Twenty-Second Session of United Nations Group of Experts on Geographical Names

New York, 20-29 April 2004

Item 11 of the provisional agenda

Report Submitted by the Islamic Republic of Iran

National Cartography Center

Iranian Committee for Standardization of Geographical Names (ICSGN)

P.O.BOX 13185-1684,Azadi Sq., Tehran, Iran

Tel: +98(21)6000031/37, Fax: +98(21)6001972

http://www.geonames.ncc.org.ir

Abstract:

Adoption of a standard method in consistent registration of geographical names is among the major duties of the organizations involved in geographical studies. This approach, prior to anything else, makes possible transparent reference to various regions and territories by name, both nationally and internationally.

This, however, should not be regarded as the sole goal for toponymy. Geographical names are among the determining parameters in the background of every culture and history: a linguistically and historically transparent and justified reference can, quite possibly, throw light on an important historical and/or cultural aspect. On the contrary, a careless registration will, no doubt, hide a related fact forever.

In registering correct geographical names, special attention must be paid to the linguistic ground on which the name in question were made. Quite usually, there may be no vigorous cultural-historical evidence in support of a special name, pronunciation, or spelling, in rejection of a rival, Obviously, if the decision is made solely according to such shaky evidence. Its reliability will fall to the level of subjective guesses. In such situations a close linguistic study of the structure of the name, both form a synchronic and a diachronic aspect, will solve the majority, if not all, of the superficially serous problems.

Standardization of toponymy methods is another independent issue, emphasizing the determining role of linguistics in this relation. The linguist, here, is taken to the stance of a planner who, joining his technical knowledge of linguistics with an acquaintance with local feasibilities, suggests the easiest and, at the same time, most efficient scientific method for this purpose.

Standard registration of geographical names, in the first place, is in need of adopting scientifically justifiable methods for preparing transcriptions and transliteration of their regional pronunciations.

The provision of a *Transcription Working Group* with in the overall body of, Iranian Committee for standardization of Geographical Names (ICSGN) affiliated to NCC, has been to fulfill this urgent need.

1- In what follows, the term transcription, as a name for the working group proposed, shall be interpreted broadly to cover both broad and narrow transcription, on the one hand, and transliteration, on the other, as the duties to be performed by the group. (of, the section titled Delineation).

Transcription Working Group: Definitions

i- *Transcription Working Group:* The Transcription Working Group is the name given to a subset of the Iranian Committee for standardization of Geographical Names (ICSGN), affiliated to NCC, responsible for (a) providing instruction notebooks to be used for geographical name transcription and transliteration, as well as (b) guaranteeing any proper utilization thereof.

ii- *Geographical Name:* The regional or official name used in reference to a geographical place.

iii- *Transcription:* The process of rendering the distinctive phonetic characteristics distinguishable by car, in regional pronunciation(s) of geographical names , by way of a standardized system (phonetic alphabet). Transcriptions can be broader or narrower, a according to the purposes of their utilizations.

iv- *Transliteration:* The process of rendering the ordinary spellings of geographical names in a standardized symbol system.

v- *Standardization:* The process of making standard and utilizable a written from to refer to a specific geographical name.

Transcription Working Group: Duties

The Transcription Working Group shall be responsible for the following:
Review of old transcription methods, national or international, to determine the advantages
and shortcomings thereof,

- Appropriate linguistic determination of the expectations on the part of the ICSGN , affiliated to NCC,
- Appropriate determination of the level of capacity on the part of the third parties, to be able to prepare the most efficient transcription method.
- Appropriate determination of the most efficient level for narrow transcription of geographical names, having considered the scientific- practical aspects, beforehand,
- Primary determination of the set of phonemes for language and dialect types of Iran (The set being considered, necessarily, as incomplete, open for further modifications resulting from new discoveries),
- Appropriate determination of the most efficient transliteration method for geographical name registration, having considered the scientific-practical aspects, beforehand.

- Appropriate determination of level of proficiency and knowledge on the part of applicants for transcription and transliteration, as well as their subsequent assignment for the job,
- Enforcing continuous supervision on the process of geographical name registration.

