

INTERNATIONAL
STANDARD

ISO
9

Second edition
1995-02-15

**Information and documentation —
Transliteration of Cyrillic characters into
Latin characters — Slavic and non-Slavic
languages**

*Information et documentation — Translittération des caractères cyrilliques en
caractères latins — Langues slaves et non slaves*



Reference number
ISO 9:1995(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 9 was prepared by Technical Committee ISO/TC 46, *Information and documentation*, Subcommittee SC 2, *Conversion of written languages*.

This second edition cancels and replaces the first edition (ISO 9:1986), of which it constitutes a technical revision.

Annexes A to D of this International Standard are for information only.

© ISO 1995

All rights reserved. Unless otherwise specified, 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

Introduction

This International Standard is one of a series of International Standards, dealing with the conversion of systems of writing. The aim of this International Standard 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 aesthetic matters nor to certain national customs: all these considerations are, indeed, ignored by the machine performing the function.

The adoption of this International Standard 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 the International Standard. The system proposed herein should make this possible, and be acceptable for international use if the graphisms it creates are such that they may be converted automatically into the graphisms used in any national system, so long as it is strict.

This International Standard 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 and thus allow 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 this International Standard 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 extent of the character sets available on different kinds of machine.

Information and documentation — Transliteration of Cyrillic characters into Latin characters — Slavic and non-Slavic languages

1 Scope

This International Standard establishes a system for the transliteration into Latin characters of Cyrillic characters constituting the alphabets of Slavic and non-Slavic languages, in accordance with the principles of stringent conversion in order to permit international information exchange, particularly by electronic means. For the transliteration of Slavic Cyrillic characters, tables 1 and 2 reproduce the tables published in the first edition of ISO 9:1986; for the transliteration of Cyrillic characters constituting the alphabets of non-Slavic languages, table 3 adopts the transliteration of tables 1 and 2 for all characters similar to those of Slavic languages and gives equivalents for all supplementary characters introduced in the alphabets of non-Slavic languages.

Table 3 includes in a single sequence, listed in the Cyrillic alphabetic order, the 118 single or diacritic-carrying characters that appear in one or another of the considered alphabets. The list of the languages written in these alphabets is given in annex C.

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, mechanical as well as electronic means.

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

2.2 Transliteration is the process which consists of representing the characters¹⁾ of an alphabetical or syllabic writing by the characters of a conversion alphabet.

In principle, the conversion should be made character by character: each character 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 cannot 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 may mentally restore unequivocally the original graphism and thus pronounce it.

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 (e.g. a letter, a syllabic sign, an ideographical character, a digit, a punctuation mark) or in combination (e.g. 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.

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 in reverse in order 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 (regarding pronunciation and orthography).

When a country uses two systems univocally convertible one into the other to write its own language, the system of transliteration thus implemented shall be taken a priori as a basis for the international standardized system, as far as it is compatible with the other principles exposed 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. 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, following as closely as possible the customs of the language(s) which use the converted writing system.

2.8 When romanizing a script which has no upper-case characters, it is usual to capitalize some words, following national usage.

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 as such without any change to meet national or regional customs as regards 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 apply by priority 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 the alphabet characters required for stringent conversion. The method of conversion may allow national or regional variants, which may not permit 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 table

