

23 June 2017

Original: English

**Eleventh United Nations Conference on the
Standardization of Geographical Names**

New York, 8-17 August 2017

Item 14 (a) of the provisional agenda*

Writing systems and pronunciation: Romanization

Rules of romanization of geographical names applied in Poland

Submitted by Poland **

* E/CONF.105/1

** Prepared by Maciej Zych, Commission on Standardization of Geographical Names Outside the Republic of Poland

Rules of romanization of geographical names applied in Poland

The Commission on Standardization of Geographical Names Outside the Republic of Poland is the authority responsible for names of geographical objects located outside Poland. One of Commission's tasks is to determine what names for these objects should be used in the Polish language. They may be both: Polish exonyms or original names – endonyms. In the case of endonyms originally written in the Latin alphabet, the Commission recommends to use original letters with diacritic marks as well as any non-standard letters (i.e. æ, ð, ə). In the case of endonyms originally written in a non-Roman writing system, it is recommended to render them in a Romanized form.

Due to above mentioned scopes of activity the Commission on Standardization of Geographical Names Outside the Republic of Poland establishes also romanization rules governing names of geographical objects originally expressed by means of non-Roman writing systems. These rules are determined primarily for names in the languages used in Commission's publications, including the *Official list of Polish geographical names of the world*, where, apart from the Polish exonyms, endonyms are given, including endonyms rendered in a Romanized form.

The Commission on Standardization of Geographical Names Outside the Republic of Poland establishes mainly two systems of romanization for each language: a simplified Polish phonetic transcription (which is used inter alia in: school atlases for lower levels of education, popular publications, the press) and a transliteration or an international romanization (which is recommended, among others, in geographic atlases for older classes, or in scientific publications). A Polish phonetic transcription is recommended for the majority of official countrywide languages using non-Roman scripts – with exceptions for languages of the East Asia and the Southeast Asia. For Chinese, Japanese, Korean, Thai, Lao, Burmese, and Khmer a transliteration/international romanization is only recommended. However, for the majority of regional/local official languages Polish phonetic transcription rules have not been developed yet (e.g. for languages of the republics of Russia, or regional official languages in China). A Polish phonetic transcription have not also been established for the Berber language (official in Morocco and Algeria).

Generally, the Commission on Standardization of Geographical Names Outside the Republic of Poland has adopted romanization rules for 37 languages (36 national languages and one local language) and applies, without formal adoption, rules for 40 another languages (2 national languages and 38 regional/local languages)¹.

In the case of a transliteration/international romanization, the Commission on Standardization of Geographical Names Outside the Republic of Poland always considers adopting the rules recommended by the United Nations (as a resolution of the United Nations Conference on the Standardization of Geographical Names – UNCSGN), as well as rules not officially adopted by the UNCSGN but applied by the UNGEGN (e.g. in the list of countries) and listed on the Working Group on Romanization Systems' web-site. However, the Commission does not always accept these UNGEGN's romanization rules. Among 30 rules

¹ Rules of romanization adopted by the Commission are available on Commission's web site: <http://ksng.gugik.gov.pl/latynizacja.php>.

adopted by the UNCSGN (19 for statewide languages and 11 for regional languages) the Commission on Standardization of Geographical Names Outside the Republic of Poland recommends rules for 12 languages. In the case of other 15 languages which are in the agenda of the Working Group on Romanization Systems, although not formally adopted by the UNCSGN (all for statewide languages), the Commission recommends rules for 6 of them.

For the following 10 national languages, the Commission recommends the same rules as adopted by the UNCSGN:

- **Arabic** – the system recommended by the UN in 1972,
- **Belarusian** – the Belarusian national system of 2007, as recommended by the UN in 2012,
- **Bulgarian** – the Bulgarian national system of 2006, as recommended by the UN in 2012,
- **Chinese** – the Pinyin phonetic system (Hànyǔ Pīnyīn), as recommended by the UN in 1977,
- **Khmer** – the BGN/PCGN 1972 system, as recommended by the UN in 1972,
- **Macedonian** – the official Yugoslav system, as recommended by the UN in 1977,
- **Russian** – the GOST 1983 system, as recommended by the UN in 1977,
- **Serbian** – Serbian Latin script (srpska latinica), as recommended by the UN in 1977,
- **Thai** (Thailand) – the modified system of the Royal Institute of Thailand, as recommended by the UN in 2002,
- **Ukrainian** – the Ukrainian national system of 2010, as recommended by the UN in 2012.

