

# **Extended Latin Alphabet Coded Character Set for Bibliographic Use**

**Abstract:** This standard establishes computer codes for an extended Latin alphabet character set to be used in bibliographic work when handling non-English items. The standard addresses special characters in languages using the Latin alphabet as well as combining marks (diacritics) required for romanization and transliteration. This standard establishes the 7-bit and 8-bit code values.

---

**Developed by the  
National Information Standards Organization**

**Approved May 3, 1993 by the  
American National Standards Institute**

---



Bethesda, Maryland, U.S.A.

**Published by  
NISO Press  
P.O. Box 1056  
Bethesda, MD 20827**

Copyright ©1993 by the National Information Standards Organization  
All rights reserved under International and Pan-American Copyright Conventions. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage or retrieval system, without prior permission in writing from the publisher. All inquiries should be addressed to NISO Press, P.O. Box 1056, Bethesda, MD 20827.

ISSN: 1041-5653  
ISBN: 1-880124-02-5

Printed in the United States of America



This paper meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).

#### **Library of Congress Cataloging-in-Publication Data**

Extended Latin alphabet coded character set for bibliographic use : American national standard extended Latin alphabet coded character set for bibliographic use : approved May 3, 1993, by the American National Standards Institute / developed by the National Information Standards Organization.

p. cm. — (National information standards series, ISSN 1041-5653)

“This standard may be identified by the use of the notation ANSEL”—Foreword.

ISBN 1-880124-02-5 (pbk)

1. Character sets (Data processing)—Standards—United States. 2. Machine-readable bibliographic data—Standards—United States. 3. Cataloging—Data processing—Standards—United States. I. American National Standards Institute. II. National Information Standards Organization (U.S.). III. Title: American national standard extended Latin alphabet coded character set for bibliographic use. IV. Title: ANSEL. V. Series.

Z699.35.C48E98 1993

025.3'16—dc20

92-7411  
CIP

## Contents

Foreword .....	iv
1. Scope, Purpose, and Application .....	1
1.1 Scope .....	1
1.2 Purpose .....	1
1.3 Application .....	1
2. Referenced Standards .....	1
2.1 American National Standards .....	1
2.2 ISO Standards .....	1
3. Definitions .....	2
4. Implementation .....	2
4.1 Unassigned Positions .....	2
4.2 Character Modifiers .....	2
5. Code Tables for the Extended Latin Alphabet Coded Character Set .....	3
5.1 7-Bit Code Table .....	3
5.2 8-Bit Code Table .....	3
6. Legend .....	6
Appendixes	
Appendix A Languages using ANSEL character modifiers or special characters .....	10
Appendix B ANSEL character modifiers and special characters with languages of occurrence .....	15
Figures	
Figure 1 7-Bit Code Table .....	4
Figure 2 8-Bit Code Table .....	5
Tables	
Table 1 Languages to which this standard may apply .....	1
Table 2 Languages to which this standard may apply in transliterations .....	2
Table 3 Character modifiers appearing in the ASCII set as spacing characters in the extended Latin alphabet coded character set .....	3
Table 4 ASCII characters indicated as having alternate names as character modifiers .....	3
Table 5 ANSEL spacing graphic characters .....	6
Table 6 ANSEL nonspacing graphic characters .....	7
Table 7 ASCII control characters .....	8
Table 8 ASCII graphic characters .....	9
Table A1 Latin alphabet script languages .....	10
Table A2 Transliterated non-Latin alphabet script languages .....	12
Table B1 Character modifiers .....	15
Table B2 Special characters .....	20

## Foreword

(This foreword is not a part of the American National Standard Extended Latin Alphabet Coded Character Set for Bibliographic Use, ANSI/NISO Z39.47-1993. It is included for information only.)

This standard for the extended Latin alphabet coded character set for bibliographic use was originally developed in 1985 by the Standards Committee on Coded Character Sets for Bibliographic Information Interchange of the American National Standards Committee on Library and Information Sciences and Related Publishing Practices, Z39, now the National Information Standards Organization.

The codes given in this standard may be identified by the use of the notation ANSEL. The notation ANSEL should be taken to mean the codes prescribed by the latest edition of this standard. To explicitly designate a particular edition of the standard when using the notation, the last two digits of the year of issue may be appended.

The standard establishes both the 7-bit and the 8-bit code values for the computer codes for characters used in bibliographic work when handling non-English items. The characters included in the codes have been selected because they are the ones needed to fully record bibliographic citations in many Latin alphabet languages and non-Latin languages transliterated into Latin alphabet characters.

The greater part of the extended Latin alphabet coded character set in this standard has been used for bibliographic work in the library community for over 20 years. Thus, it predates extended character set standardization that has recently been undertaken by the American National Standards Committee on Information Processing Systems, X3, and by the International Organization for Standardization, Technical Committee 46 on Information and Documentation and ISO-IEC JTC 1, Information Technology. The characters in column 4 (7-bit) or 12 (8-bit) are additions to this library set that have been identified as potentially useful in bibliographic work.

As of 1990, the library community used the 8-bit set specified in this standard (ASCII and ANSEL) with

the following exceptions: the characters defined in column 12 (8-bit) are not used; Greek characters a, b, and d and superscripts and subscripts for 0–9, (, ), +, - are added through escape sequences. The set just described constitutes what is commonly called the “ALA character set.” It is also the USMARC character set, and as such it is fully described in the publication *USMARC Specifications for Record Structure, Character Sets, Tapes*. It should be noted that the USMARC (or “ALA”) set could be changed from the above description, for example, to incorporate the characters in column 12, so the latest edition of the USMARC specifications document should be consulted for the exact specification.

At the time of reaffirmation, the text of ANSI/NISO Z39.47 was revised to a) delete references to six ANSI romanization standards which have been retired, b) incorporate three new definitions, c) clarify the language of the section dealing with character modifiers, and d) update/correct the appendixes. In the absence of published ANSI standards, it is recommended that the publication *ALA-LC Romanization Tables* be consulted for guidance on romanization and transliteration of non-roman scripts.

NISO acknowledges with thanks and appreciation the contributions of Randall K. Barry of the Library of Congress, Network Development and MARC Standards Office, in revising this standard.

Suggestions for improving this standard are welcome. They should be sent to the National Information Standards Organization, P.O. Box 1056, Bethesda, MD 20827, (301) 975-2814.