Table 1 — General table for Slavic Cyrillic characters

Cyrillic character				Transliteration into Latin characters from Cyrillic alphabets (Bulgarian, Byelorussian, Macedonian, Russian, Serbo-Croatian, Ukrainian)	Respective languages	Examples			
No.	printed		written						
1	а	А	<i>a</i>	<i>А</i>	a	A	all	адрес	adres
2	б	Б	<i>б</i>	<i>Б</i>	b	B	all	баба	baba
3	в	В	<i>в</i>	<i>В</i>	v	V	all	вы	vy
4	г	Г	<i>г</i>	<i>Г</i>	g	G	all	голова	golova
5	д	Д	<i>д</i>	<i>Д</i>	d	D	all	да	da
6	ђ	Ђ	<i>ђ</i>	<i>Ђ</i>	đ	Đ	sr	ђон	đon
7	ѓ	Ѓ	<i>ѓ</i>	<i>Ѓ</i>	ǰ	Ǧ	mk	ѓуѓум	ǰuǰum
8	е	Е	<i>e</i>	<i>Е</i>	e	E	all	еда	eda
9	ë	Ë	<i>ë</i>	<i>Ë</i>	ë	Ë	be ru	ëлка	ëlka
10	є	Є	<i>є</i>	<i>Є</i>	ê	Ê	uk	твоє	tvoê
11	ж	Ж	<i>ж</i>	<i>Ж</i>	ž	Ž	all	журнал	žurnal
12	з	З	<i>з</i>	<i>З</i>	z	Z	all	звезда	zvezda

Table 1 — (continued)

Cyrillic character					Transliteration into Latin characters from Cyrillic characters of Slavic alphabets (Bulgarian, Byelorussian, Macedonian, Russian, Serbo-Croatian, Ukrainian)		Respective languages	Examples	
No.	printed		written						
13	С	С	с	С	Ŝ	Ŝ	mk	свезда	ŝvezda
14	И	И	и	И	ı	ı	bg mk ru sr uk	книга	kniga
15	і	І	і	Ј	ì	Ì	be uk	білий	bilij
16	ï	Ї	ï	Ї	ï	Ï	uk	їзда	izda
17	ј	Ј	ј	Ј	ǰ	ǰ	mk sr	један	ǰedan
18	й	Й	й	Й	ı	ı	be bg ru uk	первый	pervyj
19	к	К	к	К	k	К	all	как	kak
20	л	Л	л	Л	l	L	all	липа	lipa
21	љ	Љ	љ	Љ	î	Î	mk sr	љубав	îubav
22	м	М	м	М	m	M	all	муж	muž
23	н	Н	н	Н	n	N	all	нижний	nižnij
24	њ	Њ	њ	Њ	ň	Ň	mk sr	њива	ňiva

Table 1 — (continued)

No.	Cyrillic character				Transliteration into Latin characters from Cyrillic characters of Slavic alphabets (Bulgarian, Byelorussian, Macedonian, Russian, Serbo-Croatian, Ukrainian)		Respective languages	Examples	
	printed	written							
25	О	О	о	О	o	O	all	общество	obšestvo
26	п	П	п	П	p	P	all	пара	para
27	р	Р	р	Р	r	R	all	рыба	ryba
28	с	С	с	С	s	S	all	сестра	sestra
29	т	Т	т	Т	t	T	all	товариш	tovariš
30	ћ	Ћ	ћ	Ћ	ć	Ć	sr	кућа	kuća
31	ќ	Ќ	ќ	Ќ	ќ	Ќ	mk	куќа	kuća
32	у	У	у	У	u	U	all	утро	utro
33	ў	Ў	ў	Ў	ŭ	Ŭ	be	слоўнік	sloŭnik
34	ф	Ф	ф	Ф	f	F	all	физика	fizika
35	х	Х	х	Х	h	H	all	химический	himičeskij
36	ц	Ц	ц	Ц	c	C	all	центральный	central'nyj

Table 1 — (concluded)

No.	Cyrillic character				Transliteration into Latin characters from Cyrillic characters of Slavic alphabets (Bulgarian, Byelorussian, Macedonian, Russian, Serbo-Croatian, Ukrainian)		Respective languages	Examples	
	printed		written						
37	ч	Ч	ч	Ч	č	Č	all	часы	časy
38	џ	Џ	џ	Џ	đ	Đ	mk sr	џамија	đamiĵa
39	ш	Ш	ш	Ш	š	Š	all	школа	škola
40	щ	Щ	щ	Щ	š̂	Š̂	bg ru uk	щит	šit
41	ъ	Ъ	ъ	Ъ	''	''	bg ru	объявление	ob"âvlenie
42	ы	Ы	ы	Ы	y	Y	be ru	был	byl
43	ь	Ь	ь	Ь	'	'	be bg ru uk	альбом	al'bom
44	э	Э	э	Э	è	È	be ru	это	èto
45	ю	Ю	ю	Ю	û	Û	be bg ru uk	южный	ûžnyj
46	я	Я	я	Я	â	Â	be bg ru uk	яма	âma
47	’	’	’	’	’	’	be mk uk	’рѓа	’rga

NOTE – For the diacritical signs used, see annex A.

