

INTERNATIONAL STANDARD

**ISO
843**

First edition
1997-01-15

Corrected and reprinted
1999-05-01

Information and documentation — Conversion of Greek characters into Latin characters

*Information et documentation — Conversion des caractères grecs en
caractères latins*



Reference number
ISO 843:1997(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 843 was prepared by Technical Committee ISO/TC 46, *Information and documentation*, Subcommittee SC 2, *Conversion of written languages*.

It cancels and replaces ISO Recommendation R 843:1968, of which it constitutes a technical revision.

Annex A forms an integral part of this International Standard. Annexes B and C are for information only.

© ISO 1997

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

INFORMATION AND DOCUMENTATION - CONVERSION OF GREEK CHARACTERS INTO LATIN CHARACTERS

1 SCOPE

This International Standard establishes a system for the transliteration and/or transcription of Greek characters into Latin characters. This system provides for two sets of rules, each constituting a type of conversion, which are:

- Type 1, transliteration of Greek characters into Latin characters
- Type 2, transcription of Greek characters into Latin characters

This system applies to the characters of the Greek script, independent of the period in which it is or was used, i.e. it applies to monotoniko and polytoniko scripts from all periods of Classic or Modern Greek (archaic, alexandrian, hellenistic, byzantine, katharevousa, dimotiki etc.) as well as any other form of writing using the Greek script.

Definitions and explanations for the terms used (transliteration, transcription etc.) appear in Annex A.

This International Standard does not strictly define which type of conversion shall be used in an application. An application that needs some kind of mapping of Greek characters into Latin characters can choose one, and only one, of these types for a specific purpose. The application must explicitly declare the type adopted.

This International Standard, however, recommends a preferable way of using the various types of conversion:

- Type 1 (transliteration) can be used in information exchange of bibliographic references, directory services and cases where the original text may be unavailable but still needs to be reconstituted. This type can be used in every case that requires the unique conversion of the transliterated word to its original form (Greek characters) by persons not knowing the Greek language or by machines (e.g. from information technology applications, telematic services and others from the area of communications).

- Type 2 (transcription) can be used in identity cards, passports, road signs, map citations and other cases where the correct pronunciation of the Greek word takes precedence over the need for reconstitution of the original form. In cases like these, other means of ensuring the relationship between the original and the converted word can be used.

A table showing one method for reversible transcription is given in Annex B (Informative). Other methods for reversible transcription may also be used.

The user should be aware that ordering problems may occur as a result of the transliteration/transcription process.

2 CHARACTER REPERTOIRES INVOLVED IN THE CONVERSION PROCESS

In this International Standard two character repertoires are identified:

- a. The complete graphic character repertoire used in any form of writing of the Greek script. This repertoire constitutes the collection of all non-combined characters that appear in Tables 9, 33 and 34 of ISO/IEC 10646-1:1993 plus 6 characters (Stigma, Digamma, Koppa, Sampi, Byzantino/Lunate Sigma and Yot) from Table 10 and one character (dialytika) from Table 8 of the International Standard ISO/IEC 10646-1 : 1993.
- b. A graphic character repertoire used by the Latin script. This includes graphic characters that are in usage in many Latin-oriented languages. The Latin characters are derived from table 1 and table 2 of ISO/IEC 10646-1: 1993. The character "Combining macron", derived from Table 8 of the ISO/IEC 10646-1, can be used in certain applications. (An alternative source for the Latin repertoire is ISO/IEC 8859-1).

The graphical representation of Greek and Latin characters used in the transliteration / transcription tables is indicative. The correct representation is beyond the purpose of this International Standard.

The references mentioned do not impose any requirements on the coding of the characters, nor do they restrict the ability of those who perform the conversion (human beings or machines) to handle more characters than those included in these two repertoires.

3 TRANSLITERATION AND TRANSCRIPTION TABLES

3.1 Transliteration and transcription of Alphabetical characters.

Table 1: Transliteration

	GREEK CHARACTER	LATIN CHARACTER
1	A, α	A, a
2	B, β	V, v
3	Γ, γ	G, g
4	Δ, δ	D, d
5	E, ε	E, e
6	Z, ζ	Z, z
7	H, η	$\bar{I}$, $\bar{i}$ or I^- , i^-
8	Θ, θ	TH, th
9	I, ι	I, i
10	K, κ	K, k
11	Λ, λ	L, l
12	M, μ	M, m
13	N, ν	N, n
14	Ξ, ξ	X, x
15	O, ο	O, o
16	Π, π	P, p
17	P, ρ	R, r
18	Σ, σ, ς	S, s, s (2)
19	T, τ	T, t
20	Υ, υ	Y, y
21	Φ, φ	F, f
22	X, χ	CH, ch
23	Ψ, ψ	PS, ps
24	Ω, ω	$\bar{O}$, $\bar{o}$ or O^- , o^-

Note 1 - Combinations of two or more characters are transliterated according to the provisions laid down for each independent character. The only exception to this rule for transliteration is the conversion of the Greek double vowels AY, αυ, EY, ευ, OY, ου which are mapped into Latin as AU, au, EU, eu, OU, ou respectively.

