

ISO basic Latin alphabet

The **ISO basic Latin alphabet** is a Latin-script alphabet and consists of two sets of 26 letters, codified in^[1] various national and international standards and used widely in international communication.

The two sets contain the following 26 letters each:^{[1][2]}

ISO basic Latin alphabet

Uppercase Latin alphabet	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>H</u>	<u>I</u>	<u>J</u>	<u>K</u>	<u>L</u>	<u>M</u>	<u>N</u>	<u>O</u>	<u>P</u>	<u>Q</u>	<u>R</u>	<u>S</u>	<u>T</u>	<u>U</u>	<u>V</u>	<u>W</u>	<u>X</u>	<u>Y</u>	<u>Z</u>
Lowercase Latin alphabet	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z

Contents

History

Terminology

- Name for Unicode block that contains all letters
- Names for the two subsets
- Names for the letters

Timeline for encoding standards

- Timeline for widely used computer codes supporting the alphabet

Representation

Usage

- Alphabets containing the same set of letters
- Column numbering

See also

References

History

By the 1960s it became apparent to the computer and telecommunications industries in the First World that a non-proprietary method of encoding characters was needed. The International Organization for Standardization (ISO) encapsulated the Latin script in their (ISO/IEC 646) 7-bit character-encoding standard. To achieve widespread acceptance, this encapsulation was based on popular usage. The standard was based on the already published *American Standard Code for Information Interchange*, better known as ASCII, which included in the character set the 26 × 2 letters of the English alphabet. Later standards issued by the ISO, for example ISO/IEC 8859 (8-bit character encoding) and ISO/IEC 10646 (Unicode Latin), have continued to define the 26 × 2 letters of the English alphabet as the basic Latin script with extensions to handle other letter in other languages.^[1]

Terminology

Name for Unicode block that contains all letters

The Unicode block that contains the alphabet is called "C0 Controls and Basic Latin".

Names for the two subsets

In Unicode 7.0 two subheadings exist:^[3]

- "Uppercase Latin alphabet", individual letters contain the string `UPPERCASE LATIN CAPITAL LETTER` in their descriptions
- "Lowercase Latin alphabet", individual letters contain the string `UPPERCASE LATIN SMALL LETTER` in their descriptions

Names for the letters

The letters are also contained in "Halfwidth and Fullwidth Forms" FF00 to FFE4^[4]

```
FF21 A FULLWIDTH LATIN CAPITAL LETTER A
FF41 a FULLWIDTH LATIN SMALL LETTER A
```

Timeline for encoding standards

- 1865 International Morse Code was standardized at the International Telegraphy Congress in Paris, and was later made the standard by the International Telecommunication Union (ITU)
- 1950s Radiotelephony Spelling Alphabet by ICAO ^[1]

Timeline for widely used computer codes supporting the alphabet

- 1963: ASCII (7-bit character-encoding standard from the American Standards Association which became ANSI in 1969)
- 1963/1964: EBCDIC (developed by IBM and supporting the same alphabetic characters as ASCII, but with different code values)
- 1965-04-30: Ratified by ECMA as **ECMA-6**^[5] based on work the ECMA's Technical Committee TC1 had carried out since December 1960^[5]
- 1972: ISO 646 (ISO 7-bit character-encoding standard, using the same alphabetic code values as ASCII, revised in second edition ISO 646:1983 and third edition ISO/IEC 646:1991 as a joint ISO/IEC standard)
- 1983: ITU-T Rec. T51 | ISO/IEC 6937 (a multi-byte extension of ASCII)
- 1987: ISO/IEC 8859-1:1987 (8-bit character encoding)
 - Subsequently, other versions and parts of ISO/IEC 8859 have been published.
- Mid-to-late 1980s: Windows-1250, Windows-1252, and other encodings used in Microsoft Windows (some roughly similar to ISO/IEC 8859-1)
- 1990: Unicode 1.0 (developed by the Unicode Consortium),^{[6][7]} contained in the block 'C0 Controls and Basic Latin' using the same alphabetic code values as ASCII and ISO/IEC 646
 - Subsequently, other versions of Unicode have been published and it later became a joint ISO/IEC standard as well, as identified below
- 1993: ISO/IEC 10646-1:1993, ISO/IEC standard for characters in Unicode 1.1
 - Subsequently, other versions of ISO/IEC 10646-1 and one of ISO/IEC 10646-2 have been published. Since 2003, the standards have been published under the name "ISO/IEC 10646" without the separation into two parts.
- 1997: Windows Glyph List 4