Table 2 — Complementary table for the Slavic Cyrillic characters used by some communities established outside the boundaries of their native countries

Cyrillic character				Transliteration into Latin characters from Cyrillic characters of Slavic alphabets	Examples	
No.	printed		written			
48	Г	Г	г Г	ǵ Ǵ	густ	ǵust
49	Ѣ	Ђ	ѣ Ђ	ě Ě	ѣсть	ěstʹ
50	Ѧ	Ѧ	ѧ Ѧ	ǎ Ǟ	мѧка	mǎka
51	Ѧ	Ѧ	ѧ Ѧ	f̂ F̂	каѦдра	kaf̂dra
52	Ѧ	Ѧ	ѧ Ѧ	ÿ ÿ̂	мѦро	mÿro

NOTE – Character 48 is also used in the Ukraine.

Table 3 — Table for Cyrillic characters of non-Slavic languages

No.	Cyrillic character	Transliteration into Latin characters from Cyrillic characters	No.	Cyrillic character	Transliteration into Latin characters from Cyrillic characters	No.	Cyrillic character	Transliteration into Latin characters from Cyrillic characters	No.	Cyrillic character	Transliteration into Latin characters from Cyrillic characters	No.	Cyrillic character	Transliteration into Latin characters from Cyrillic characters
1	а	a	25	ѐ	ẓ	49	М	m	73	ć	ś	97	Ч	ç
2	ä	ä	26	ë	z̄	50	Н	n	74	т	t	98	ч	ç̄
3	ǎ	ǎ	27	з	ż	51	ң	ŋ	75	т̣	ṭ	99	ч̣	ç̣
4	ạ̌	ạ̌	28	И	i	52	ң̣	ŋ̣	76	т̇	ṫ	100	ч̇	ç̇
5	ā	ā	29	й	ī	53	ң̣̇	ŋ̣̇	77	т̣̇	ṭ̇	101	е	ç̣̇
6	æ	æ	30	й̇	í	54	ң̣̣̇	ŋ̣̣̇	78	у	u	102	е̇	ç̣̣̇
7	á	á	31	й̇̇	î	55	ң̣̇̇	ŋ̣̇̇	79	ү	ü	103	ш	ç̣̇̇
8	å	å	32	й̇̇̇	j̇	56	ң̣̇̇̇	ŋ̣̇̇̇	80	ү̇	ū	104	щ	ș
9	б	b	33	і	ì	57	ң̣̇̇̇̇	ŋ̣̇̇̇̇	81	ү̇̇	ú	105	ъ	”
10	в	v	34	і̇	ï	58	о	o	82	ү̇̇̇	ú̇	106	ы	у̇
11	г	g	35	і̇̇	ĳ	59	ö	ö	83	ү̇̇̇̇	ú̇̇	107	ы̇	ü̇
12	г̇	ġ	36	і̇̇̇	j̇̇	60	ø	ø	84	ү̇̇̇̇̇	ü̇̇̇	108	ы̇̇	ü̇̇̇
13	г̇̇	ġ̇	37	і̇̇̇̇	j̇̇̇	61	ö̇	ö̇	85	ү̇̇̇̇̇̇	ú̇̇̇̇	109	ь	·
14	ђ	g̣	38	к	k	62	ö̇̇	ö̇̇̇	86	ү̇̇̇̇̇̇̇	ú̇̇̇̇̇	110	э	è
15	ђ̣	h	39	қ	q̣	63	ö̇̇̇	ö̇̇̇̇	87	ү̇̇̇̇̇̇̇̇	ū̇̇̇̇̇	111	э̇	è̇
16	д	d	40	қ̣̇	ḳ̇	64	ó	ó	88	w	w	112	э̇̇	à
17	е	e	41	қ̣̇̇	ḳ̇̇	65	ō	ō	89	ф	f	113	ю	û
18	е̇	ë	42	қ̣̇̇̇	ḳ̇̇̇	66	œ	œ	90	х	h	114	ю̇	û̇
19	е̇̇	ë̇	43	қ̣̇̇̇̇	q̣̇̇̇̇	67	п	p	91	х̣	ḥ	115	я	â
20	ж	ž	44	қ̣̇̇̇̇̇	ḳ̇̇̇̇̇	68	п̣	p̣	92	ц	c	116	я̇	â̇
21	ж̣	ẓ̌	45	қ̣̇̇̇̇̇̇	q̣̇̇̇̇̇̇	69	п̣̇	ṗ̣	93	ц̣	c̣	117	я̇̇	·
22	ж̣̇	ż	46	л	l	70	р	r	94	ц̣̇	ç̇	118	”	..
23	ж̣̇̇	ż̇	47	л̇	l̇	71	р̇	s	95	ч	ç			
24	з	z	48	л̇̇	l̇̇	72	ç	ş	96	ч̣	ç̣			

