

# References

Citations are given for the standards and dictionaries that were used as the actual resources for *The Unicode Standard*, primarily for Version 1.0. Where a Draft International Standard is known to have progressed to International Standard status, the entry for the DIS has been revised. *A revised or reaffirmed edition, with a different date, may have been published subsequently.* For the current version of a standard, see the catalog issued by the standards organization. The Web site of the International Organization for Standardization (<http://www.iso.ch/>) includes the *ISO Catalogue* and links to the sites of member organizations. Many of the ISO character set standards were originally developed by ECMA and are also ECMA standards.

In general, American library practice has been followed for the romanization of titles and names written in non-Roman script. Exceptions are when information supplied by an author had to be used because the name or title in the original script was unavailable.

---

## R.1 Source Standards

**ANSI X3.4:** American National Standards Institute. *Coded character set—7-bit American national standard code for information interchange*. New York, 1986. (ANSI X3.4-1986).

**ANSI X3.32:** American National Standards Institute. *American national standard graphic representation of the control characters of American national standard code for information interchange*. New York, 1973. (ANSI X3.32-1973).

**ANSI Y10.20:** American National Standards Institute. *Mathematic signs and symbols for use in physical sciences and technology*. New York, 1988. (ANSI Y10.20-1975 (R1988)).

**ANSI Z39.47:** American National Standards Institute. *Extended Latin alphabet coded character set for bibliographic use*. New York, 1985. (ANSI Z39.47-1985).

**ANSI Z39.64:** American National Standards Institute. *East Asian character code for bibliographic use*. New Brunswick, Transaction, 1991. (ANSI Z39.64-1989).

**ASMO 449:** Arab Organization for Standardization and Metrology. *Data processing 7-bit coded character set for information interchange*. [s.l.], 1983. (Arab standard specifications; 449-1982). Authorized English translation.

**CCCI:** *Zhongwen Zixun Jiaohuanma (Chinese Character Code for Information Interchange)*. Revised edition. Taipei, Xingzhengyuan Wenhua Jianshe Xiaozu (Executive Yuan Committee for Cultural Construction), 1985.

**CNS 11643-1986:** *Tongyong hanzi biao zhun jiaohuanma (Han character standard interchange code for general use)*. Taipei, Xingzhengyuan (Executive Yuan), 1986.

**CNS 11643-1992:** *Zhongwen biao zhun jiaohuanma (Chinese standard interchange code)*. Taipei, 1992.

Obsoletes 1986 edition.

**EACC:** (See ANSI Z39.64.)

**ECMA Registry:** (See ISO Register.)

**GB 2312:** *Xinxi jiaohuanyong hanzi bianmaji, jibenji (Code of Chinese graphic character set for information interchange, primary set)*. Beijing, Jishu Biaozhun Chubanshe (Technical Standards Press), 1981. (GB 2312-1980).

**GB 12345:** *Xinxi jiaohuanyong hanzi bianmaji, fuzhuji (Code of Chinese ideogram set for information interchange, supplementary set)*. Beijing, Jishu Biaozhun Chubanshe (Technical Standards Press), 1990. (GB 12345-1990).

**GB 13000:** *Xinxi jishu—Tongyong duobawei bianma zifuji (UCS)—Diyi bufen: Tixi jiegou yu jiben duowenzhong pingmian (Information technology—Universal multiple-octet coded character set (UCS)—Part 1: Architecture and basic multilingual plane)*. Beijing, Jishu Biaozhun Chubanshe (Technical Standards Press), 1993. (GB 13000.1-93) (ISO/IEC 10646.1-1993).

**GB 13134:** *Xinxi jiaohuanyong yiwen bianma zifuji (Yi coded character set for information interchange)*, [prepared by] Sichuansheng Minzushiwu Weiyuanhui. Beijing, Jishu Biaozhun Chubanshe (Technical Standards Press), 1991. (GB 13134-1991).

**GBK:** *Xinxi jiaohuanyong hanzi bianma kuozhan guifan (Extended Code of Chinese graphic character set for information interchange)*. Beijing, Zhongguo dianzi gongyebu [and] Guojiao jishu jianduju, 1995.

The Chinese-specific subset of GB 13000.1-93

GBK is implemented on Windows platforms as Code Page 936.