Representation

In ASCII the letters belong to the printable characters and in Unicode since version 1.0 they belong to the block "C0 Controls and Basic Latin". In both cases, as well as in ISO/IEC 646, ISO/IEC 8859 and ISO/IEC 10646 they are occupying the positions in hexadecimal notation 41 to 5A for uppercase and 61 to 7A for lowercase.

Not case sensitive, all letters have code words in the CAO spelling alphabet and can be represented with Morse code.

Usage

All of the lowercase letters are used in the International Phonetic Alphabet (IPA). In X-SAMPA and SAMPA these letters have the same sound value as in IPA. In Kirshenbaum they have the same value except for the letter r



The letters of the ISO basic Latin alphabet on a 16-segment display (plus the Arabic numerals).

Alphabets containing the same set of letters

The list below only includes alphabets that lack:

- letters whose diacritical marks make them distinct letters.

- multigraphs that constitute distinct letters.

alphabet	diacritic	multigraphs (not constituting distinct letters)	ligatures
<u>Afrikaans alphabet</u>	á, é, è, ê, ë, í, î, ï, ó, ô, ú, û, ý		
<u>Catalan alphabet</u>	à, é, è, í, ï, ò, ò, ú, ü, ç		
<u>Dutch alphabet</u>	ä, é, è, ë, ï, ō, ū	The digraph (ij) is sometimes considered to be a separate letter. When that is the case, it usually replaces or is intermixed with (y).	
<u>English alphabet</u>	-none-	<sh>, <ch>, <ea>, <ou>, <th>, <ph>, <ng>	æ, œ
<u>French alphabet</u>	à, â, ç, é, è, ê, ë, î, ï, ô, û, û, ü, ý	<ai>, <au>, <ei>, <eu>, <oi>, <ou>, <eau>, <ch>, <ph>, <gn>, <an>, <am>, <en>, , <in>, <im>, <on>, <om>, <un>, <um>, <yn>, <ym>, <ain>, <aim>, <ein>, <oin>, <aï>, <eï>	æ, œ
<u>German alphabet</u>	ä, ö, ü	<sch>, <qu>, <ch>, <ph>, <ng>, <ie>, <ck>, <ei>, <eu>, <äu>	ß
<u>Italian alphabet</u>	à, è, é, ì, ò, ù	<ch>, <ci>, <gh>, <gi>, <gl>, <gli>, <gn>, <sc>, <sc>	
<u>Ido alphabet</u>	-none-	<qu>, <ch>, <sh>	-none-
<u>Indonesian alphabet</u>	-none-	<kh>, <ng>, <ny>, <sy>	
<u>Interlingua alphabet</u>	-none-	<qu>	-none-
<u>Luxembourgish alphabet</u>	ä, é, ë		
<u>Malay alphabet</u>	-none-	<gh>, <kh>, <ng>, <ny>, <sy>	-none-
<u>Portuguese alphabet</u>	ã, õ, á, é, í, ó, ú, â, ê, ô, ä, ç	<ch>, <lh>, <nh>, <rr>, <ss>, <am>, , <im>, <om>, <um>, <ãe>, <ão>, <õe>	-none-

English is the only major modern European language requiring no diacritics for native words (although a diaeresis is used by some publishers in words such as "coöperation"):^{[8][9]}

Note for Portuguese: *k*, *w* and *y* were part of the alphabet until several spelling reforms during the 20th century, the aim of which was to change the etymological Portuguese spelling into an easier phonetic spelling. These letters were replaced by other letters having the same sound: thus *psychologia* became *psicologia*, *kioske* became *quiosque*, *martyr* became *mártir*, etc. Nowadays *k*, *w*, and *y* are only found in foreign words and their derived terms and in scientific abbreviations (e.g. *gkm*, *byronismo*). These letters are considered part of the alphabet again following the 1990 Portuguese Language Orthographic Agreement which came into effect on January 1, 2009, in Brazil. See Reforms of Portuguese orthography