Annex A

(informative)

Diacritical signs used for the transliteration of Slavic Cyrillic characters (tables 1 and 2)

The diacritical signs used in tables 1 and 2 of this International Standard are taken from the code table of ISO 5426.

Table A.1 indicates their position in that code table.

Table A.1 — Diacritical signs used in tables 1 and 2

No.	Position in the code table	No.	Position in the code table
7	4/2	40	4/3
10	4/3	41	3/14
11	4/15	43	3/13
13	4/3	44	4/1
15	4/1	45	4/3
16	4/8	46	4/3
17	4/15	1)	
21	4/3	48	4/1
24	4/3	49	4/15
30	4/2	50	4/15
31	4/2	51	4/1
33	4/15	52	4/1
37	4/15		
38	4/3		
39	4/15		

1) Character 47 is an apostrophe coded 2/7 in ISO/IEC 646.

Annex B (informative)

Diacritical signs used for the transliteration of Cyrillic characters of non-Slavic languages (table 3)

The diacritical signs used are taken from the code table of ISO 5426. Table B.1 indicates their position in the code table.

Table B.1 — Diacritical signs used in table 3

No. of transliterated characters	Position in the code table	Name
2, 19, 26, 34, 59, 79, 98, 107	4/8	Diaeresis
3, 62, 84	4/8 + 5/6	Diaeresis + Dot below
4, 14, 18, 23, 56, 101	4/6	Breve
5, 22, 29, 42, 57, 65, 80, 93, 108	4/5	Macron
6	7/1	Ligature ae – Small letter
7, 12, 27, 30, 37, 47, 55, 64, 68, 83	4/2	Acute accent
8, 114	4/10	Circle above
13, 28, 32, 53, 86	4/7	Dot above
15, 39, 52, 97	5/6	Dot below
102	4/6 + 5/3	Breve + Hook to right
20, 35, 36, 41, 76, 81, 95, 103	4/15	Háček
21	4/15 + 5/2	Háček + Hook to left
25, 43, 48, 51, 72, 77, 91, 96	5/2	Hook to left
31, 40, 60, 94, 99, 104, 113, 115	4/3	Circumflex accent
33, 44, 54, 63, 69, 73, 75, 85, 100, 110, 112	4/1	Grave accent
61, 82, 111	4/13	Double acute accent
87	4/5 + 5/6	Macron + Dot below
66	7/10	Ligature oe – Small letter
105	3/14	Tverdyj znak
109	3/13	Mjagkij znak
116	3/6	Double dagger
117	4/1	Grave accent (used alone)
118	4/8	Diaeresis (used alone)