Note 2 - The Greek character ς (σίγμα τελικό – sigma final) is transliterated into Latin s, the same as for the Greek character σ (σίγμα – sigma). The character σ is used at the beginning or in the middle of a word, while the character ς is used at the end of the word.

The transliterating medium (human or machine) should, when in the process of converting the transliterated text back to the original Greek text, read the character of the string that follows s. In cases where the Latin s is followed by an alphabetic character or hyphenation then the s is transliterated into σ. In any other case (for example when it is followed by a space, mark, symbol, etc.) the Latin s is transliterated into the Greek character ς.

Table 2: Transcription

GREEK CHARACTERS			LATIN CHARACTERS
	1	2	3
	Greek characters	Combinations of Greek characters	Transcription
1	A, α		A, a
2		(AI, αι) (8)	AI, ai
3		(AI, άι) (8)	AI, ai
4		(AI, αι̃) (8)	AI, aĩ
5		AY, αυ (3) (10)	AV, av
6		AY, αυ (4) (10)	AF, af
7		AY, αυ (9)	AY, ay
8	B, β		V, v
9	Γ, γ		G, g
10		ΓΓ, γγ	NG, ng
11		(ΓK, γκ) (8)	GK, gk
12		ΓΞ, γξ	NX, nx
13		ΓΧ, γχ	NCH, nch
14	Δ, δ		D, d
15	E, ε		E, e
16		(EI, ει) (8)	EI, ei
17		(EI, έι) (8)	EI, ei
18		(EI, ει̃) (8)	EI, eĩ
19		EY, ευ (3) (10)	EV, ev
20		EY, ευ (4) (10)	EF, ef
21		EY, ευ (9)	EY, ey
22	Z, ζ		Z, z
23	H, η		I, i
24		HY, ηυ (3) (10)	IV, iv
25		HY, ηυ (4) (10)	IF, if
26		(HY, ηυ) (9)	IY, iy
27	Θ, θ		TH, th
28	I, ι		I, i
29	K, κ		K, k
30	Λ, λ		L, l
31	M, μ		M, m
32		MP, μπ (5) (7)	B, b
33		MP, μπ (6)	MP, mp
34	N, ν		N, n
35		(NT, ντ) (8)	NT, nt
36	Ξ, ξ		X, x
37	O, ο		O, o
38		(OI, οι) (8)	OI, oi
39		(OI, όι) (8)	OI, oi
40		(OI, οι̃) (8)	OI, oĩ
41		OY, ου	OU, ou
42		(OY, ου) (9)	OY, oy
43	Π, π		P, p
44	P, ρ		R, r
45	Σ, σ, ς		S, s, s (2)
46	T, τ		T, t
47	Υ, υ		Y, y
48		(YI, υι) (8)	YI, yi
49	Φ, φ		F, f
50	X, χ		CH, ch
51	Ψ, ψ		PS, ps
52	Ω, ω		O, o

Notes to table 2

Note 3 - Used before the characters β, γ, δ, ζ, λ, μ, ν, ρ and all the vowels.

Note 4 - Used before the characters θ, κ, ξ, π, σ, τ, φ, χ, ψ and at the end of the word.

Note 5 - At the beginning of the word.

Note 6 - In the middle of the word.

Note 7 - At the end of the word.

Note 8 - The combinations of Greek characters, that are in parentheses, are presented in Table 2 only for clarification. They are converted according to the provisions laid down for each independent character.

Note 9 - They are converted according to the provisions laid down for each independent character when the vowel before υ has an accent or when the υ has dialytika.

Note 10 - In the conversion of the combinations αυ, ευ, ηυ, when they have an accent, the accent is transferred to the vowel (e.g. αυ, αf, ιv, ιf, vυ, vι).

General notes applying to both transliteration and transcription:

Note 11 - If the conversion of a Greek letter gives rise to a double capital in Latin which is followed by lowercase text (for example at the beginning of a sentence after a full stop mark) the second, third, etc. Latin capital letter may for aesthetic reasons be rendered in lowercase (e.g. **Chara** instead of **CH**ara, **Thalassa** instead of **TH**alassa, **Psari** instead of **PS**ari).

