INTERNATIONAL STANDARD

**ISO
233-2**

First edition
1993-08-15

Information and documentation — Transliteration of Arabic characters into Latin characters —

Part 2:
Arabic language — Simplified transliteration

Information et documentation — Translittération des caractères arabes en caractères latins —

Partie 2: Langue arabe — Translittération simplifiée

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 233-2 was prepared by Technical Committee ISO/TC 46, *Information and documentation*, Sub-Committee SC 2, *Conversion of written languages*.

ISO 233 consists of the following parts, under the general title *Information and documentation — Transliteration of Arabic characters into Latin characters*:

— *Part 1: Arabic language — Stringent transliteration* (Actually published as ISO 233:1984, *Documentation — Transliteration of Arabic characters into Latin characters.*)

— *Part 2: Arabic language — Simplified transliteration*

— *Part 3: Other than Arabic languages*

Annex A of this part of ISO 233 is for information only.

Introduction

ISO 233 is one of a series of International Standards, dealing with the conversion of systems of writing. The aim of ISO 233 and others in the series is to provide a means for international communication of written messages in a form which permits the automatic transmission and reconstitution of these by men or machines. The system of conversion, in this case, must be univocal and entirely reversible.

This means that no consideration should be given to phonetic and esthetic matters nor to certain national customs: all these considerations are, indeed, ignored by the machine performing the function.

The adoption of ISO 233 for international communication leaves every country free to adopt for its own use a national standard which may be different, on condition that it be compatible with ISO 233. The system proposed herein should make this possible; and be acceptable to international use if the graphisms it creates are such that they may be converted automatically into the graphisms used in any strict national systems.

ISO 233 may be used by anyone who has a clear understanding of the system and is certain that it can be applied without ambiguity. The result obtained will not give a correct pronunciation of the original text in a person's own language, but it will serve as a means of finding automatically the original graphism thus allowing anyone who has a knowledge of the original language to pronounce it correctly. Similarly one can only pronounce correctly a text written in, for example, English or Polish, if one has a knowledge of English or Polish.

The adoption of national standards compatible with ISO 233 will permit the representation, in an international publication, of the morphemes of each language according to the customs of the country where it is spoken. It will be possible to simplify this representation in order to take into account the number of the character sets available on different kinds of machines.

This page intentionally left blank

Information and documentation — Transliteration of Arabic characters into Latin characters —

Part 2:
Arabic language — Simplified transliteration

1 Scope

This part of ISO 233 establishes a simplified system for the transliteration of Arabic characters into Latin characters. This simplification of the stringent rules established by ISO 233:1984 is especially intended to facilitate the processing of bibliographic information (e.g. catalogues, indices, citations, etc.).

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. 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.

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 or mechanical means as well as by electronic means.

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

2.2 Transliteration is the process of representing the characters of an alphabetical or syllabic system of writing by the characters of a conversion alphabet.

In principle, the conversion shall be made character by character: each character[1] of the converted graphical system is rendered by only one character of the conversion alphabet, this being the easiest way to ensure the complete and unambiguous reversibility of the conversion alphabet in the converted system.

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 digraphs or diacritical marks. In this case, arbitrary choices and the use of purely conventional marks shall be avoided as far as possible, and a certain phonetic logic shall be maintained in order to give the system a wide acceptance.

However, it must be accepted that the graphism obtained will 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 shall be such that the reader who has a knowledge of the converted language can mentally restore unequivocally the original graphism and thus pronounce it.

2.3 Retransliteration is the process whereby the characters of a conversion alphabet are transformed back into those of the converted writing system. It is the exact opposite of the transliteration process in that the rules of a transliteration system are applied

---

1) A character is an element of an alphabetical or other type of writing system that graphically represents a phoneme, a syllable, a word or even a prosodical characteristic of a given language. It is used either alone (for example a letter, a syllabic sign, an ideographical character, a digit, a punctuation mark) or in combination (for example an accent, a diacritical mark). A letter having an accent or a diacritical mark, for example â, è, ö, is therefore a character in the same way as a basic letter.

in reverse so as to reconvert the transliterated word to its original form.

2.4 Transcription is the process whereby the pronunciation of a given language is noted by the system of signs of a conversion language.

A transcription system is of necessity based on the orthographical conventions of the conversion language. 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 or syllabic and for all ideophonographical systems of writing like Chinese.

2.5 To carry out **romanization** (the conversion of non-Latin writing systems to the Latin alphabet), either transliteration or transcription or a combination of the two may be used depending on 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 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, regarding pronunciation, orthography, etc.).

When a country uses two systems to write its own language thus that are univocally convertible into each other, the system of transliteration implemented shall be taken a priori as a basis for the international standardized system, as far as it is compatible with the other principles addressed hereafter.

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

2.8 When romanizing a script which does not have upper case characters, it is usual to capitalize some words, following national usages.

3 Principles of conversion for alphabetical writing systems

3.1 The conversion may be made at various levels.

The first level is that of completely reversible **stringent transliteration** which is necessary to attain, in full, the aim given in 2.2. This conversion applies all principles of transliteration without exception. It does not permit variants. The conventional systems of stringent transliteration should be applied without any change to meet national or regional customs regarding pronunciation or orthography. They permit the univocal international transmission of messages by mechanical or electronic means.

