

E

UNITED
NATIONS



Economic and Social Council

Distr.
LIMITED

E/CONF.91/CRP.39
16 January 1998

ENGLISH ONLY

SEVENTH UNITED NATIONS CONFERENCE
ON THE STANDARDIZATION OF
GEOGRAPHICAL NAMES
New York, 13-22 January 1998
Item 12 (a) of the provisional agenda*

WRITING SYSTEMS AND GUIDES TO PRONUNCIATION:
ROMANIZATION

• International Romanization Standards

— Paper submitted by the
International Organization for Standardization (ISO) **

* E/CONF.91/1
** Prepared by Mr. John Clews, Chair of ISO/TC46/SC2: Conversion
of Written Languages

International romanization standards

1. ISO/TC46/SC2 (Conversion of Written Languages)

ISO/TC46/SC2 (Conversion **of** Written Languages) is the International Organization for Standardization subcommittee responsible for romanization. This last met from 12-14 May 1997 at the British Standards Institution in Chiswick, London, to review international standards in this area - both already published and under development. It next meets in Athens in May 1998.

This report provides an overview of ISO/TC46/SC2 activities. To avoid repetition, further details are available in the more formal Summary report on the work of ISO/TC46/SC2, June 1997 (also circulated to the Seventh United Nations Conference on the Standardization of Geographical Names in New York, 12-23 January 1998). It also outlines advantages of cooperation between the United Nations Conference on the Standardization of Geographical Names and/or the United Nations Group of Experts on Geographical Names and ISO/TC46/SC2.

2. Conversion of Geographical Names

It has always been the intention of ISO/TC46/SC2 that its standards provide for Conversion of Geographical Names (as well as other purposes). Most ISO/TC46/SC2 standards have text stating that "the procedure is often used for historical or geographical texts, cartographical documents and in particular bibliographical work where characters must be converted from different writing systems into a single alphabet to allow for alphabetical intercalation in bibliographies, catalogues, indexes, toponymic lists, etc." Most of these aims of ISO/TC46/SC2 are also aims shared by the United Nations Conference on the Standardization of Geographical Names, and the United Nations Group of Experts on Geographical Names. The irony is that both "sides" have been working in isolation (both at international and national levels), often without knowing of the existence of the other group. For some languages, the two groups will have produced standards which result in identical or near-identical romanized text (as in Chinese and Greek), and in others there will be considerable differences, but for no good reason other than they were "not invented here." Where there are differences for other scripts, this could be despite input from the "host" country for that script, as different bodies would have been involved from within that country.

The United Nations Conference on the Standardization of Geographical Names and the United Nations Group of Experts on Geographical Names have produced romanization standards for some scripts, and ISO/TC46/SC2 has produced standards for some other scripts, but both have gaps where input from the other "side" might be beneficial. The United Nations standards in this area are not recognized as ISO standards by the International Organization for Standardization, despite the fact that these UN standards may be more widely used than some of the ISO standards in this area.

The rest of this report summarises how ISO/TC46/SC2 works, and makes suggestions for more productive liaison between ISO/TC46/SC2, the United Nations Conference on the Standardization of Geographical Names and the United Nations Group of Experts on Geographical Names.

3. How ISO/TC46/SC2 works

The International Organization for Standardization (ISO) is an international organization based in Geneva. Its membership is international, although not all UN members are members of ISO, and consists of the national bodies responsible for standardization in each member country. Many of these national member bodies are governmental organizations or quasi-governmental organizations.

ISO standards are developed by working groups of nominated experts, reviewed by national member bodies of ISO (national standards bodies) and published by ISO Central Secretariat, in an attempt to ensure quality and consensus in ISO standards, in a process which takes at least a couple of years.

ISO standards are developed by consensus, requiring strict review by working group members and by national member bodies during development and during revision.

ISO/TC46/SC2 currently has ten working groups, each dealing with one or more scripts. Over the last year or so, some draft standards have been developed faster, and with input from more experts, using electronic communications, via the World Wide Web and email. This is important in gaining consensus, and ISO/TC46/SC2 now has a register of over 350 experts and users of romanization from 45 countries worldwide.

4. Standards and draft standards developed by ISO/TC46/SC2

ISO/TC46/SC2 has produced 14 standards or mature draft standards that have covered several stages of voting, (including an ISO technical report and some additional simplified and phonemic versions of the base standard).

These are Cyrillic (ISO 9:1995); Arabic (ISO 233:1984; ISO 233-2:1993; ISO/DIS 233-3); Hebrew (ISO 259:1984; ISO 259-2:1994; ISO/DIS 259-3); Greek (ISO 843:1997); Japanese (ISO 3602:1989); Chinese (ISO 7098:1991); Georgian (ISO 9984:1996); Armenian (ISO 9985:1996); Thai (ISO/DIS 11940); and Korean (ISO/TR 11941:1996).

There are also some less mature draft standards:

ISO NP 15922 (Simplified transcription of Thai) and ISO NP 15923 (Phonemic transcription of Persian) are under development, but has not yet reached voting stages.

ISO CD 14522 (Transliteration of Mongolian) was sent for initial voting, but failed its initial voting stages, and a new draft is under development.