**ISCII:** India Department of Electronics. *Indian standard code for information interchange*. New Delhi, 1988.

**ISO Register:** International Organization for Standardization. *ISO international register of character sets to be used with escape sequences*. [Geneva], 1990.

The current register is available at <http://www.itscj.ipsj.or.jp/ISO-IR/>.

**ISO/IEC 646:** International Organization for Standardization. *Information technology—ISO 7-bit coded character set for information interchange*. [Geneva], 1991. (ISO/IEC 646:1991).

**ISO 2047:** International Organization for Standardization. *Information processing—Graphical representations for the control characters of the 7-bit coded character set*. [Geneva], 1975. (ISO 2047:1975).

**ISO/IEC 2022:** International Organization for Standardization. *Information processing—ISO 7-bit and 8-bit coded character sets—Code extension techniques*. 3d ed. [Geneva], 1986. (ISO 2022:1994).

Edition 4 (ISO/IEC 2022:1994) has title: *Information technology—Character code structure and extension techniques*.

**ISO 2033:** International Organization for Standardization. *Information processing—Coding of machine-readable characters (MICR and OCR)*. 2d ed. [Geneva], 1983. (ISO 2033:1983).

**ISO/IEC 4873:** International Organization for Standardization. *Information technology—ISO 8-bit code for information interchange—Structure and rules for implementation*. [Geneva], 1991. (ISO/IEC 4873:1991).

**ISO 5426:** International Organization for Standardization. *Extension of the Latin alphabet coded character set for bibliographic information interchange*. 2d ed. [Geneva], 1983. (ISO 5426:1983).

**ISO 5426-2:** International Organization for Standardization. *Information and documentation—Extension of the Latin alphabet coded character set for bibliographic information*

*interchange—Part 2: Latin characters used in minor European languages and obsolete typography*. [Geneva], 1996. (ISO 5426-2:1986).

**ISO 5427:** International Organization for Standardization. *Extension of the Cyrillic alphabet coded character set for bibliographic information interchange*. [Geneva], 1984. (ISO 5427:1984).

**ISO 5428:** International Organization for Standardization. *Greek alphabet coded character set for bibliographic information interchange*. [Geneva], 1984. (ISO 5428-1984).

**ISO/IEC 6429:** International Organization for Standardization. *Information technology—Control functions for coded character sets*. 3d. ed. [Geneva], 1992. (ISO/IEC 6429:1992).

**ISO 6438:** International Organization for Standardization. *Documentation—African coded character set for bibliographic information interchange*. [Geneva], 1983. (ISO 6438:1983).

**ISO 6862:** International Organization for Standardization. *Information and documentation—Mathematics character set for bibliographic information interchange*. [Geneva], 1996. (ISO 6862:1996).

**ISO/IEC 6937:** International Organization for Standardization. *Information processing—Coded character sets for text communication*. [Geneva], 1984.

Edition 2 (ISO/IEC 6937:1994) has the following title: *Information technology—Coded graphic set for text communication—Latin alphabet*.

**ISO/IEC 8859:** International Organization for Standardization. *Information processing—8-bit single-byte coded graphic character sets*. [Geneva], 1987–.

These parts of ISO/IEC 8859 predate the Unicode Standard and were used as resources: Part 1, *Latin alphabet no. 1*; Part 2, *Latin alphabet no. 2*; Part 3, *Latin alphabet no. 3*; Part 4, *Latin alphabet no. 4*; Part 5, *Latin/Cyrillic alphabet*; Part 6, *Latin/Arabic alphabet*; Part 7, *Latin/Greek alphabet*; Part 8, *Latin/Hebrew alphabet*; Part 9, *Latin alphabet no. 5*.

**ISO 8879:** International Organization for Standardization. *Information processing—Text and office systems—Standard generalized markup language (SGML)*. [Geneva], 1986. (ISO 8879:1986).

**ISO 8957:** International Organization for Standardization. *Information and documentation—Hebrew alphabet coded character sets for bibliographic information interchange*. [Geneva], 1996. (ISO 8957:1996).

**ISO 9036:** International Organization for Standardization. *Information processing—Arabic 7-bit coded character set for information interchange*. [Geneva], 1987. (ISO 9036:1987).

**ISO/IEC 10367:** International Organization for Standardization. *Information technology—Standardized coded graphic character sets for use in 8-bit codes*. [Geneva], 1991. (ISO/IEC 10367:1991).

