

INTERNATIONAL STANDARD

**ISO
7098**

Second edition
1991-12-15

Information and documentation — Romanization of Chinese

Information et documentation — Romanisation du chinois



Reference number
ISO 7098 : 1991 (E)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

International Standard ISO 7098 was prepared by Technical Committee ISO/TC 46, *Information and documentation*, Sub-Committee SC 2, *Conversion of written languages*.

This second edition cancels and replaces the first edition (ISO 7098 : 1982), which has been technically revised.

Annex A forms an integral part of this International Standard.

© ISO 1991

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the publisher.

International Organization for Standardization

Case postale 56 • CH-1211 Genève 20 • Switzerland

Printed in Switzerland

Information and documentation — Romanization of Chinese

1 Scope

This International Standard explains the principles of romanization of Modern Chinese or *pǔtōnghuà*, the official language of the People's Republic of China as defined in the *Directives for the promotion of pǔtōnghuà*, promulgated in 1956-02-06 by the State Council.

2 General principles of conversion of writing systems

2.1 The words in a language, which are written according to a given script (the converted system), sometimes have to be rendered according to a different system (the conversion system), normally used for a different language. This operation is often followed for historical or geographical texts, cartographical documents and in particular for bibliographical work in every case where it is necessary to write words supplied in various alphabets in a manner that allows intercalation with other words in a single alphabet so as to enable a uniform alphabetization to be made in bibliographies, catalogues, indices, toponymic lists, etc. It is indispensable in that it permits the univocal transmission of a written message between two countries using different writing systems or exchanging a message the writing of which is different from their own. It thereby permits transmission by manual as well as mechanical or electronic means.

The two basic methods of conversion of a system of writing are transliteration and transcription.

2.2 Transliteration is the operation which consists of representing the characters¹⁾ of an entirely alphabetical system of writing by the characters of the conversion alphabet.

In principle, this conversion should be made character by character: each character of the converted alphabet is rendered by one character, and one only of the conversion alphabet, this being to ensure the complete and unambiguous reversibility of the conversion alphabet into the converted alphabet.

When the number of characters used in the conversion system is smaller than the number of characters of the converted system, it is necessary to use digraph or diacritical marks. In this case one must avoid as far as possible arbitrary choice and the use of purely conventional marks and try to maintain a certain phonetic logic in order to give the system a wide acceptance.

However, it must be accepted that the graphism obtained may not always be correctly pronounced according to the phonetic habits of the language (or of all the languages) which usually use(s) the conversion alphabet. On the other hand, this graphism must be such that the reader who knows the converted language may mentally restore unequivocally the original graphism and thus pronounce it correctly.

2.3 Retransliteration is the operation which consists of converting the characters of a conversion alphabet to those of the converted alphabet. This operation is the exact opposite of transliteration; it is carried out by applying the rules of a system of transliteration in reverse order so as to reconstitute the transliterated word to its original form.

2.4 Transcription is the operation which consists of representing the characters of a language, whatever the original system of writing, by the phonetic system of letters or signs of the conversion language.

A transcription system is of necessity based on the orthographical conventions of a conversion language and its alphabet. The users of a transcription system must therefore have a knowledge of the conversion language to be able to pronounce the characters correctly. Transcription is not strictly reversible.

Transcription may be used for the conversion of all writing systems. It is the only method that can be used for systems that are not entirely alphabetical and for all ideophonographic writing systems (Chinese, Japanese, etc.).

1) A character is an element of a writing system, whether or not alphabetical, that represents a phoneme, a syllable, a word or even prosodical characteristics of the language, by using graphical symbols (letters, diacritical marks, syllabic signs, punctuation marks, prosodical accents, etc.) or a combination of these signs (a letter having an accent or a diacritical mark, for example à, é or ò, is therefore a character in the same way as a basic letter).

2.5 To carry out romanization (the conversion of non-Latin writing systems to the Latin alphabet) it is possible to use either transliteration or transcription or a combination of these two methods, according to the nature of the converted system.

2.6 A conversion system proposed for international use may call for compromise and the sacrifice of certain national customs. It is therefore necessary for each national community of users to accept concessions, fully abstaining in every case from imposing as a matter of course solutions that are actually justified only by national practice (for example as regards pronunciation, orthography, etc.). However, these concessions would obviously not relate to the use that a country makes of its national writing system: when this national system is not converted, the characters constituting it must be accepted in the form in which they are written in the national language. When a country uses two systems univocally convertible one into the other to write its own language, the system of transliteration thus implemented must be taken *a priori* as a basis for the international standardized system, as far as it is compatible with the other principles mentioned hereafter.

2.7 Where necessary, the conversion systems should specify an equivalent for each character, not only the letters but also the punctuation marks, numbers, etc. They should similarly take into account the arrangement of the sequence of characters that make up the text, for example the direction of the script, and specify the way of distinguishing words and of using separation signs and capital letters, following as closely as possible the customs of the language(s) which use the converted writing system.