This standard was processed and approved for submittal to ANSI by the National Information Standards Organization. NISO approval of this standard does not necessarily imply that all Voting Members voted for its approval. At the time it approved this standard, NISO had the following members:

### NISO Voting Members

American Association of Law Libraries  
Gary J. Bravy

American Chemical Society  
Robert S. Tannehill, Jr.  
Leon R. Blauvelt (Alt)

American Library Association  
Myron Chace  
Glenn Patton (Alt)

American Psychological Association  
Maurine F. Jackson

American Society for Information Science  
Nolan Pope

American Society of Indexers  
Jessica Milstead  
Patricia S. Kuhr (Alt)

American Theological Library Association  
Myron Chace

Apple Computer, Inc.  
Karen Higginbottom

## Art Libraries Society of North America

Patricia J. Barnett  
Pamela J. Parry (Alt)

## Association of American University Presses

Mary Lou Menches

## Association of Information and Dissemination Centers

Bruce H. Kiesel

## Association for Information Management and Image Management

Marilyn Courtot

## Association of Jewish Libraries

Bella Hass Weinberg  
Pearl Berger (Alt)

## The Association for Recorded Sound Collections

Donald McCormick  
Barbara Sawka (Alt)

## Association of Research Libraries

Duane Webster

## AT&amp;T Bell Labs

M.E. Brennan

## Baker &amp; Taylor Books

Christian K. Larew  
Stephanie Lanzalotto (Alt)

## The Boeing Company

Steven C. Hill  
Michael Crandall (Alt)

## Book Manufacturers' Institute

Douglas Horner

## Catholic Library Association

Michael B. Finnerty

## CLSI, Inc.

Robert Walton  
Andy Lukes (Alt)

## Colorado Alliance of Research Libraries

Ward Shaw

## Data Research Associates, Inc.

Michael J. Mellinger  
James Michael (Alt)

## Dynix

Rick Wilson

## EBSCO Subscription Services

Sharon Cline McKay  
Mary Beth Vanderpoorten (Alt)

## Engineering Information Inc.

Eric Johnson  
Mary Berger (Alt)

## The Faxon Co., Inc.

Fritz Schwartz  
Joe Santosuosso (Alt)

## Gaylord Information Systems

Robert Riley  
Bradley McLean (Alt)

## IBM Corporation

Peggy Federhart

## Indiana Cooperative Library Services Authority

Barbara Evans Markuson  
Janice Cox (Alt)

## Library Binding Institute

Sally Grauer

## Library of Congress

Winston Tabb  
Sally H. McCallum (Alt)

## Mead Data Central

Peter Ryall  
Dave Withers (Alt)

## Medical Library Association

Rick B. Forsman  
Raymond A. Palmer (Alt)

## MINITEX

Anita Anker Branin  
William DeJohn (Alt)

## Music Library Association

Lenore Coral  
Geraldine Ostrove (Alt)

## National Agricultural Library

Joseph H. Howard  
Gary K. McCone (Alt)

## National Archives and Records Administration

Alan Calmes

## National Federation of Abstracting and Information Services

Ann Marie Cunningham  
Sarah Syen (Alt)

## National Institute of Standards and Technology, Office of Information Services

Jeff Harrison  
Marietta Nelson (Alt)

## National Library of Medicine

Lois Ann Colaanni

## OCLC, Inc.

Kate Nevins  
Don Muccino (Alt)

## OHIONET

Joel Kent  
Greg Pronevitz (Alt)

## Optical Publishing Association

John Nairn  
R. Bowers (Alt)

## PALINET

James E. Rush

## Pittsburgh Regional Library Center

Mary Lynn Kingston

## Readmore Academic Services

Sandra J. Gurshman  
Dan Tonkery (Alt)

## The Research Libraries Group, Inc.

Wayne Davison  
Kathy Bales (Alt)

## Society of American Archivists

Christine Ward  
Victoria Irons Walch (Alt)

## Software AG of North America, Inc.

James J. Kopp  
James E. Emerson (Alt)

## ANSI/NISO Z39.47-1993

Special Libraries Association  
Audrey N. Grosch

SUNY/OCLC Network  
Glyn T. Evans  
David Forsythe (Alt)

UMI  
Don Willis  
John Brooks (Alt)

Unisys Corporation  
Bill Payne

U.S. Department of Commerce, Printing and  
Publishing Division  
William S. Lofquist

U.S. Department of Defense, Defense Technical  
Information Center  
Margaret Brautigam  
Gretchen Schlag (Alt)

U.S. Department of Energy, Office of Scientific and  
Technical Information  
Mary Hall  
Nancy Hardin (Alt)

U.S. ISBN Maintenance Agency  
Emery Koltay

U.S. National Commission on Libraries and Information  
Science  
Peter Young  
Sandra N. Milevski (Alt)

VTLS  
Vinod Chachra

H.W. Wilson Company  
George I. Lewicky  
Ann Case (Alt)

## NISO Board of Directors

At the time NISO approved this standard, the following individuals served on its Board of Directors:

James E. Rush, Chairperson  
PALINET

Michael J. Mellinger, Vice Chair/Chair-elect  
Data Research Associates

Paul Evan Peters, Immediate Past Chairperson  
Coalition for Networked Information

Heike Kordish, Treasurer  
New York Public Library

Patricia R. Harris, Executive Director  
National Information Standards Organization

### *Directors Representing Libraries:*

Lois Ann Colaianne  
National Library of Medicine

Susan Vita  
Library of Congress

Shirley Kistler Baker  
Washington University

### *Directors Representing Information Services:*

Lois Granick  
American Psychological Association

Michael J. McGill  
University of Michigan

Wilhelm Bartenbach  
Engineering Information

### *Directors Representing Publishing:*

Peter J. Paulson  
OCLC/Forest Press

Constance U. Greaser  
American Honda

Marjorie Hlava  
Access Innovations, Inc.

# American National Standard Extended Latin Alphabet Coded Character Set for Bibliographic Use

## 1. Scope, Purpose, and Application

### 1.1 Scope

This standard specifies 63 graphic characters contained in a 94-byte set that can be invoked in a 7-bit or 8-bit environment. They are intended for use with the 128 graphic and control characters of ASCII, the American National Standard Code for Information Interchange, ANSI X3.4-1986 (R1992), and are therefore fully compatible with the 7-bit coded character set defined in that standard. (The relationship and use of multiple character sets is described in the American National Standard Code Extension Techniques for Use with the 7-Bit Coded Character Set of American National Standard Code for Information Interchange, ANSI X3.41-1990.)

