



UNITED NATIONS
ECONOMIC
AND
SOCIAL COUNCIL



Distr.
LIMITED

E/CONF.53/L.31
15 August 1967

ORIGINAL: ENGLISH

UNITED NATIONS CONFERENCE ON
THE STANDARDIZATION OF
GEOGRAPHICAL NAMES
Geneva, 4-22 September 1967
Item 11 of the provisional agenda

WRITING SYSTEMS

- Some general principles of transliteration and transcription -

Submitted by the Government of
the United States of America*

* Limited number of copies issued only to participants.

One and only one Roman-letter symbol or combination of symbols should be used for a given sound or letter in a transcription or transliteration system and not more than one sound or letter should be represented by a given Roman-letter symbol. For transcription systems, this requires an accurate linguistic analysis of the sounds and ranges of distinctive sounds (phonemes), tones, accent patterns and other significant phenomena of a language before a transcription system can be drawn up. For transliteration systems, there is required an arrangement of the letters of the alphabet concerned over against Roman-letter symbols in a manner such that the Roman letters would be those that the user of the transliterated name would naturally associate with those sounds. For the use of English-speaking countries, at least, only after the Roman-letter stock of reasonably appropriate letters is exhausted should one resort to the use of diacritical marks or modified letters.

In general, transcription should not be combined with transliteration. Only confusion would result from the mixture of graphic and phonemic substitutive symbolization, since transcription is the process of substituting (in the Roman-alphabet areas) Roman-letter symbols for the sounds of a given language, whereas transliteration is not concerned in the strictest sense with sounds, but with Roman-letter equivalents of non-Roman letters. Combination of transcription with transliteration and *vice versa* leads to the temptation of levelling out sound distinctions which do not exist in the language of the transcriber. A notable example would be, in a transliteration system for Arabic, transliterating the Arabic velarized (usually called emphatic) consonant “*hā*” in the same way as the nonvelarized consonant “*hā*”, “*ṣād*” as “*sin*”, “*ḷā*” as “*tā*”, “*ḡād*” as “*dāl*” and “*ẓā*” as “*zāy*”. Nor should the long-short vowel distinctions of languages such as Arabic, Amharic or Mongolian be neglected in a transcription or transliteration system because they do not occur in the transliterator’s language.

Automatic or nondistinctive sound features such as the fixed initial syllable accents in Mongolian should be disregarded in transcription systems. They need not be marked in the transcriptions themselves because their occurrence can be defined for the users of toponyms in general explanations of pronunciation. Likewise, graphic symbols of non-Roman alphabets can be disregarded in transliteration where they are automatically replaceable in transliteration, as initial “*hamzah*” in Arabic, or the “smooth breathing” in Greek. Morphophonemic symbols (symbols which represent now one, now another sound, depending on the phonetic environment), can be resolved

into Roman-letter symbols for the sounds they represent. Thus the “*lām*” of the Arabic definite article can be assimilated to the rules of Arabic, or the “*e*” in Russian Cyrillic can be transliterated as “*ye*” and not “*e*” initially, after vowels, and after the so-called hard and soft consonants.

For transcription and transliteration systems to be used in writing the place names of an area in Roman letters, groups of toponyms should be dealt with area by area. The sound system for the main or nationally official language or dialect of an area should be the basis of transcription and divergent dialects should be disregarded. Otherwise the management of a toponymic programme is likely to become very complex, since a number of sound systems must be analysed instead of only one; moreover, the problem of dialect boundaries, which is often exceedingly difficult, must be resolved. An example is the situation in regard to Chinese names: a satisfactory treatment is possible in terms of the pronunciation of Mandarin Chinese, which is understood throughout most of China, whereas a policy of differentiating between Mandarin, Wu, Hakka and Cantonese, to say nothing of subdialects, would at present at least be almost impossible. Similarly, the transliteration of a non-Roman alphabet by different systems according to the pronunciation of local dialects presupposes the accurate knowledge of local pronunciation which is not strictly speaking, the concern of transliteration.

Since United States Board on Geographic Names transcription and transliteration systems are designed to strike a balance between scientific accuracy and intelligibility for the general public of the United States, diacritical marks and modified letters are kept to a minimum. Some linguistic features totally foreign to European languages, such as the tones of the Sino-Tibetan-Burmese languages, must be sacrificed when they are completely incomprehensible to the user in the United States, or when their presence or even their nature cannot be determined with any accuracy at the present state of our knowledge.

Many problems in the Roman-letter spelling of toponyms from non-Roman writing systems still remain unresolved. The number of areas for which sufficient geographical and linguistic source materials are as unavailable is still larger than one would desire. Such matters as the scarcity of Arabic and Persian toponyms in Arabic script complete with vowel points, and the great variation in the geographic and toponymic data for many areas throughout the world present perplexing problems. However, progress is being made and better results are continually being achieved through the application of toponymic principles to the problem of spelling the place names of the world in Roman letters.

SOME GENERAL PRINCIPLES OF TRANSLITERATION AND TRANSCRIPTION

Paper presented by the United States of America¹

Every country that uses geographical names on maps or charts or in books or documents of any sort is faced with two general categories of such names: names in its own language and names in languages other than its own.

A country can use without further ado names within its own boundaries or names within other countries using the same language. This is especially true if names are avail-

able in easily accessible gazetteers, lists or other publications that provide positive determination of the identity and writing form of the names themselves. A faithful copying of the names from such definitive sources will in each and every case reproduce the body of names without change of any sort.