Annex C

(informative)

List of languages covered by table 3

Russian name	English name	French name
абазинский	Abaza	abaza
абхазский	Abkhaz	abkhaze
аварский	Awar	avar
адыгейский (черкесский)	Adyge (Circassian)	adyghé (circassien)
азербайджанский	Azeri	azéri
алтайский	Altay	altaïen
балкарский	Balkar	balkar
башкирский	Bashkir	bachkir
бурятский	Buryat	bouriate
гагаузский	Gagauzi	gagaouze
даргинский	Dargwa	dargwa
долганский	Dolgan	dolgane
дунганский	Dungan	doungane
ингушский	Ingush	ingouche
кабардинский (черкесский)	Kabardian (Circassian)	kabarde (circassien)
казахский	Kazakh	kazakh
калмыцкий	Kalmyk	kalmouk
караимский	Karaim	caraiite
каракалпакский	Karakalpak	karakalpak
карачаевский	Karachay	karatchaï
карельский	Karelian	carélien
кетский	Ket	kète
коми-зыряцкий	Komi-Zyrian	komi-zyriène
коми-пермяцкий	Komi-Permian	komi-permien
корякский	Koryak	koriak
крымско-татарский	Crimean Tatar	tatar de Crimée
кумыкский	Kumyk	koumyk
курдский	Kurdish	kurde
кыргызский	Kirgiz	kirghize
лакский	Lak	lak
лезгинский	Lezgian	lezghien
мансийский	Mansi	mansi
марийский (горный)	Mari (high)	mari (haut)
марийский (луговой)	Mari (low)	mari (bas)
молдавский	Moldavian	moldave
монгольский	Mongolian	mongol
мордовско-мокшанский	Mordvin-Moksha	mordve-mokcha
мордовско-эрзянский	Mordvin-Erza	mordve-erza
нанайский	Nanay	nanai
нганасанский	Nganasan	nganassane
ненецкий	Nenets	nénetse
нивхский	Nivkh	nivkhe
ногайский	Noghay	nogai

Russian name	English name	French name
осетинский	Ossetic	ossète
саамский	Saam	same
селькупский	Selkup	selkoupe
табасаранский	Tabasaran	tabassarane
таджикский	Tajik	tadjik
татарский	Tatar	tatar
татский	Tat	tate
тофаларский	Tafalar	tofalar
тувинский	Tuva	touva
туркменский	Turkmen	turkmène
удмуртский	Udmurt	oudmourte
удэ(ге)йский	Udegey	oudégué
узбекский	Uzbek	ouzbek
уйгурский	Uyghur	ouïgour
ульчский	Ulch	oultche
хакасский	Khakass	khakasse
хантыйский (ваховский)	Vakh Khanty	khanty de Vakh
хантыйский (казымский)	Kazym Khanty	khanty de Kazym
хантыйский (сургутский)	Surgut Khanty	khanty de Sourgout
хантыйский (шурькшарский)	Shurykshar Khanty	khanty de Chourykchar
цыганский	Tsigan	tsigane
чеченский	Chechen	tchéthène
чувашский	Chuvash	tchouvache
чукотский	Chukcha	tchouktche
шорский	Shor	chor
шугнанский	Shugnan	chougnane
эвенкийский	Evenki	evenke
эвенский	Even	evène
эскимосский	Eskimo	esquimau
юкагирский	Yukagir	ioukaguir
язгулямский	Yazgulam	iazgoulame
якутский	Yakut	iakoute

Annex D
(informative)

Bibliography

- [1] ISO/IEC 646:1991, *Information technology — ISO 7-bit coded character set for information interchange*.
- [2] ISO 5426:1983, *Extension of the Latin alphabet coded character set for bibliographic information interchange*.

ICS 01.140.10

Descriptors: documentation, transliteration, letters (symbols), Cyrillic characters, Latin characters.

Price based on 13 pages