3 Principles for converting ideophonographic scripts

3.1 The structure of ideophonographic scripts, where conveyance of meaning is of greater importance than that of pronunciation, entails the existence of a large number of characters (more than 40 000 in the case of Chinese), thus making sign by sign transliteration impossible and resulting in the need to devise a system of transcription. Each character must therefore be transcribed by one or more Latin letters standing for the pronunciation or pronunciations of the character in question. This means that the transcriber must be familiar with the reading or readings of the text to be transcribed.

3.2 Inasmuch as the transcription of ideophonographic scripts is merely a matter of phonetic notation of characters in Latin letters of the languages which use them, the identical characters will require different transcriptions depending on whether they are found in Chinese, Japanese or Korean texts.

3.3 On the other hand, the same character within the same language must always be transcribed in the same way regardless of the type of graphic representation utilized (full or simplified form of a Chinese character) except where a single character has more than one pronunciation.

3.4 Reversibility of romanization systems of ideophonographic scripts is impossible due to the following factors:

- the disparity in pronunciation of a given character in two different languages or within a single language;
- the high frequency of homophones within the same language;
- the possible coexistence of several writing systems within a given text.

3.5 In the case of those languages which use, even within the same text, more than one kind of script (for example *kana* and Chinese characters in Japanese, *hangul* and Chinese characters in Korean) both the transcription of the ideophonographic characters and the conversion of the other types of characters (for example *hangul*, *kana*) should yield a consistent and homogeneous system of romanization.

3.6 Although, as a rule, spacing between characters is regular, it is usual to transcribe the characters forming a single word by linking them together. Principles and rules for formation of words shall be standardized in other texts related to every language concerned.

3.7 Although there are no capital letters in ideophonographic scripts, it is usual when romanizing to capitalize some words, following the national uses.

4 Pinyin

Hànyǔ pīnyīn fāng'àn (Chinese phonetic system) or *pīnyīn*, which was officially adopted on 1958-02-11 by the National Assembly of the People's Republic of China, is used to transcribe Chinese. The transcriber writes down the pronunciation of the characters according to their readings in *pǔtōnghuà*.

5 Syllabic forms

5.1 Each Chinese character represents one syllable. One word may consist of one or more syllables.

5.2 The table of syllabic forms is given in annex A.

5.3 The following dictionary

現代漢語詞表 / 劉原主編.—北京：中國標準，1984.

Xiàndài Hànyǔ cíbiǎo / Liú Yuán zhǔbiān. — Běijīng : Zhōngguó biāozhǔn, 1984,

gives the transcription in *pīnyīn* of more than 100 000 words.

It is completed, for words for which the pronunciation is difficult, by

普通話異讀詞審音表 (1985 年 12月修訂) / 普通話審音委員會編.—北京：文字改革，1987.

Pǔtōnghuà yìdúcí shěnyīnbǎo (1985 nián 12 yuè xiūdìng) / Pǔtōnghuà shěnyīn wěiyuánhui biān.— Běijīng : Wénzì gǎigé, 1987.

A reference work for modern Chinese is the dictionary

現代漢語詞典 / 中國社會科學院語言研究所詞典編輯室編.—北京：商務印書館，1982.

Xiàndài Hànyǔ cídiǎn / Zhōngguó shèhuì kēxuéyuàn yǔyán yánjiūsuǒ cídiǎn biānjíshì biān.— Běijīng : Shāngwù yìnshūguǎn, 1982.

These works can be completed by

新華字典.—北京：商務印書館，1971.

Xīnhuá zìdiǎn.— Běijīng : Shāngwù yìnshūguǎn, 1971,

which gives the transcription of 8 500 characters, and by a concordance table of simplified characters and non-simplified characters

簡化字總表 (1986 年 新版).—北京：語文出版社，1986.

Jiǎnhuàzì zǒngbiǎo (1986 nián xīnbǎn).— Běijīng : Yǔwén chūbǎnshè, 1986.

6 Transcription of tones

6.1 Each syllable may have one of four tones or may be toneless. The four tones are marked by the following diacritic signs:

- first tone —
- second tone /
- third tone ✓
- fourth tone \

6.2 In the table of syllabic forms (see annex A), the syllables carry the mark of the first tone on the peak of the syllable. The first tone has been arbitrarily chosen to indicate on which letter the diacritic sign noting the tone must be written down.

6.3 Atony, sometimes called the fifth tone, is indicated by the lack of a diacritic sign.

6.4 Changes of tone induced by the tone of the next syllable are not shown.

6.5 Normally, tone marks are written; however, they may be omitted for convenience.

7 Punctuation

Punctuation marks similar to those existing in the sets of Latin characters are transcribed as such.

Chinese specific punctuation marks are transcribed as follows:

Chinese marks	Latin marks
。	. (full stop)
、	, (comma)
•	- (hyphen)
「	« (quotation mark left)
」	» (quotation mark right)

8 Numerals

Numerals written in Chinese characters are transcribed in *pīnyīn*. Numerals written in Arabic or Roman characters are kept as such.