Also for two regional languages, the Commission on Standardization of Geographical Names Outside the Republic of Poland applies the rules recommended by the UNCSGN:

- **Tibetan** (China) – the Pinyin phonetic system for the Tibetan language in China (Zàngwén Pīnyīn), as recommended by the UN in 1977 (and the Wylie system of 1959),
- **Uighur** (China) – the Pinyin phonetic system for the Uighur language in China, as recommended by the UN in 1977.

On the other hand, for 9 national languages and 9 regional languages the Commission recommends different rules than those adopted by the UNCSGN.

Official languages of states:

- **Amharic** – the BGN/PCGN 1967 system – this system is widely used, whereas the Imperial Ethiopian Mapping and Geography Institute system, recommended by UNCSGN in 1972, is not currently used even in Ethiopia itself (as reported the convenor of the Working Group of Romanization Systems during the WG meeting in Prague in April 2017).
- **Bengali** – the ISO 15919:2001 system (the remark as for Hindi below),
- **Greek** – the ISO 843:1997 system, which is an international standard, and which is traditionally used in Poland for transliteration of Greek names; the ELLOT 743 system, recommended by UNCSGN in 1987, however, has not been implemented in Poland,
- **Hebrew** (Israel) – the official Israeli system of 1957, as recommended by the UN in the years 1977-2007; the new system of 2006 of the Academy of the Hebrew Language was adopted by the UNCSGN in 2007, the Commission on Standardization of Geographical Names Outside the Republic of Poland has postponed decision on changing the recommended system until actual implementation in Israel this new

system (according to information provided by the convenor of the Working Group of Romanization Systems during the WG meeting in Prague in April 2017, there is no clear evidence of implementation this new system in Israel),

- **Hindi** – the ISO 15919:2001 system, which is an international standard, and which has been used in Poland by Indologists for a long time; the UNCSGN system adopted in 1972 and based on a report prepared by D.N. Sharma, has been neither implemented in India nor widely used in cartographic publications – its adoption by the Commission on Standardization of Geographical Names Outside the Republic of Poland and a recommendation for use in publications published in Poland would have no logical grounds,
- **Nepali** – the ISO 15919:2001 system (the remark as for Hindi above),
- **Persian** – the system of the Commission on Standardization of Geographical Names Outside the Republic of Poland (minor modification of the UN romanization system 1967); the new system of romanization of the Persian language was approved by the UNCSGN in 2012, but the Commission has postponed decision on changing the recommended system until its implementation in Iran and a popularize in international cartographic publications,
- **Tamil** – the ISO 15919:2001 system (the remark as for Hindi above),
- **Urdu** – the system of the Commission on Standardization of Geographical Names Outside the Republic of Poland (based on J.T. Platts² and G.A. Zograf³ transliteration schemes); the UNCSGN system adopted in 1972 and based on a report prepared by D.N. Sharma, has been neither implemented in Pakistan or India, nor widely used in cartographic publications – its adoption by the Commission on Standardization of Geographical Names Outside the Republic of Poland and a recommendation for use in publications published in Poland would have no logical grounds,

Regional languages:

- **Assamese** (India) – the ISO 15919:2001 system (the remark as for Hindi above),
- **Gujarati** (India) – the ISO 15919:2001 system (the remark as for Hindi above),
- **Kannada** (India) – the ISO 15919:2001 system (the remark as for Hindi above),
- **Malayalam** (India) – the ISO 15919:2001 system (the remark as for Hindi above),
- **Marathi** (India) – the ISO 15919:2001 system (the remark as for Hindi above),
- **Mongolian** (China) – the Mongolian National Center for Standardization and Metrology system of 2003; UNCSGN adopted in 1977 the Chinese Pinyin phonetic system for the Mongolian language; the Commission on Standardization of Geographical Names Outside the Republic of Poland has recognized, however, that there is no reason for using a different systems for the Mongolian language in China and for the Mongolian language in Mongolia,
- **Oriya** (India) – the ISO 15919:2001 system (the remark as for Hindi above),
- **Punjabi** (India) – the ISO 15919:2001 system (the remark as for Hindi above),
- **Telugu** (India) – the ISO 15919:2001 system (the remark as for Hindi above).