This standard consists of code tables and a legend giving a name and an example for each of the extended Latin graphic characters.

### 1.2 Purpose

The character set in this standard is intended for the interchange of bibliographic information among data processing systems and within message transmission systems. It is suitable for bibliographic citations, including their annotations, in the Latin alphabet.

### 1.3 Application

The character set in this standard is intended to handle recorded information written in the Latin

alphabet in the languages listed in Table 1, among others. It is also intended to handle romanized forms of the languages listed in Table 2, among others.

Appendix A contains two tables showing the character modifiers and special characters coded in this standard as they are used for each language. Appendix B contains two tables showing the languages that use each character modifier or special character.

## 2. Referenced Standards

### 2.1 American National Standards

This standard is intended for use in conjunction with the following American National Standards. When these standards are superseded by a revision approved by the American National Standards Institute, Inc., the revision shall apply:

ANSI X3.4-1986 (R1992), Coded Character Set—7-Bit American National Standard Code for Information Interchange

ANSI X3.41-1990, Code Extension Techniques for Use with the 7-Bit Coded Character Set of ASCII

### 2.2 ISO Standards

This standard is intended for use in conjunction with Data Processing — Procedure for Registration of Escape Sequences, ISO 2375:1985.<sup>1</sup>

**Table 1. Languages to which this standard may apply**

Afrikaans	Esperanto	Indonesian	Slovak
Albanian	Estonian	Italian	Slovene
Anglo-Saxon	Faroese	Latvian	Spanish
Catalan	Finnish	Lithuanian	Swedish
Croatian	French	Navaho	Tagalog
Czech	German	Norwegian	Turkish (Modern)
Danish	Hawaiian	Polish	Vietnamese
Dutch	Hungarian	Portuguese	Wendic
English	Icelandic (Modern)	Romanian	

<sup>1</sup>Published by the International Organization for Standardization (ISO) and available from the American National Standards Institute (ANSI), 11 W. 42nd St., New York, NY 10036.

Table 2. Languages to which this standard may apply in transliterations

Amharic	Greek	Malayalam	Russian
Arabic	Gujarati	Marathi	Sanskrit
Armenian	Hebrew	Manipuri	Serbian
Assamese	Hindi	Mewari	Sindhi
Belorussian	Japanese	Nepali	Sinhalese
Bengali	Kannada	Oriya	Tamil
Braj	Khmer	Pahari	Telugu
Bulgarian	Konkani	Pali	Thai
Burmese	Korean	Panjabi	Tibetan
Chinese	Lahnda	Persian	Ukrainian
Church Slavic	Lao	Prakrit	Urdu
Dogri	Macedonian	Pushto	Yiddish
Georgian	Maithili	Rajasthani	

### 3. Definitions

**Character**—A member of a set of elements used for the organization, control, or representation of data.

**Character modifiers (diacritics)**—A mark, point, or sign used with alphabetic graphic characters to distinguish them in form or sound.

**Coded character set**—A set of unambiguous rules that establishes a character set and the one-to-one relationship between the characters of the set and their bit combinations.

**Control character**—A character whose occurrence in a particular context initiates, modifies, or stops an action that affects the recording, processing, transmission, or interpretation of data.

**Extended character set**—A set that includes characters that supplement those contained in a basic character set for a script.

**Graphic character**—A character, other than a control character, that has a visual representation normally handwritten, printed, or displayed.

**Nonspacing graphic character (combining character)**—A graphic character whose use is not followed by the forward movement of the output device. For the purpose of this standard, the term includes character modifiers.

**Punctuation mark**—A mark that indicates the structure of sentence parts for clarity; for example, . : ; .

**Spacing graphic character.** A graphic character whose use is followed by the forward movement of the output device to the next character position. For the purposes of this standard, the term includes special characters, special symbols, and punctuation marks.

**Special character**—An alphabetic character other than A-Z or other spacing graphic character; for example, æ.

**Special symbol**—A conventional sign used in place of words or word groups; for example, © % £.

### 4. Implementation

The implementation of this American National Standard is in accordance with the provisions of ANSI X3.4-1986. The 7-bit supplementary set is identified by escape sequences assigned by the ISO Registration Authority in accordance with procedures given in ISO 2375-1985.

#### 4.1 Unassigned Positions

The unassigned positions in the code tables shall not be used.

#### 4.2 Character Modifiers

Character modifiers, which are always used in conjunction with other characters, appear in columns 6 and 7 of the 7-bit extended Latin alphabet coded character set (columns 14 and 15 in an 8-bit environ-



ment). All characters in these two columns are designated nonspacing graphic characters. Thus, the backspace character, column 0/row 8 (col/row, 0/8), is not used with these character modifiers.

In a character string, these nonspacing characters shall precede the character that they modify. When a character requires multiple character modifiers, they are to be entered in the order in which they appear, reading left to right or top to bottom.

The character modifiers in Table 3 appear in the ASCII set as spacing characters in the extended Latin alphabet coded character set. The nonspacing form shall be used.

The ASCII characters in Table 4 are indicated as having alternative names (uses) as character modifiers in ANSI X3.4-1977. These ASCII characters shall only be used according to their primary names and not as character modifiers (requiring the backspace). Corresponding nonspacing character modifiers in the extended Latin alphabet coded character

set shall be used when needed. For example, use 2/2 for the quotation mark and 14/8 for the character modifier diaeresis.

## 5. Code Tables for the Extended Latin Alphabet Coded Character Set

### 5.1 7-Bit Code Table

The extended Latin alphabet coded character set in a 7-bit environment shall be as shown in Figure 1, 7-Bit Code Table. For the legend, see 6. Legend.

### 5.2 8-Bit Code Table

The extended Latin alphabet coded character set in an 8-bit environment shall be as shown in Figure 2, 8-Bit Code Table.

In columns 0 through 7, the table shows the characters of the 7-bit coded character set ASCII adapted for 8 bits. The extended Latin alphabet coded character set is shown in columns 10 through 15. For the legend, see 6. Legend.