Transcription Working Group: Achievements

In the end, the Transcription Working Group shall come up with three handbooks of instructions, to be utilized in practice, for the following purposes:

1. **Handbook of Instructions for Narrow Transcription:** This will be prepared based on the standard approach put forward by the International Phonetic Association(IPA); any result thereof to be kept in the related data-base and used in detailed investigations.

2. **Handbook of Instructions for Broad(or Phonemic)Transcription:** This shall be prepared according to a Transcription system to be approved in this relation by the same working group, the results of which to be kept in the related data- base and used in related public research.

3. **Handbook of instructions for Transliteration:** This shall be Latin-based and particulary constructed for utilization in relation to Persian orthography, and the results shall be kept in the related data-base and used in related research.

Now this working group prepared their documents and submitted three handbooks of instruction to ICSGN and ICSGN submite the Transcription Proceduee for Iranian Toponymic geographical Names to UNGEGEN .

IN THE NAME OF GOD

ISLAMIC REPUBLIC OF IRAN
NATIONAL CARTOGRAPHY CENTER (NCC)

IRANIAN COMMITTEE FOR FOR STANDARDIZATION OF
GEOGRAPHICAL NAMES(ICSGN)

TRANSCRIPTION WORKING GROUP FOR GEOGRAPHICAL
NAMES

TRANSCRIPTION PROCEDURE
FOR
IRANIAN GEOGRAPHICAL NAMES

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 Iranian Committee for 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 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. «ی» represents both [i] and [y]; «ه» stands for both [e] and [h]; and «و» 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 «ز», «ذ», «ض» and «ظ», all stand in correspondence with [z]; and the three letters «س», «ص» and «ث» 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 «گرد» can be read as either [gærd] "powder", [gerd] "round" or [gord] "hero").

6

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

- Most of the Persian characters change their shape in different collocations (e.g. the first letter in «غار» "cave", the second letter «تغییر» "change", the final letter in «جیغ» "scream", and the final letter in «باغ» "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» standing for [b], «پ» standing for [p], «ت» standing for [t], and «ث» standing for [s]; or «ک» standing for [k] and «گ» 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.

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 in 1988, 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 the way to utilize the 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 ethnic communities, each with its own different language, live within the borders. In addition to these, there are numerous other communities each with a different dialect of Persian. 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. 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 1988 transcription system, proposed, 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 Consonant Symbol / Standard Persian Consonant	Technical Articulatory Description
1	/p/	voiceless bilabial stop
2	/b/	voiced bilabial stop
3	/t/	voiceless denti-alveolar stop
4	/d/	voiced denti-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 labiodental fricative
10	/v/	voiced labiodental fricative
11	/s/	voiceless denti-alveolar fricative
12	/z/	voiced denti-alveolar fricative
13	/š/	voiceless alveopalatal fricative
14	/ž/	voiced alveopalatal fricative
15	/x/	voiceless uvular fricative
16	/h/	voiceless glottal fricative
17	/č/	voiceless alveopalatal affricate
18	/j/	voiced alveopalatal affricate
19	/r/	voiced alveolar thrill
20	/m/	voiced bilabial nasal
21	/n/	voiced denti-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)/Symbol(s) in Persian Writing System	Examples			
				Narrow IPA Transcription	Broad NCC Transcription	Persian Spelling	Meaning
1	/p/	[p]	پ	[pʰɑːmíɾ]	/pāmir/	پامیر	Proper Name
2	/b/	[b]	ب	[bɑːbʷoˈl]	/bābol/	بابٔل	Proper Name
3	/t/	[t]	ت، ط	[tʰæbær+estʰɑ́n̩] [tʰækʰ+ɑ́b̩]	/takāb/ /tabarestān/	تَکاب طَبَرِستان	Proper Name Proper Name