For the following 6 national languages, the Commission on Standardization of Geographical Names Outside the Republic of Poland recommends the same rules as those applied by the UNGEGN (although not formally adopted by the UNCSGN):

- **Armenian** – the BGN/PCGN 1981,
- **Burmese** – the BGN/PCGN 1970 system,
- **Georgian** – the Georgian national system of 2002 (the BGN/ PCGN 2009 system),

² J.T. Platts, *A Dictionary of Urdu, Classical Hindi, and English*, Oxford 1984 (1st ed., 1884).

³ Г.А. Зорграф, *Языки Южной Азии*, Москва 1990 [G.A. Zograf, *Jazyki Južnoj Azii*, Moskva 1990].

- **Korean** – the McCune-Reischauer 1939 system (and the South Korean system of 2000 – revised romanization of Ministry of Education),
- **Lao** – the BGN/PCGN 1966 system,
- **Maldivian** – the Maldivian Government 1987 System.

For the following 9 languages, the Commission uses different rules from the ones used by the UNGEGN (although not formally adopted by the UNCSGN):

- **Dzongkha** – the Roman Dzongkha system of 1994; UNGEGN uses the newer Roman Dzongkha 1997 system; at present the Commission on Standardization of Geographical Names Outside the Republic of Poland works on verifying the system used for a romanization of the Dzongkha language,
- **Japanese** – the modified Hepburn system with diacritical symbols; no information which system is used by the UNGEGN,
- **Kazakh** – the ISO 9:1995 system; the UNGEGN uses the BGN/PCGN 1979 System,
- **Kirghiz** – the ISO 9:1995 system; the UNGEGN uses the BGN/PCGN 1979 System,
- **Mongolian** – the Mongolian National Center for Standardization and Metrology system of 2003; the UNGEGN uses the BGN/PCGN 1964 System,
- **Pashto** – the system of the Commission on Standardization of Geographical Names Outside the Republic of Poland (the modification of the BGN/PCGN 1968 system); the UNGEGN uses the BGN/PCGN 1968 System,
- **Sinhalese** – the ISO 15919:2001 system; the UNGEGN uses a system based on a report prepared by D. N. Sharma,
- **Tajik** – the system of the Commission on Standardization of Geographical Names Outside the Republic of Poland (the modification of the ISO 9:1995 system); the UNGEGN uses the BGN/PCGN 1994 System,
- **Tigrinya** – the BGN/PCGN 1994 system; the UNGEGN uses the BGN/PCGN 2007 System.

In addition, the Commission on Standardization of Geographical Names Outside the Republic of Poland recommends for using in Poland the following romanization rules for languages not listed on a web-site of the Working Group on Romanization Systems.

For 4 official languages of states:

- **Berber** (Morocco, Algeria) – the Moroccan Royal Institute for Amazigh Culture System,
- **Dari** – the system of the Commission on Standardization of Geographical Names Outside the Republic of Poland (the modification of the UN romanization system 1967 for Persian),
- **Fiji Hindi** – the ISO 15919:2001 system for Hindi,
- **Kurdish** – the romanization system based on the Kurmancî writing system.