Column numbering

The Roman (Latin) alphabet is commonly used for column numbering in a table or chart. This avoids confusion with row numbers using Arabic numerals. For example, a 3-by-3 table would contain Columns A, B, and C, set against Rows 1, 2, and 3. If more columns are needed beyond Z (normally the final letter of the alphabet), the column immediately after Z is AA, followed by AB, and so on. This can be seen by scrolling far to the right in a spreadsheet program such as Microsoft Excel or LibreOffice Calc.

These are double-digit "letters" for table columns, in the same way that 10 through 99 are double-digit numbers. The Greek alphabet has a similar extended form that uses such double-digit letters if necessary but it is used for chapters of afaternity as opposed to columns of a table.

Such double-digit letters for bullet points are AA, BB, CC, etc., as opposed to the number-like place value system explained above for table columns.

See also

- Hebrew alphabet
- Greek alphabet
- Latin alphabet, Latin-script alphabets
- Early Cyrillic alphabet, Cyrillic alphabets
- Windows code pages

References

1. "Internationalisation standardization of 7-bit codes, ISO 646"(<http://www.terena.org/activities/multiling/euoml/section04.html>) Trans-European Research and Education Networking Association (TERENA) Retrieved 2010-10-03.
2. "RFC1815 – Character Sets ISO-10646 and ISO-10646-J-1"(<http://www.faqs.org/rfcs/rfc1815.html>) Retrieved 2010-10-03.
3. "CO Controls and Basic Latin"(<https://www.unicode.org/charts/PDF/U0000.pdf>) (PDF). *Unicode.org*. Retrieved 2016-08-08.
4. "Halfwidth and Fullwidth Forms"(<https://www.unicode.org/charts/PDF/UFF00.pdf>) (PDF). *Unicode.org*. Retrieved 2016-08-08.
5. *Standard ECMA-6: 7-Bit Coded Character Set* (<http://www.ecma-international.org/publications/files/ECMA-STWITHDRAWN/ECMA-6,%205th%20Edition,%20March%201985.pdf>) (PDF) (5th ed.). Geneva, Switzerland:European Computer Manufacturers Association (Ecma). March 1985. Archived (<https://web.archive.org/web/20160529230908/http://www.ecma-international.org/publications/files/ECMA-STWITHDRAWN/ECMA-6%2C%205th%20Edition%2C%20March%201985.pdf>) PDF) from the original on May 29, 2016 Retrieved 2016-05-29. "The Technical Committee TC1 of ECMA met for the first time in December 1960 to prepare standard codes for Input/Output purposes. On April 30, 1965, Standard ECMA-6 was adopted by the General Assembly of ECMA."
6. "Unicode character database"(<https://www.unicode.org/ucd>) *The Unicode Standard* Retrieved 2013-03-22.
7. *The Unicode Standard Version 1.0, Volume 1*. Addison-Wesley Publishing Company Inc. 1990. ISBN 0-201-56788-1
8. As an example, an article containing **diaeresis** in "coöperate" and **cedilla** in "façades" as well as **acircumflex** in the word "crêpe" (Grafton, Anthony (2006-10-23). "Books: The Nutty Professors, The history of academic charisma"(http://www.newyorker.com/archive/2006/10/23/061023crbo_book?currentPage=all) *The New Yorker*.)
9. "The New Yorker's odd mark — the diaeresis" (<https://web.archive.org/web/20101216160024/http://dsciber.com/news/121-the-new-yorkers-odd-mark-the-diaeresis>) 16 December 2010.

Retrieved from 'https://en.wikipedia.org/w/index.php?title=ISO_basic_Latin_alphabet&oldid=865862139

This page was last edited on 26 October 2018, at 17:31(UTC).

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. By using this site, you agree to the [Terms of Use](#) and [Privacy Policy](#). Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc](#), a non-profit organization.