Note 12 - The converting medium must be aware that the special mark (macron) above or after the Latin characters is obligatory for the correct conversion of the character. However, the size of this special mark and its accurate graphical representation or position is not the subject of this International Standard. Use of another character to substitute macron where it is unavailable, is not prohibited but it is done at the user's risk.

3.2 Conversion of non-alphabetical characters (used in transliteration and transcription)

The non-alphabetical characters, as defined in this International Standard, are converted as follows (There are no differences in transliteration or transcription concerning the conversion of the non - alphabetical characters. The non-alphabetical characters are divided into those used in the modern monotoniko form of writing and those used in any other form of writing using the Greek script):

Table 3: Conversion of non-alphabetical characters

GREEK CHARACTER			LATIN CHARACTER
	Monotoniko	Polytoniko	Transliteration / Transcription
1	´ (TONOS)		´ (ACUTE ACCENT)
2		´ (OXEIA)	´ (ACUTE ACCENT)
3		` (VAREIA)	` (GRAVE ACCENT)
4		˜ (PERISPOMENI)	ˆ (CIRCUMFLEX)
5	¨ (DIALYTIKA)	¨ (DIALYTIKA)	¨ (DIAERESIS)
6	; (EROTIMATIKO)	; (EROTIMATIKO)	? (QUESTION MARK)
7	. (TELEIA)	. (TELEIA)	. (FULL STOP)
8	· (ANO TELEIA)	· (ANO TELEIA)	; (SEMI COLON)
9	: (ANO - KATO TELEIA)	: (ANO - KATO TELEIA)	: (COLON)
10	, (KOMMA)	, (KOMMA)	, (COMMA)
11	! (THAYMASTIKO)	! (THAYMASTIKO)	! (EXCLAMATION MARK)
12	' (APOSTROFOS)	' (APOSTROFOS)	' (APOSTROPHE)
13		' (KORONIS)	' (APOSTROPHE)
14		' (PSILI)	' (APOSTROPHE)
15		ˆ (DASEIA)	h
16		˘ (YPOGEGRAMMENI)	˘ (CEDILLA)
17		˘ (PROSGEGRAMMENI)	˘ (CEDILLA)
18		˘ (ENOTIKON)	- (HYPHEN)

Table 4: Conversion of Archaic characters

	Archaikoi Charaktires	Transliteration / Transcription
1	Ɔ (DIGAMMA)	W
2	Ͱ (GIOT)	j
3	Ϛ (BYZANTINO SIGMA)	S

Table 5: Greek numerical symbols and their correspondence to decimal numbers

Α' α' =1	Ι' ι' =10	Ρ' ρ' =100
Β' β' =2	Κ' κ' =20	Σ' σ' =200
Γ' γ' =3	Λ' λ' =30	Τ' τ' =300
Δ' δ' =4	Μ' μ' =40	Υ' υ' =400
Ε' ε' =5	Ν' ν' =50	Φ' φ' =500
Ζ' ζ' =6	Ξ' ξ' =60	Χ' χ' =600
Ζ' ζ' =7	Ο' ο' =70	Ψ' ψ' =700
Η' η' =8	Π' π' =80	Ω' ω' =800
Θ' θ' =9	Ϛ' =90	ϛ' =900
Α α =1000	Ι ι =10000	Ρ ρ =100000
Β β =2000	Κ κ =20000	Σ σ =200000
Γ γ =3000	Λ λ =30000	Τ τ =300000
Δ δ =4000	Μ μ =40000	Υ υ =400000
Ε ε =5000	Ν ν =50000	Φ φ =500000
Ζ ζ =6000	Ξ ξ =60000	Χ χ =600000
Ζ ζ =7000	Ο ο =70000	Ψ ψ =700000
Η η =8000	Π π =80000	Ω ω =800000
Θ θ =9000	Ϛ =90000	ϛ =900000

Note 13 - Laws, decrees and other documents that have been indexed by Greek numerical symbols are converted using decimal numbers for their indexing and their correspondence is given in table 5.

Note 14 - Any other non-alphabetical character that may be found in the Greek text (e.g. makron, vrachy, etc.) should not be converted.

4 SPECIAL CASES

First names, surnames, names of places or street names that are of non-Greek origin are transcribed without taking into consideration their original writing with Latin or other characters.

Annex A (Normative)

A INTRODUCTION

A.1 Standards on conversion of systems of writing

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 to international use if the graphisms it creates are such that they may be converted automatically into the graphisms used in any strict national system.

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 number of the character sets available on different kinds of machines.

A.2 General principles of conversion of writing systems

A.2.1 Definition and methods

A.2.1.1 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.