For 29 regional/local languages:

- **Abaza** (Russia) – the ISO 9:1995 system,
- **Abkhaz** (Georgia) – the ISO 9:1995 system,
- **Adyghe** (Russia) – the ISO 9:1995 system,
- **Altai** (Russia) – the ISO 9:1995 system,
- **Bashkir** (Russia) – the ISO 9:1995 system,
- **Buryat** (Russia) – the ISO 9:1995 system,
- **Cantonese dialect of Chinese** (China) – the Jyutping phonetic system of 1993 without tone marks, as recommended by Linguistic Society of Hong Kong,
- **Chechen** (Russia) – the ISO 9:1995 system,

- **Chuvash** (Russia) – the ISO 9:1995 system,
- **Hill Mari** (Russia) – the ISO 9:1995 system,
- **Ingush** (Russia) – the ISO 9:1995 system,
- **Kabardian (Cherkess)** (Russia) – the ISO 9:1995 system,
- **Kalmyk** (Russia) – the ISO 9:1995 system,
- **Karachay-Balkar** (Russia) – the ISO 9:1995 system,
- **Khakas** (Russia) – the ISO 9:1995 system,
- **Komi** (Russia) – the ISO 9:1995 system,
- **Konkani** (India) – the ISO 15919:2001 system,
- **Manipuri** (India) – the ISO 15919:2001 system,
- **Meadow Mari** (Russia) – the ISO 9:1995 system,
- **Mordvin Erzya** (Russia) – the ISO 9:1995 system,
- **Mordvin Moksha** (Russia) – the ISO 9:1995 system,
- **Ossetian** (Georgia, Russia) – the ISO 9:1995 system,
- **Rusyn (Lemko)** (Poland) – the system of the Commission on Standardization of Geographical Names Outside the Republic of Poland (based on the international Slavistics system for the transliteration of Cyrillic),
- **Rusyn (Pannonian Rusyn)** (Serbia) – the system of the Commission on Standardization of Geographical Names Outside the Republic of Poland (the modification of the international Slavistics system for the transliteration of Cyrillic),
- **Rusyn (Prešov Rusyn)** (Slovakia) – the system of the Commission on Standardization of Geographical Names Outside the Republic of Poland (the modification of the international Slavistics system for the transliteration of Cyrillic),
- **Tatar** (Russia) – the ISO 9:1995 system,
- **Tuvan** (Russia) – the ISO 9:1995 system,
- **Udmurt** (Russia) – the ISO 9:1995 system,
- **Yakut** (Russia) – the ISO 9:1995 system.

As it was presented above, in the most cases the Commission on Standardization of Geographical Names Outside the Republic of Poland uses systems other than those recommended by the UNCSGN – only 12 systems recommended by the Commission are the ones adopted by the UNCSGN. There are several reasons for this. Firstly, some of the systems recommended by the UNCSGN are systems that are not widely used, and in some cases their acceptance by UNCSGN seems to be hasty (especially in the case of Indian languages). The UNGEGN should withdraw the recommendation for these unimplemented systems.

Secondly, in the agenda of the UNGEGN Working Group on Romanization Systems, there are lack of romanization systems for many languages. There are mainly regional or local official languages (e.g. official languages of republics in Russia), however, there are also no rules for several national languages:

- **Berber** – since 2011 it is the official language in Morocco and since 2016 also in Algeria, where it has formally received the same status as the Arabic language,
- **Dari** – one of the two nationwide official languages of Afghanistan, the variety of the Persian language (sometimes not recognized as a separate language, hence some sources indicate that the official language of Afghanistan is Persian; in 2012 the UNCSGN adopted the new romanization systems for Persian, which explicitly referred the Persian language to Iran solely),
- **Fiji Hindi** – one of the three official languages of Fiji,

- **Kurdish** – one of the two nationwide official languages of Iraq (since 2005).

The lack of uniform, up-to-date and credible romanization rules adopted on the UNGEGN forum for a large number of languages means that different romanization rules are applied to these languages in different countries, and these countries choose the rules most favorable to them from their own point of view. This situation can be observed also in Poland. It would therefore be important to increase on the UNGEGN forum the effort into establishing romanization rules for more languages. And such rules should have a real the potential to be more widespread in both the proposing country and in international cartographic publications. At the same time, the UNGEGN should withdraw unimplemented romanization systems. This would facilitate the Commission on Standardization of Geographical Names Outside the Republic of Poland to recommend rules adopted by the UNCSGN for use in Poland.