ISO NP 15919 (Transliteration of Devanagari and related Indic scripts) is under development, but has not yet reached voting stages. There are also three more generalized standards under development which are not limited to any one script, and which have not yet had draft standards sent for voting:

ISO NP 15920: Use of diacritical marks for conversion purposes;

IS0 NP 15921: Generalized conversion methods; and

IS0 NP 15924: Codes for representation of names of scripts

5. Underlying assumptions in ISO/TC46/SC2 standards

The following criteria have been used in developing recent romanization standards: pronunciation, convention, readability and visual similarity. These are also criteria adopted by CEN/TC304, the relevant European Standards Committee, in its standards relating to transformation of non-Latin characters.

To date, only conversion into Latin script (romanization) has been undertaken by ISO/TC46/SC2, although there is latent interest into conversion into Cyrillic script.

The two basic methods of romanization are transliteration (converting characters in one script into characters in another script) and transcription (converting the sounds of one language into characters associated with that sound in another language which uses another script). Transcription is not strictly reversible. For romanization in some languages there may be no differences between transliterated and transcribed text: for romanization in other languages there may be major differences, and in these scripts it is important to appreciate the differences. Published ISO/TC46/SC2 standards have all used transliteration as the basis, although additional tables, or additional parts in some draft standards, also provide transcription information, which may result in a different romanization. In general, a one-to-one ratio between source characters and target characters has been aimed at, although this is under review, and is not the case for romanization of any Chinese, Japanese and Korean characters.

ISO/TC46/SC2's standards have also aimed at ensuring reversibility where possible, so that the correct non-Latin characters can be correctly deduced from transliterated Latin text.

Recently there has been a trend away from a lot of heavily-accented characters, and towards as much use as possible of plain letters in texts, with digraphs (two Latin letters representing one original letter) where necessary. This trend is converging with the practices of the United Nations Conference on the Standardization of Geographical Names and the United Nations Group of Experts on Geographical Names and its Working group on Romanization.

6. The importance of script conversion

To some extent, increased availability of non-Latin software means that geographical names can be input, printed and displayed using the original characters used for spelling those names, including non-Latin script characters. However, despite computing standards like ISO/IEC 10646 and Unicode, there will always be a need for transliteration as long as people do not have the same level of competence in all scripts besides the script used in their mother-tongue, and may have a need to deal with these languages, or when they have to deal with mechanical or computerized equipment which does not provide all the scripts of characters that they need.

7. Can improved romanization standards be achieved by cooperation?

ISO/TC46/SC2 was represented at an earlier United Nations Conference on the Standardization of Geographical Names, but there appears to have been no contact since then.

Cooperation between the UN-based experts and the ISO-based experts could allow more expertise to be brought to bear on the languages and/or scripts under consideration, with benefit to each expert.

It might also be possible to arrive at identical romanizations which suited each group, and would therefore suit most users. This might resolve the anomaly for users and staff in some map libraries, for instance, where the catalogue of maps applied one form of romanization, as used in libraries, while the maps themselves applied another form of romanization.

8. Initiatives already taken by ISO/TC46/SC2

I have already proposed in ISO/TC46/SC2 that new standards should include capital city or country names as examples, rather than random words in the source language, in order to provide examples known to most users, as in the draft on Mongolian script.

I have also made some contacts on email lists concerning maps to try and identify experts who could play a role in the work of ISO/TC46/SC2.

On the UK national committee for ISO/TC46/SC2, which I also chair, Paul Woodman of the UK's permanent Committee on Geographical Names is also a committee member.

I also hope that it may be possible for a representative of ISO/TC46/SC2 to take part in the UNGEGN Working Group on Romanization.

9. Possible initiatives by UNGEGN and the UN Conference

It would be useful for the United Nations Conference on the Standardization of Geographical Names and the United Nations Group of Experts on Geographical Names to consider Liaison membership of ISO/TC46/SC2. As can be seen from the Summary report on the work of

ISO/TC46/SC2, June 1997 (also circulated to the Seventh United Nations Conference on the Standardization of Geographical Names in New York, 12-23 January 1998) as well as country members there are a number of Liaison members, including various organizations of the United Nations System.

This would give the United Nations Conference on the Standardization of Geographical Names and/or the United Nations Group of Experts on Geographical Names the right to influence and comment on drafts and revisions of ISO/TC46/SC2 standards, and also the possibility of "Fast Tracking" some of its own standards within ISO/TC46/SC2.

10. ISO/TC46/SC2 Contact details:

John Clews
(Chair of ISO/TC46/SC2: Conversion of Written Languages)
SESAME Computer Projects,
8 Avenue Road,
Harrogate,
HG2 7PG,
United Kingdom.
Email: Converse@sesame.demon.co.uk;
tel: +44 (0) 1423 888 432

Evangelos Melagrakis
(Secretary of ISO/TC46/SC2: Conversion of Written Languages)
ELOT,
313 Acharnon Str.,
GR-111 45 Athens,
Greece
Email: eem@elot.gr
tel: +30 1 201 9890

Web site: <http://www.elot.gr/tc46sc2>;
Electronic mailing list for ISO/TC46/SC2 (tc46sc2@elot.gr).