**Table 3. Character modifiers appearing in the ASCII set as spacing characters in the extended Latin alphabet coded character set**

Name	ASCII	ANSEL	
	Col/Row	7-Bit Col/Row	8-Bit Col/Row
Circumflex (^)	5/14	6/3	14/3
Underline (_)	5/15	7/6	15/6
Tilde (~)	7/14	6/4	14/4

**Table 4. ASCII characters indicated as having alternate names as character modifiers**

Primary Name	ASCII	ANSEL	
	Col/Row	7-bit Col/Row	8-bit Col/Row
Quotation mark (diaeresis)	2/2	6/8	14/8
Apostrophe (acute accent)	2/7	6/2	14/2
Comma (cedilla)	2/12	7/0	15/0
Opening single quotation mark (grave accent)	6/0	6/1	14/1

Figure 1. 7-Bit code table

Bits 7→			0				1			
Bits 6,5→			00	01	10	11	00	01	10	11
Hex 1→			0	1	2	3	4	5	6	7
00	00	0	▨	▨	■	‘	°	▩	?	„
	01	1	▨	▨	Ł	ł	ł	▩	`	„
	10	2	▨	▨	Ø	ø	®	▩	’	.
	11	3	▨	▨	Ð	ð	©	▩	^	..
01	00	4	▨	▨	Þ	þ	#	▩	~	°
	01	5	▨	▨	Æ	æ	ı	▩	-	=
	10	6	▨	▨	Œ	œ	ı	▩	~	-
	11	7	▨	▨	/	//	▩	▩	.	„
10	00	8	▨	▨	·	ı	▩	▩	“	„
	01	9	▨	▨	ƒ	£	▩	▩	˘	„
	10	10	▨	▨	•	ð	▩	▩	°	˘
	11	11	▨	▨	±	▩	▩	▩	˘	˘
11	00	12	▨	▨	σ	σ	▩	▩	˘	▩
	01	13	▨	▨	U	u	▩	▩	,	▩
	10	14	▨	▨	,	▩	▩	▩	˘	,
	11	15	▨	▨	▩	▩	▩	▩	°	■
↑   ↑   ↑			↑							
4,3   2,1   2			↑							
Bits   Hex			ANSEL							



Reserved for control characters

Reserved for future standardization

Corners (reserved)

Figure 2. 8-Bit code table

Bits 8,7→  Bits 6,5→  Hex 1→			00				01				10				11			
			00	01	10	11	00	01	10	11	00	01	10	11	00	01	10	11
			0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
00	00	0	NUL	DLE	SP	0	@	P	`*	p	▨	▨	■	‘	°	☒	’	„
	01	1	SOH	DC1	!	1	A	Q	a	q	▨	▨	Ł	ł	/	☒	`	ˆ
	10	2	STX	DC2	"	2	B	R	b	r	▨	▨	Ø	ø	Ⓔ	☒	’	.
	11	3	ETX	DC3	#	3	C	S	c	s	▨	▨	Ɖ	đ	©	☒	^	..
01	00	4	EOT	DC4	\$	4	D	T	d	t	▨	▨	Ɔ	þ	#	☒	~	°
	01	5	ENQ	NAK	%	5	E	U	e	u	▨	▨	Æ	æ	ı	☒	-	=
	10	6	ACK	SYN	&	6	F	V	f	v	▨	▨	Œ	œ	ı	☒	~	-
	11	7	BEL	ETB	'	7	G	W	g	w	▨	▨	/	"	☒	☒	'	„
10	00	8	BS	CAN	(	8	H	X	h	x	▨	▨	·	ı	☒	☒	"	ˆ
	01	9	HT	EM	)	9	I	Y	i	y	▨	▨	ƒ	£	☒	☒	ˆ	ˆ
	10	10	LF	SUB	*	:	J	Z	j	z	▨	▨	Ⓢ	ð	☒	☒	°	ˆ
	11	11	VT	ESC	+	;	K	[	k	{	▨	▨	±	☒	☒	☒	ˆ	ˆ
11	00	12	FF	FS	,	<	L	\	l		▨	▨	Œ	σ	☒	☒	ˆ	☒
	01	13	CR	GS	-	=	M	]	m	}	▨	▨	Ů	ur	☒	☒	'	☒
	10	14	SO	RS	.	>	N	^*	n	~*	▨	▨	,	☒	☒	☒	ˆ	'
	11	15	SI	US	/	?	O	_*	o	DEL	▨	▨	☒	☒	☒	☒	ˆ	■
↑ ↑ ↑			↑				↑				↑				↑			
4,3 2,1 2			C0				G0 (ASCII)				C1				G1 (ANSEL)			
Bits Hex																		

- \* Redefined in the extended Latin alphabet coded character set
- /// Reserved for control characters
- ☒ Reserved for future standardization
- Corners (reserved)

## 6. Legend

Table 5 and Table 6 list 7-bit and 8-bit codes, sample graphics, names, and examples of use for each of the characters defined in this standard. Note: The legend for ASCII characters in Table 7 and Table 8 is adapted from ANSI X3.41-1986 and is included here only for completeness. For information on the use of ASCII refer to the latest edition of that standard.

Table 5. ANSEL spacing graphic characters

7-bit col/row	8-bit col/row	Graphic	Name	Example of use
2/1	10/1	Ł	slash L - uppercase	Łódź
2/2	10/2	Ø	slash O - uppercase	Øst
2/3	10/3	Đ	slash D - uppercase	Đuro
2/4	10/4	Þ	thorn - uppercase	Þann
2/5	10/5	Æ	ligature AE - uppercase	Ægir
2/6	10/6	Œ	ligature OE - uppercase	Œuvre
2/7	10/7	/	soft sign (mīāgkiī znak)	Fakul'tet
2/8	10/8	.	middle dot	novel·la
2/9	10/9	♭	musical flat	B♭
2/10	10/10	®	patent mark	ABC®
2/11	10/11	±	plus or minus	A±B
2/12	10/12	Œ	hook O - uppercase	BŒ
2/13	10/13	Ů	hook U - uppercase	XŮA
2/14	10/14	'	alif	Un'yusho
3/0	11/0	‘	‘ayn	fa‘il
3/1	11/1	ł	slash l - lowercase	rozbił
3/2	11/2	ø	slash o - lowercase	høj
3/3	11/3	đ	slash d - lowercase	đavola
3/4	11/4	þ	thorn - lowercase	þann
3/5	11/5	æ	ligature ae - lowercase	skæg
3/6	11/6	œ	ligature oe - lowercase	œuvre
3/7	11/7	"	hard sign (tvěrdyī znak)	ob"iāvlenie
3/8	11/8	ı	dotless i - lowercase	masalı
3/9	11/9	£	British pound	£5.00
3/10	11/10	ð	eth	verður
3/12	11/12	σ	hook o - lowercase	Sσ
3/13	11/13	υ	hook u - lowercase	Tυ Dυc
4/0	12/0	°	degree sign	10°C
4/1	12/1	ℓ	script l*	25 ℓ.
4/2	12/2	®	phono copyright mark	Decca ®
4/3	12/3	©	copyright mark	© 1974
4/4	12/4	#	musical sharp	D#
4/5	12/5	¿	inverted question mark	¿Qué?
4/6	12/6	¡	inverted exclamation mark	¡Esta!