A.2.1.2 Transliteration is the process which consists of representing the characters (1) of an alphabetical or syllabic system of writing by the characters of a conversion alphabet, this being the easiest way to ensure the complete and unambiguous reversibility of the conversion alphabet in the converted system.

In exceptional cases, e.g. 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, 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 has a knowledge of the converted language may mentally restore unequivocally the original graphism and thus pronounce it.

A.2.1.3 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.

A.2.1.4 Transcription is the process whereby the sounds of a given language are 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.

A.2.1.5 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.

A.2.2 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 as regards pronunciation, orthography, etc.).

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 exposed hereafter.

A.2.3 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.

A.2.4 When romanizing a script which does not have upper-case characters, it is usual to capitalize some words, following national usage.

(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.

A.3 Principles of conversion for alphabetical writing systems

A.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 clause A.2.1.2. This conversion applies all principles of transliteration without exception. However, whenever it is useful to distinguish the end or beginning of a syllable (a morpheme or a word) variants may be used. The conventional systems of stringent transliteration shall 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 shall apply by priority the principle of stringent conversion. They 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 may 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.

A.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.

A.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.

Annex B

(informative)

Column 4 defines reversible transcription. Column 5 gives guidance for pronunciation according to the International Phonetic Alphabet (IPA).

GREEK CHARACTERS			LATIN CHARACTERS		
	1	2	3	4	5
	Greek characters	Combinations of Greek characters	Transcription	Reversible Transcription	Pronunciation according to IPA
1	A, α		A, a	A, a	α
2		(ΑΙ, αι) (8)	AI, ai	AI, ai	ε
3		(ΑΙ, άι) (8)	AI, ai	AI, ai	αι
4		(ΑΙ, αι̃) (8)	AI, aĩ	AI, aĩ	αι̃
5		AY, αυ (3) (10)	AV, av	AV, av	αv
6		AY, αυ̃ (4) (10)	AF, af	AF, af	af
7		AY, αυ̃ (9)	AY, ay	AY, ay	α̃
8	B, β		V, v	V, v	v
9	Γ, γ		G, g	G, g	γ
10		ΓΓ, γγ	NG, ng	NG, ng	ηg
11		(ΓΚ, γκ) (8)	GK, gk	GK, gk	g (5), ηg (6) (7)
12		ΓΞ, γξ	NX, nx	NX, nx	ηx
13		ΓΧ, γχ	NCH, nch	NCH, nch	ηχ
14	Δ, δ		D, d	D, d	θ
15	E, ε		E, e	E, e	ε
16		(ΕΙ, ει) (8)	EI, ei	EI, ei	ει
17		(ΕΙ, έι) (8)	EI, ei	EI, ei	ει
18		(ΕΙ, ει̃) (8)	EI, eĩ	EI, eĩ	ει̃
19		EY, ευ (3) (10)	EV, ev	EV, ev	ev
20		EY, ευ̃ (4) (10)	EF, ef	EF, ef	ef
21		EY, ευ̃ (9)	EY, ey	EY, ey	ει
22	Z, ζ		Z, z	Z, z	z
23	H, η		I, i	I, i or I, i	i
24		HY, ηυ (3) (10)	IV, iv	IV, iv	iv
25		HY, ηυ̃ (4) (10)	IF, if	IF, if	if
26		HY, ηυ̃ (9)	IY, iy	IY, iy	ii
27	Θ, θ		TH, th	TH, th	θ
28	I, ι		I, i	I, i	ι
29	K, κ		K, k	K, k	k
30	Λ, λ		L, l	L, l	l
31	M, μ		M, m	M, m	m
32		ΜΠ, μπ (5) (7)	B, b	B, b	b
33		ΜΠ, μπ̃ (6)	MP, mp	MP, mp	mb
34	N, ν		N, n	N, n	n
35		(ΝΤ, ντ) (8)	NT, nt	Nt, nt	d (5), nd (6), (7)
36	Ξ, ξ		X, x	X, x	χ
37	O, ο		O, o	O, o	ο
38		(ΟΙ, οι) (8)	OI, oi	OI, oi	i
39		(ΟΙ, όι) (8)	OI, oi	OI, oi	oi
40		(ΟΙ, οι̃) (8)	OI, oĩ	OI, oĩ	oĩ
41		OY, ου	OU, ou	OU, ou	u
42		OY, ου̃ (9)	OY, oy	OY, oy	oi
43	Π, π		P, p	P, p	p
44	P, ρ		R, r	R, r	r
45	Σ, σ, ς		S, s, s (2)	S, s, s (2)	s, z, s (15)
46	T, τ		T, t	T, t	t
47	Υ, υ		Y, y	Y, y	I