-5-2-2-1- Broad Transcription

In this Procedure, the term *broad transcription* has been registered for any of the strings of symbols, with or without diacritics, selected from the set of NCC broad transcription symbols, to register linguistic material (specifically, toponymic items) for various applications (including, registration of geographical names on maps), having investigated, beforehand, the equivalent narrow transcriptions kept within NCC's Narrow Transcriptions Database.

2-5-2-2-2- Broad Transcriber

The broad transcriber has been deemed, here, as an individual who is familiar with the International Phonetic Alphabet (IPA) in detail, and is competent enough in converting IPA strings to their NCC broad counterparts.

Note: Broad Transcribers have to pass special oral and written tests on IPA alphabet and methodology as well as the methodology to convert it into NCC.

2-5-2-3- Domain

In the present procedure, the raw material for broad transcriber shall be the narrow transcriptions available at NCC's Broad Transcriptions Database. The broad transcriber shall:

a) Reinvestigate each narrow transcription carefully and make the necessary proposals for possible corrections;

b) Convert the approved narrow transcriptions into their NCC counterparts, in course of a practical-library work.

The result, shall, no doubt, be a rewriting of the narrow transcriptions kept in NCC's Narrow Transcriptions Database into Standard Persian phonemic system.

2-5-2-4- Modifications

In the course of preparing NCC broad transcriptions, the Broad Transcriber shall pay special attention to the following items to make the result as realistic and practical as possible:

a) the accuracy of the converted strings, in comparison with their narrow equivalents,

b) the need for optimizing modifications, including, among others, the methodology and the inventory of NCC broad symbols.

Note: The outcomes of the abovementioned shall be prepared in the form of written proposals by the supervisor to the team of broad transcribers to the Transcription Working Group for final approval. Any modification in any part of the overall program shall be included, only, after prior reinvestigation and approval by the Transcription Working Group and final enactment by the Specialist Committee for Romanization and Standardization of Geographical Names.

2-5-2-5- Approval and Finalization

In order for NCC broad transcriptions prepared based on the method above to be included in NCC's Broad Transcriptions Database, they have to pass the following tests:

a) Regular reinvestigations by other broad transcribers selected and authorized by the supervisor to the team of broad transcribers (each broad transcription shall be accompanied by the approval by, at least, two broad transcribers;

b) Regular random reinvestigations by the supervisor to the team of broad transcribers;

c) Regular random reinvestigations by non-member specialists selected and authorized by the Head of the Transcription Working Group.

Each broad transcription shall enter NCC's Broad Transcriptions Database, only with the final approval by the Head of the Transcription Working Group.

2-5-2-6- Utilization

The toponymic items kept in NCC's Broad Transcriptions Database shall be available for any practical purpose at both national and international levels. Any utilization of the database shall be materialized, however, based on a prior permit in writing by the head of Transcription Working Group, approved by the Committee for Romanization and Standardization of Geographical Names.