\*In bibliographic work, the script l, ℓ, is commonly used as an abbreviation for the term "leaves." It shall not be used as a symbol for the unit of measure "liter."

Table 6. ANSEL nonspacing graphic characters

7-bit col/row	8-bit col/row	Graphic	Name	Example of use
6/0	14/0	◌̡	low rising tone mark	củi
6/1	14/1	◌̀	grave accent	règle
6/2	14/2	◌́	acute accent	está
6/3	14/3	◌̂	circumflex accent	même
6/4	14/4	◌̃	tilde	niño
6/5	14/5	◌̄	macron	gājējs
6/6	14/6	◌̆	breve	altā
6/7	14/7	◌̇	dot above	žaba
6/8	14/8	◌̈	umlaut (diaeresis)	öppna
6/9	14/9	◌̋	háček (caron)	vždy
6/10	14/10	◌̉	circle above (angstrom)	hår
6/11	14/11	◌̏	ligature, left half	akademiā
6/12	14/12	◌̐	ligature, right half	akademiā
6/13	14/13	◌̑	high comma, off center	rozdeľovac
6/14	14/14	◌̒	double acute accent	időszaki
6/15	14/15	◌̓	candrabindu	Aliiev
7/0	15/0	◌̔	cedilla	ça
7/1	15/1	◌̕	right hook	vieta
7/2	15/2	◌̖	dot below	teđa
7/3	15/3	◌̗	double dot below	khuṭbah
7/4	15/4	◌̘	circle below	Ṣaṃskṛta
7/5	15/5	◌̙	double underscore	Ghulam
7/6	15/6	◌̚	underscore	ṣamar
7/7	15/7	◌̛	left hoof	dārziṇa
7/8	15/8	◌̜	right cedilla	khոց
7/9	15/9	◌̝	half circle below	humantuṣ
			(upadhmaniya)	
7/10	15/10	◌̞	double tilde, left half	ṅalan
7/11	15/11	◌̟	double tilde, right half	ṅalan
7/14	15/14	◌̠	high comma, centered	g'eotermika

Table 7. ASCII control characters

Column/row	Mnemonic	Meaning
0/0	NUL	null
0/1	SOH	start of heading
0/2	STX	start of text
0/3	ETX	end of text
0/4	EOT	end of transmission
0/5	ENQ	enquiry
0/6	ACK	acknowledge
0/7	BEL	bell
0/8	BS	backspace
0/9	HT	horizontal tabulation
0/10	LF	line feed
0/11	VT	vertical tabulation
0/12	FF	form feed
0/13	CR	carriage return
0/14	SO	shift out
0/15	SI	shift in
1/0	DLE	data link escape
1/1	DC1	device control 1
1/2	DC2	device control 2
1/3	DC3	device control 3
1/4	DC4	device control 4
1/5	NAK	negative acknowledge
1/6	SYN	synchronous idle
1/7	ETB	end of transmission block
1/8	CAN	cancel
1/9	EM	end of medium
1/10	SUB	substitute
1/11	ESC	escape
1/12	FS	file separator
1/13	GS	group separator
1/14	RS	record separator
1/15	US	unit separator
7/15	DEL	delete

Table 8. ASCII graphic characters

Column/row	Graphic	Primary name (alternative name)
2/0		space (blank) [normally nonprinting]
2/1	!	exclamation point
2/2	“	quotation marks (diaeresis)
2/3	#	number sign
2/4	\$	dollar sign
2/5	%	percent sign
2/6	&	ampersand
2/7	'	apostrophe (closing single quotation mark; acute accent)
2/8	(	opening parenthesis
2/9	)	closing parenthesis
2/10	*	asterisk
2/11	+	plus
2/12	,	comma (cedilla)
2/13	-	hyphen (minus)
2/14	.	period (decimal point)
2/15	/	slant (slash)
3/0 to 3/9	0...9	digits 0 through 9
3/10	:	colon
3/11	;	semicolon
3/12	<	less than
3/13	=	equals
3/14	>	greater than
3/15	?	question mark
4/0	@	commercial at
4/1 to 5/10	A...Z	Latin letters A through Z - uppercase
5/11	[	opening bracket
5/12	\	reverse slant (backslash)
5/13	]	closing bracket
5/14	^	circumflex
5/15	_	underline
6/0	`	opening single quotation mark (grave accent)
6/1 to 7/10	a...z	Latin letters a through z - lowercase
7/11	{	opening brace (opening curly bracket)
7/12		vertical line (pipe)
7/13	}	closing brace (closing curly bracket)
7/14	~	tilde

## Appendix A

### Languages using ANSEL character modifiers or special characters

(This appendix is not part of American National Standard  
ANSI/NISO Z39.47-1993 and is included for information only.)

In Tables A1 and A2, the Latin alphabet script languages and transliterated non-Latin alphabet script languages are listed with ANSEL character modifiers and special characters found in each language. Only lowercase letters are shown unless the uppercase form would not be obvious, in which case, it is shown in parentheses. In cases in which character modifiers are used as pronunciation signs (as in Afrikaans or Vietnamese), only a few of the possible character modifiers with alphabetic character combinations are shown.