A country is confronted with other problems when it renders names from languages other than its own, whether such names are written in a variation of the alphabet or writing system of its own language or in an entirely different alphabet or writing system.

¹ The original text of this paper, prepared by J. G. Mutziger, Office of Geography, Department of the Interior, appeared as document E/CONF.53/L.31.

As long as names foreign to a country are written in a variation of the writing system of its own language, total agreement between the donor country's written forms and the written forms used by the receiver country will be achieved by the receiver country if all diacritical marks and modified letters or characters of the donor language are faithfully reproduced.

For example, English language names need merely be copied by Czech and Polish users, whereas Czech and Polish names can be kept in their original form by English-language users only if such un-English Czech symbols as "č", "á", "ř" and "ů" and such un-English Polish symbols as "ą", "ł", "ś" and "ź" are faithfully copied. On the other hand, Czech and Polish users must in their turn reproduce such French symbols as "à", "â" and "ç" if their revisions of French names are to maintain identity with the original written forms.

As a further example, Iranian users of Arabic names need merely copy them as they are, since no letters occur in the writing of Arabic that do not occur in Persian. Arabic users of Persian names, however, would have to copy carefully the diacritical marks of the specifically Persian letters "pe", "che", "zhe" and "gāf" if they wished to maintain the integrity of the original Persian written forms.

When a country has to use names from a writing system other than its own,² it is faced with the necessity of using either a transliteration system, that is, a one-to-one substitution of graphic symbols, or a transcription system, that is, a one-to-one substitution of symbols representing the sounds of the donor language.

In practice, the use of transliteration is indicated when the receiver language is written by means of an alphabet and the alphabet of the donor language represents a good phonemic notation for that language.

The use of transcription is indicated when the donor language's alphabet is not one in which letters correspond

to the language's phonemes or when the receiver language is not written alphabetically; for example, Chinese.

In transliteration, the existence of graphic distinctions which the receiver alphabet does not have necessitates the use of diacritical marks so that there may be a one-to-one correspondence between donor and receiver graphs. For instance, since Persian has four letters representing the "z" sound, diacritical marks must be devised to differentiate the four (cf. the z-graphs "z̄", "z", "ż" and "z̈" in the BGN/PCGN³ system for Persian).

In transcription, the existence in the donor language of individual sounds or classes of sounds that the receiver language does not have necessitates the use of diacritical marks to account for all the sounds of the donor language. (cf. in the BGN (Modified Wade-Giles system for Chinese) the distinction between "chu" and "chü", "ch'u" and "ch'ü", "lu", and "lü" and "yu" and "yü").

As far as possible, in both transliteration and transcription, diacritical marks should be used in such a way that classes of sounds or contrasts such as that between short and long vowels will be systematically distinguished (cf. in the BGN/PCGN Arabic system the contrast by means of the cedilla of the nonvelarized consonants "d", "h", "s", "t" and "z" and the corresponding velarized consonants "ḍ", "ḥ", "ṣ", "ṭ" and "ẓ" and also the contrast by means of the macron of the short vowels "a", "i" and "u" and the corresponding long vowels "ā", "ī" and "ū").

Theoretical considerations such as those brought forward above cannot always determine the nature of a transliteration or transcription system, especially when systems already in use have almost universal currency within a country or throughout a writing system area. Nonetheless, they are of great value in the evaluation or improvement of existing systems and should always be kept in mind when new systems are worked out.

² See discussion of writing systems in *United Nations Conference on the Standardization of Geographical Names*, vol. 1 (United Nations publication, Sales No.: E.68.1.9), p. 22.

³ Board of Geographic Names/Permanent Committee on Geographic Names.

ROMANIZATION OF CHINESE CHARACTERS FOR PLACE NAMES

Paper presented by the United States of America*

The China Topographic Service and the United States Army Map Service have combined their efforts to produce a manuscript entitled "Modified readings of Chinese characters for place names Romanization based on the modified Wade-Giles system", which will be published soon. It represents an important step forward in research on Chinese geographical names in that it largely eliminates the need for English-speaking persons to consult Chinese lexical works, many of which are not readily available and most of which do not give an accurate description of the pronunciation of their entries.

English-speaking nations have relied almost exclusively on the Wade-Giles system for the transcription of Chinese, and sources utilizing this system have generally used the Romanization appearing in the Giles dictionary. The Giles dictionary was based essentially on Mandarin

pronunciation. The *Kuo-yin Ch'ang-yung Tz'u-hui* (*Manual of Chinese national Romanization of frequently used characters*, Shanghai, 1932), which uses basically the same pronunciation but with some modifications, has been designated by all Chinese Governments since its publication as the official standard for the pronunciation of Chinese.

Using this official and national standard pronunciation, we are able to assign Roman-letter equivalents to Chinese characters that accurately reflect the segmental phonemes that appear in the pronunciation of these characters, for which the Wade-Giles system is adequate. Thus, while the Romanization used in the Giles dictionary is inadequate for national standardization, its sound-to-symbol system is adequate.

It is this current sound-to-symbol relationship that is important in *Modified readings of Chinese characters for place names Romanization based on the modified Wade-Giles system*. Implicit in its publication is the assumption

* The original text of this paper, prepared by G. F. Beasley, Office of Geography, Department of the Interior, appeared as document E/CONF.53/L.29.