4	/d/	[d]	د	[dɑːrɑˊb̥]	/**d**ārāb/	داراب	Proper Name
5	/k/	[c] / [k]	ک	[cʰævíˊr̥]	/**k**avir/	کَویر	"desert"
6	/g/	[ɟ] / [g]	گ	[ɟædʒeɾeˊ]	/**g**ajere/	گَجِره	Proper Name
7	/q/	[ɢ]	ق، غ	[tʰɑːleɢɑ̃ˊn̥] [dãmɢɑ̃ˊn̥]	/tāle**q**ān/ /dām**q**ān/	طالقان دامغان	Proper Name
8	/ʔ/	[ʔ]	ع، ء، ئ، ؤ، أ	[ʔæli+j+ɑːbɑˊd̥]	/**ʔ**aliyābād/	عَلی‌آباد	Proper Name
9	/f/	[f]	ف	[fɑːmʷuˊr̥]	/**f**āmur/	فامور	Proper Name
10	/v/	[v]	و	[vɑl̥]	/**v**āl/	وال	"whale"
11	/s/	[s]	ث، س، ص	[sæn+ɑːbɑˊd̥] [sæɾ+ɑˊb̥] [sæbɾɑˊ]	/**s**arāb/ /**s**anābād/ /**s**abrā/	سَراب صَناباد صَبرا	Proper Name
12	/z/	[z]	ز، ذ، ض، ظ	[zɑb̥] [zʷorˋræˊtʰ] [zæɾɢɑ̃ˊm̥]	/**z**āb/ /**z**orrat/ /**z**arqām/	زاب ذُرَّت ضَرغام	Proper Name "corn" Proper Name
13	/š/	[ʃ]	ش	[ʃẽmʃ+æˊcʰ]	/**š**emšak/	شِمشَک	Proper Name
14	/ž/	[ʒ]	ژ	[ʔɑːʒíˊr̥]	/ʔā**ž**ir/	آژیر	"alarm"
15	/x/	[χ]	خ	[χʷormíˊz̥]	/**x**ormiz/	خورمیز	Proper Name
16	/h/	[h]	ﻩ، ح	[hæɾɑˊtʰ] [hɑːdʒi+j+ɑːbɑˊd̥]	/**h**arāt/ /**h**ājiyābād/	هَرات حاجی‌آباد	Proper Name
17	/č/	[tʃ]	چ	[tʃɑːlʷuˊs]	/**č**ālus/	چالوس	Proper Name
18	/j/	[dʒ]	ج	[dʒʷordʒɑ̃ˊn̥]	/**j**orjān/	جُرجان	Proper Name
19	/r/	[r]	ر	[rʷuː+dæˊʃtʰ]	/**r**udašt/	رودَشت	Proper Name
20	/m/	[m]	م	[mæɾ̥ɣ̥]	/**m**arv/	مَرو	Proper Name
21	/n/	[n]	ن	[nʷur̥]	/**n**ur/	نور	Proper Name
22	/l/	[l]	ل	[lɑr̥]	/**l**ār/	لار	Proper Name
23	/y/	[j]	ی	[jervɑ̃ˊn̥]	/**y**ervān/	یِروان	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 / Standard Persian Consonant	Technical Articulatory Description
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	**/ā/**	low back unrounded vowel
7	**/ey/**	mid front closing diphtong
8	**/ou/**	mid back closing diphtong

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/	[i]	ای، ی	[χʷormíʐ]	/xormiz/	خورمیز	Proper Name
2	/e/	[e]	ا، ◌ِ، ه	[ɟædʒeré]	/gaĵere/	گَجرِه	Proper Name
3	/a/	[æ]	اَ، ◌َ	[mæɾʁ̥]	/marv/	مَرو	Proper Name
4	/u/	[u]	او، و	[nʷuɾ̥]	/nur/	نور	Proper Name
5	/o/	[o]	اُ، ◌ُ، و	[ɑːbʷól]	/bābol/	بائل	Proper Name
6	/ā/	[ɑ]	آ، ا	[dɑːɾáb̥]	/dārāb/	داراب	Proper Name
7	/ey/	[ej]	◌ِی	[mej̃l̥]	/meyl/	مِیل	"inclination"
8	/ou/	[ow]	◌ُو	[gʷowhæɾ̥]	/gouhar/	گوهر	"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
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,
to make the result as realistic and practical as possible.

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 ICSGN.

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 competent individuals selected and authorized by the supervisor to the team of broad transcribers;

b) Random reinvestigations by the supervisor to the team of broad transcribers.

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