The transliteration characters are those required by the ALA-LC romanization tables.<sup>2</sup>

<sup>2</sup>The romanization schemes adopted by the American Library Association and the Library of Congress are published under the title: *ALA-LC Romanization Tables*. For further information contact the Cataloging Policy and Support Office, Library of Congress, Washington, DC 20540-4305.

**Table A1. Latin alphabet script languages**

Language	Character modifiers	Special characters
Afrikaans	ê î ô û ë ï è ò (etc.)	
Albanian	á é ó ú í â ê ë ç	
Anglo-Saxon	ā æ ē ī ō ū ę ƿ	æ œ þ ð
Catalan	à è ò é í ó ú ï ü ç	·
Croatian	č š ž ć í ʹ	đ
Czech	á é í ó ú ý č ď ě ň ř š ť ž d' (Ď) t' (Ť) ů ý ŷ	
Danish	å ĝ	æ ø
Dutch	â ê ë ï á é	
Esperanto	ĉ ĝ ĥ ĵ ŝ ŭ	
Estonian	ä ö ü š ž õ	
Faroese	á í ó ú ý	ø æ ǿ
Finnish	ä å	
French	à è ù â ê î ô û é ç ë ü ï	œ
German	ä ö ü	
Hawaiian	á é í ó ú ā ē ī ō ū	‘
Hungarian	á é í ó ú ő ű ö ü á	
Icelandic (Modern)	á é í ó ú ý ö	
Indonesian		,
Italian	à è ì ò ì	
Latvian	ā č ē ī ū ķ ļ ņ ģ š ž ģ (G)	

(Continued)



**Table A1. Latin alphabet script languages (Concluded)**

Language	Character modifiers	Special characters
Lithuanian	ą ę į ū č š è ē	
Navaho	è ñ ħ ą į á ê t' ġ ā	ł æ
Norwegian	å ä ö	æ ø
Polish	ą ę í ó ś ź ę ż ȳ	ł
Portuguese	ã ã ĩ õ ù á à ç ô ü	
Romanian	â î ș ȣ ț ă	
Slovak	á é í ó ú ý ľ í č ě ň ř š ž ď ť d' (Ď) l' (Ľ) t' (Ť) ů ä ô	
Slovene	č ě š ž á ô v'	
Spanish	á é í ó ú ñ ü	
Swedish	ä ö ř å	
Tagalog	á à â ñģ	
Turkish (Modern)	ö ü ç ş â û ğ i (İ) î (Î)	ı (I)
Vietnamese	â ê ô ã à ò á ă ă ā ư ã ă ậ ò, (etc.)	đ ơ ư
Wendic	ć ń ó p í ś w ź č ě š ž l' (Ľ) ò	

**Table A2. Transliterated non-Latin alphabet script languages**

[illegible]

Table A2. Transliterated non-Latin alphabet script languages (Continued)

Language	Character Modifiers	Special Characters
Lao	ā ī ū ē ā ē ō ō ē q̄ q	æ œ ' ‘
Macedonian	š č ž k g c	đ
Maithili	ā ī ū r̄ r̄ l̄ t̄ d̄ m̄ r̄ ṇ ḥ ṭ ṣ ḥ ẽ ă ǒ ê ô ñ ñ ñ ñ ś ḳ ḥ g̣ g̣ ḥ	,
Malayalam	ā ī ū ē ō r̄ l̄ r̄ t̄ l̄ l̄ t̄ d̄ ṇ ṣ ṁ ḥ á ñ ñ ś ṛ	,
Manipuri	ā ī ū r̄ l̄ t̄ d̄ r̄ m̄ ḥ ñ t̄ ȳ ñ ś ñ ñ	,
Marathi	ā ī ū r̄ r̄ l̄ l̄ t̄ d̄ ṇ ḥ ê ô ñ ñ ś ă ẽ ĩ ǒ ũ ă ĩ ũ ẽ ǒ	,
Mewari	ā ī ū r̄ r̄ l̄ t̄ d̄ m̄ r̄ ṇ ḥ ṭ ṣ ḥ ẽ ă ǒ ê ô ñ ñ ñ ñ ś ḳ ḥ g̣ g̣ ḥ	,
Nepali	ā ī ū r̄ r̄ l̄ t̄ d̄ m̄ r̄ ṇ ḥ ṭ ṣ ḥ ẽ ă ǒ ê ô ñ ñ ñ ñ ś ḳ ḥ g̣ g̣ ḥ	,
Oriya	ā ī ū r̄ r̄ l̄ l̄ t̄ d̄ ṇ ṁ ḥ ñ ȳ ñ ñ ñ ś	,
Pahari	ā ī ū r̄ r̄ l̄ t̄ d̄ m̄ r̄ ṇ ḥ ṭ ṣ ḥ ẽ ă ǒ ê ô ñ ñ ñ ñ ś ḳ ḥ g̣ g̣ ḥ	,
Pali	ā ī ū ñ t̄ d̄ ṇ ṁ ḥ ñ ñ ĩ	
Panjabi	ā ī ū t̄ d̄ ṇ ṁ ṛ ṁ ñ ñ ñ ḳ ḥ g̣	
Persian	š z ḥ š t̄ z̄ z̄ ī ā ū á	, ‘ /
Prakrit	ā ī ū r̄ r̄ l̄ t̄ d̄ m̄ ṇ ṣ ḥ ñ ñ ṁ ḥ ś ñ ḥ	,
Pushto	ā ū ī t̄ ḥ d̄ r̄ z̄ ṣ ṇ ṁ z̄ ṭ ṣ z̄ ś z̄ á	, ‘ /
Rajasthani	ā ī ū r̄ r̄ l̄ t̄ d̄ m̄ r̄ ṇ ḥ ṭ ṣ ḥ ẽ ă ǒ ê ô ñ ñ ñ ñ ś ḳ ḥ g̣ g̣ ḥ	,
Russian	ě f̄ ȳ ts̄ tē iū iā ī ĭ	/ //
Sanskrit	ā ī ū r̄ r̄ l̄ t̄ d̄ ṇ ṣ ṁ ḥ ṁ ḥ ñ ñ ś ñ ḥ	,
Serbian	š č ž k g c	đ

(Continued)