To permit an international unequivocal communication, International Standards on transliteration must first apply the principle of stringent conversion. These can then be used as a basis for the establishment of rules for simplified conversion and for preparation of national standards.

The second level is that of **simplified conversion**. The simplification can be made necessary, for example, by the use of machines that do not accept all of the alphabetic characters required for stringent conversion. The method of conversion may allow national or regional variants, which can preclude complete reversibility. The simplified conversion may be the subject of International Standards or agreements.

The third level is that of **popular conversion** which, for example, should enable the same foreign names to be written in a uniform manner in the newspapers of a given country. It is obliged to take into account phonetic or graphic practices and therefore can only be national.

3.2 In cases where the same characters appear in one alphabet used with some differences by different languages, these characters would be transliterated in the same way, irrespective of the language they belong to.

3.3 If the converted alphabet gives a different form to the same character according to its place in the word (as is the case for example in the Arabic, Hebrew and Greek alphabets), the conversion alphabet will use only one character of constant form.

4 Transliteration tables

4.1 Preliminary notes

4.1.1 In order to render a transliterated text more legible, the vowels are supplied [method 2.1 c) of ISO 233:1984]. However sukūn (⁰) is omitted, as well as the vowels and diphthongs which have only a flexional function in nominal forms.

4.1.2 The initial alif (١) is not represented: the presence of an initial vowel in the transliterated word is enough to indicate an alif in the original script.

4.1.3 A character bearing a hamzaẗ (ﺀ), which depends on the vocalic context, is not represented.

4.1.4 The definite article (ال), is always represented by the characters "al-", whatever its vocalization.

4.1.5 The prepositions (li, bi, ka) as well as the conjunction wa, which in Arabic are joined to the word, are separated by a hyphen in the transliterated script.

4.1.6 It should be remembered that, as a general principle of transliteration, the phonetic value of the characters is not necessarily taken into account.

4.2 Tables

(The notes referred to in the tables are given in 4.3.)

Table 1 — Consonants

No.	Arabic character	Transliteration into Latin character	No.	Arabic character	Transliteration into Latin character
1	ا	when it bears a hamzaẗ and is in initial position: omitted when lengthening an a: ◌̄ (See table 2)	16	ض	ḍ
			17	ط	ṭ
2	ء	◌̕ (See note 1)	18	ظ	ẓ
3	ب	b	19	ع	◌̔
4	ت	t	20	غ	ġ
5	ث	ṯ	21	(ڢ) ف ¹⁾	f
6	ج	ǧ	22	(ڧ) ق ¹⁾	q
7	ح	ḥ	23	ك	k
8	خ	ẖ	24	ل	l
9	د	d	25	م	m
10	ذ	ḏ	26	ن	m
11	ر	r	27	ه	h
12	ز	z	27a	ة	ẗ (See note 4)
13	س	s	28	و	w (See also No. 31)
14	ش	š	29	ي	y (See also No. 32)
15	ص	ṣ	29a	ى	◌́ (See also No. 30b)

1) The letters in brackets are the Maghrebi forms.

Table 2 — Vowels and diphthongs

No.	Character		Character		Character		Character	
	Arabic	**Latin**	**Arabic**	**Latin**	**Arabic**	**Latin**	**Arabic**	**Latin**
30	´	a	آ	ā	آ̋ ̋	ā		
30a			ا	â				
30b			ـى	á	ـًى	á		
			ـٰى	á				
31	̦	u	ـو	ū	ـٰ	omitted	ـَوْ	aw
32	̦	i	ـري	ī	̋	omitted	ـَيْ	ay

Table 3 — Other conventional signs

No.	Arabic designation	Arabic character	Transliteration	Examples	
33	sukūn	ْ	omitted	أَلْ	al
34	šaddaẗ	ّ	duplication of the consonant	شّ	šš
35	hamzat al-waṣl	ٱ	omitted	وَفِي نُورِ الْقَمَرِ	
				fī nūr al-qamar	

Table 4 — Punctuation

Arabic character	Designation	Latin character
،	comma	,
؛	semicolon	;
؟	interrogation mark	?

4.3 Notes to tables 1 to 4

Note	Arabic character	Position	Transliteration	Examples	
1	hamzaẗ	with or without bearer (the bearer is not represented)	ʾ	بِــئْـرى	bi'r
				سَـأَلَ	sa'ala
				سَـاءَلَ	sāʾala
		except in initial position	omitted	أَخْـبَـارُ	aḫbār
2	maddaẗ		ʾā	قُـرآنْ	qur'ān
		except in initial position	ā	آدَابْ	ādāb
3	Definite article		al-	أَلشَّـمْـسُ	al-šams
				بِـاللَّـيْـلِ	bi-al-layl
		only when preceded by the preposition li-	l-	لِـلْـوَلَـدِ	li-l-walad
4	tā' marbūṭaẗ	in all positions	ẗ	مَـدِيـنَـةٌ	madīnaẗ

5 Capitalization

Arabic writing does not capitalize. In a transliterated text, capitalization may follow national practice.

The minimum rules are that the following are capitalized:

— the first letter of the first word of the title (even if it is the article al-);

— the first letter of personal names;

— in cases of a corporate body, at least the first letter of the first name.