**Table A2. Transliterated non-Latin alphabet script languages** (*Concluded*)

Language	Character Modifiers	Special Characters
Sindhi	ā ī ū ṭ ḥ ḍ ṛ ṣ ṣ ṛ ḥ ḍ ṣ ṭ ṣ ṣ ḥ ḥ ḡ ḡ ṇ ṇ	‘
Sinhalese	ā ī ū ē ō ṛ ṛ ṛ ṛ ṛ ṛ ṛ ṛ ṭ ḍ ṇ ṣ ṣ ṣ ṣ ṣ ṣ	’
Tamil	ā ī ū ē ō ḥ ḥ ḥ ḥ ḥ ḥ ḥ ḥ ṇ ṇ ṣ	
Telugu	ā ī ū ē ō ṛ ṛ ṛ ṛ ṛ ṛ ṛ ṛ ṣ ṣ ṣ ṣ ṣ ṣ ṣ ṣ ṣ ṣ ṣ ṣ	
Thai	ā ī ū ṛ ē æ ō ṛ œ ɔ ɕ	ʊ æ œ ‘
Tibetan	ā ī ū ṛ ṛ ṛ ṛ ṛ ṛ ṛ ṛ ḥ ṇ ṇ ṣ ṣ ṣ	’ /
Ukrainian	ī ī ī ē ū ī ā ts zh	/
Urdu	ā ī ū ṭ ḥ ḍ ṛ ṣ ṣ ṣ ḥ ṣ ḡ ṇ ṣ ṣ ṣ	’ ‘
Yiddish	ḥ ḥ ṭ ṣ ṣ ṣ	/

## Appendix B

### ANSEL Character modifiers and special characters with languages of occurrence

(This appendix is not part of American National Standard  
ANSI/NISO Z39.47-1993 and is included for information only.)

In the following tables, ANSEL character modifiers and special characters are listed with the languages in which each occurs. Double character modifiers that are sometimes used in Vietnamese and other languages are not given a separate entry.

The transliterations take into consideration the tables included in Appendix A.

**Table B1. Character modifiers**

Name	Examples	Languages
acute accent	á ċ é ĝ í ĳ ĭ Ĺ ŋ ó ő ř š ů ŵ ŷ ž	Albanian, Catalan, Croatian, Czech, Dutch Faroese, French, Hawaiian, Hungarian, Icelandic (Modern), Navaho, Polish, Portuguese, Slovak, Slovene, Spanish, Tagalog, Vietnamese, Wendic  Transliterated: Church Slavic, Gujarati, Hebrew, Hindi, Kannada, Khmer, Konkani, Maithili, Malayalam, Manipuri, Marathi, Mewari, Nepali, Oriya, Pahari, Persian, Prakrit, Pushto, Rajasthani, Sanskrit, Serbian, Sindhi, Sinhalese, Tamil, Telugu, Tibetan, Urdu, Yiddish
breve	ă ẽ ģ ĭ ǻ ǫ ŭ	Esperanto, Navaho, Romanian, Turkish (Modern), Vietnamese  Transliterated: Belorussian, Braj, Bulgarian, Chinese, Church, Slavic, Dogri, Georgian, Hindi, Khmer, Korean, Lahnda, Maithili, Mewari, Nepali, Pahari, Panjabi, Rajasthani, Russian, Sindhi, Ukrainian (Continued)

Table B1. Character modifiers (Continued)

Name	Examples	Languages
candrabindu	ṁ ṇ ṡ	Transliterated: Assamese, Bengali, Braj, Bulgarian, Hindi, Maithili, Manipuri, Mewari, Nepali, Oriya, Pahari, Prakrit, Rajasthani, Sanskrit, Sindhi, Telugu, Tibetan
cedilla	ç ş	Albanian, Catalan, French, Portuguese, Turkish (Modern)
circle above (angstrom)	Å å	Czech, Danish, Norwegian, Slovak, Swedish
circle below	ᳵ ᳶ	Transliterated: Assamese, Bengali, Braj, Gujarati, Hindi, Kannada, Khmer, Konkani, Maithili, Malayalam, Manipuri, Marathi, Mewari, Nepali, Oriya, Pahari, Prakrit, Rajasthani, Sanskrit, Sinhalese, Telugu, Tibetan
circumflex accent	â ê ê ĝ ĥ ĵ ô û û ŵ ŷ	Afrikaans, Albanian, Dutch, Esperanto, French, Navaho, Portuguese, Romanian, Slovak, Slovene, Tagalog, Turkish (Modern), Vietnamese  Transliterated: Amharic, Braj, Chinese, Gujarati, Hindi, Khmer, Maithili, Marathi, Mewari, Nepali, Pahari, Rajasthani, Sindhi, Sinhalese, Telugu
dot above	ḃ ḥ ḥ ḡ ḥ ḥ ṁ ṇ ṡ ṡ ṡ ṡ Ẃ Ẃ	Lithuanian, Navaho, Turkish (Modern)  Transliterated: Amharic, Assamese, Belorussian, Bengali, Braj, Church Slavic, Dogri, Georgian, Gujarati, Hindi, Kannada, Khmer, Konkani, Lahnda, Maithili, Malayalam, Manipuri, Marathi, Mewari, Nepali, Oriya, Pahari, Pali, Panjabi, Prakrit, Pushto, Rajasthani, Russian, Sanskrit, Sindhi, Sinhalese, Tamil, Telugu, Tibetan

(Continued)

Table B1. Character modifiers (Continued)

Name	Examples	Languages
dot below	ạ ɗ ẹ ɣ ɦ ɓ ɓ ɪ ɓ ɪ ɓ ɪ ɹ ɹ ɹ ɹ ɹ ɹ ɹ	Vietnamese  Transliterated: Amharic, Arabic, Assamese, Bengali, Braj, Dogri, Georgian, Gujarati, Hebrew, Hindi, Kannada, Khmer, Konkani, Maithili, Malayalam, Manipuri, Marathi, Mewari, Nepali, Oriya, Pahari, Panjabi, Persian, Prakrit, Pushto, Rajasthani, Sanskrit, Sindhi, Sinhalese, Tamil, Telugu, Tibetan, Urdu, Yiddish
double acute accent	ǎ ő ű	Hungarian
double dot below	ḥ ḍ ḥ ḷ ṭ ṣ	Transliterated: Braj, Hindi, Kannada, Konkani, Maithili, Mewari, Nepali, Pahari, Persian, Pushto, Rajasthani, Sindhi, Urdu
double tilde	ñ	Tagalog
double underscore	ḡ ḥ	Transliterated: Braj, Hindi, Maithili, Mewari, Nepali, Pahari, Rajasthani
grave accent	à è ì ñ ò ù ù	Afrikaans, Catalan, French, Italian, Portuguese, Tagalog, Vietnamese, Wendic  Transliterated: Khmer, Yiddish
háček	ǎ č ď ě ě ě ǎ ě ř š ť ů ž	Czech, Latvian, Lithuanian, Croatian, Slovak, Slovene, Wendic  Transliterated: Amharic, Arabic, Armenian, Church Slavic, Georgian, Macedonian, Serbian, Sinhalese, Thai (Continued)

Table B1. Character modifiers (*Continued*)

Name	Examples	Languages
half circle below (upadhmaniya)	ḥ	Transliterated: Prakrit, Sanskrit
high comma, centered	ḡ	Latvian
high comma, off center	d' g' k' l' t'	Croatian, Czech, Navaho, Slovak, Slovene, Wendic
left hook	ķ ļ ņ š ţ	Latvian, Romanian
ligature	īā īē īō īū t̄s z̄h ōt k̄s p̄s īē īē	Transliterated: Belorussian, Bulgarian, Church Slavic, Russian, Ukrainian
low rising tone mark	ǎ ǣ ǔ	Vietnamese
macron	ā æ ē ī k̄ ī n̄ ō œ p̄ r̄ s̄ ū	Anglo-Saxon, Hawaiian, Latvian, Lithuanian  Transliterated: Amharic, Arabic, Armenian, Assamese, Bengali, Braj, Burmese, Church Slavic, Dogri, Georgian, Greek, Gujarati, Hindi, Japanese, Kannada, Khmer, Konkani, Korean, Lao, Lahnda, Maithili, Malayalam, Manipuri, Marathi, Mewari, Nepali, Oriya, Pahari, Panjabi, Persian, Prakrit, Pushto, Rajasthani, Russian, Sanskrit, Sindhi, Sinhalese, Tamil, Telugu, Thai, Tibetan, Urdu
right cedilla	ç	Transliterated: Thai
right hook	æ e i q s t u	Anglo-Saxon, Lithuanian, Navaho, Polish, Vietnamese  Transliterated: Church Slavic, Lao

*(Continued)*



Table B1. Character modifiers (*Concluded*)

Name	Examples	Languages
tilde	ã ẽ ĩ ñ õ õ̃ š ů	Estonian, Navaho, Portuguese, Spanish, Vietnamese  Transliterated: Amharic, Assamese, Bengali, Braj, Dogri, Gujarati, Hindi, Kannada, Khmer, Konkani, Lahnda, Lithuanian, Maithili, Malayalam, Manipuri, Marathi, Mewari, Nepali, Oriya, Pahari, Panjabi, Prakrit, Rajasthani, Sanskrit, Sindhi, Sinhalese, Tamil, Telugu, Tibetan
umlaut (diaeresis)	ä ë ħ ĩ j ö ü v ŷ	Afrikaans, Albanian, Catalan, Dutch, Estonian, French, German, Hungarian, Icelandic, Norwegian, Portuguese, Slovak, Spanish, Swedish, Turkish  Transliterated: Chinese, Russian, Sindhi, Ukrainian
underscore	k l n r s t z g h d	Transliterated: Assamese, Bengali, Braj, Dogri, Greek, Hindi, Kannada, Konkani, Lahnda, Maithili, Malayalam, Manipuri, Mewari, Nepali, Pahari, Panjabi, Persian, Prakrit, Pushto, Rajasthani, Sanskrit, Sindhi, Tamil, Telugu, Urdu

Table B2. Special characters

Name	Examples	Languages
alif	’	Indonesian, Turkish  Transliterated: Amharic, Arabic, Assamese, Bengali, Braj, Burmese, Dogri, Georgian, Gujarati, Hebrew, Hindi, Japanese, Khmer, Korean, Lao, Maithili, Malayalam, Manipuri, Marathi, Mewari, Nepali, Oriya, Pahari, Persian, Prakrit, Rajasthani, Sanskrit, Sinhalese, Tamil, Telugu, Tibetan, Urdu
‘ayn	‘	Navaho  Transliterated: Amharic, Arabic, Armenian, Burmese, Chinese, Georgian, Hebrew, Khmer, Korean, Lao, Persian, Pushto, Sindhi, Thai, Urdu
degree sign	°	Transliterated: Khmer
dotless i	ı (I)	Turkish (Modern)
eth	ð (Ð)	Anglo-Saxon, Faroese, Icelandic (Modern)
hard sign (tvěrdyĭ znak)	//	Transliterated: Burmese, Church Slavic, Russian, Ukrainian
hook o	ơ (Ơ)	Vietnamese
hook u	ư (Ư)	Vietnamese  Transliterated: Thai

(Continued)

Table B2. Special characters (*Concluded*)

Name	Examples	Languages
ligature ae	æ (Æ)	Anglo-Saxon, Danish, Faroese, Icelandic (Modern), Navaho, Norwegian  Transliterated: Lao, Thai
ligature oe	œ (Œ)	Anglo-Saxon, French  Transliterated: Lao, Thai
middle dot	·	Catalan  Transliterated: Khmer
slash d	đ (Đ)	Croatian, Vietnamese  Transliterated: Macedonian, Serbian
slash l	ł (Ł)	Navaho, Polish, Wendic
slash o	ø (Ø)	Danish, Faroese, Norwegian
soft sign (м'ягкий знак)	'	Transliterated: Arabic, Belorussian, Church, Slavic, Hebrew, Khmer, Persian, Pushto, Russian, Tibetan, Ukrainian, Yiddish
thorn	þ (Þ)	Anglo-Saxon, Icelandic (Modern)

**ANSI/NISO Z39.47-1993 (R2003)**  
**ERRATA SHEET**

1. ANSI/NISO Z39.47-1993 is Registration # 231 in the ISO International Register of Coded Character Sets to be Used with Escape Sequences. It is available at this url:  
<http://www.itscj.ipsj.or.jp/ISO-IR/231.pdf>
2. Page 7, Table 6 the character encoded as "7/7" (7-bit) and 15/7 (8-bit) is named "left hook" (not "left hoof").
3. Page 11, Table A1 for the Vietnamese language delete the character modifier a with a macron.