**ISO 10585:** International Organization for Standardization. *Information and documentation—Armenian alphabet coded character set for bibliographic information interchange*. [Geneva], 1996. (ISO 10585:1996).

**ISO 10586:** International Organization for Standardization. *Information and documentation—Georgian alphabet coded character set for bibliographic information interchange*. [Geneva], 1996. (ISO 10586:1996).

**ISO/IEC 10646:** International Organization for Standardization. *Information Technology—Universal Multiple-Octet Coded Character Set (UCS)—Part 1: Architecture and Basic Multilingual Plane*. [Geneva], 1993. (ISO/IEC 10646-1:1993).

ISO 10754: International Organization for Standardization. *Information and documentation—Extension of the Cyrillic alphabet coded character set for non-Slavic languages for bibliographic information interchange*. [Geneva], 1996 (ISO 10754:1996).

ISO DIS 11548-1: International Organization for Standardization. *Communication aids for blind persons—Identifiers, names and assignation to coded character sets for 8-dot Braille characters—Part 1: General guidelines for Braille identifiers and shift marks*. [Geneva], 1997. (ISO DIS 11548-1).

ISO DIS 11548-2: International Organization for Standardization. *Communication aids for blind persons—Identifiers, names and assignation to coded character sets for 8-dot Braille characters—Part 2: Latin alphabet based character sets*. [Geneva], 1997. (ISO DIS 11548-2).

ISO 11822: International Organization for Standardization. *Information and documentation—Extension of the Arabic coded character set for bibliographic information interchange*. [Geneva], 1996. (ISO 11822:1996).

JIS X 0208: Japanese Standards Association. *Jouhou koukan you kanji fugoukei (Code of the Japanese Graphic Character Set for Information Interchange)*. Tokyo, 1990. (JIS X 0208-1990).

Revision of the 1983 edition.

JIS X 0212: Japanese Standards Association. *Jouhou koukan you kanji fugou-hojo kanji (Code of the supplementary Japanese graphic character set for information interchange)*. Tokyo, 1990. (JIS X 0212-1990).

1990 extensions to JIS.

JIS X 0221: Information Technology—Universal Multiple-Octet Coded Character Set (UCS)—Part 1: Architecture and Basic Multilingual Plane. Tokyo, 1995. (JIS X0221-1995).

KS C 5601-1987: Korea Industrial Standards Association. *Chongbo kyohwanyong puho (Han'gul mit Hancha)*. Seoul, 1989. (KS C 5601-1987).

KS C 5601-1992: Korea Industrial Standards Association. *Chongbo kyohwanyong puho (Han'gul mit Hancha)*. Seoul, 1992. (KS C 5601-1992).

KS C 5657: Korea Industrial Standards Association. *Chongbo kyohwanyong puho hwakchang set u*. Seoul, 1991. (KS C 5657-1991).

SI 1311.1: Standards Institution of Israel. *Information technology: ISO 8-bit coded character set with Hebrew points*. [Tel Aviv, 1996] (SI 1311.1 (1996)).

SI 1311.2: Standards Institution of Israel. *Information technology: ISO 8-bit coded character set with Hebrew accents*. [Tel Aviv, 1996] (SI 1311.2 (1996)).

TIS 620-2529: Thai Industrial Standards Institute, Ministry of Industry. *Thai Industrial Standard for Thai Character Code for Computer*. Bangkok, 1986. (TIS 620-2529-1986).

TIS 620-2533: Thai Industrial Standards Institute, *Standard for Thai Character Codes for Computers*. Bangkok, 1990. (TIS 620-253 -1990). ISBN 974-606-153-4.

In Thai. Online version at <http://www.nectec.or.th/it-standards/std620/std620.htm>.

## R.2 Source Dictionaries for Han Unification

*Dae Jaweon*. Seoul, Samseong Publishing Co. Ltd., 1988.

*Dai Kan-Wa Jiten* / Morohashi Tetsuji cho. Shu teiban. Tokyo, Taishukan Shoten, Showa 59-61 [1984-86].

*Hanyu Da Zidian*. 1st edition. Chendu, Sichuan Cishu Publishing, 1986.

*KangXi Zidian*. 7th edition. Beijing, Zhonghua Bookstore, 1989.

---

## R.3 Other Sources for the Unicode Standard

*ALA-LC romanization tables: transliteration schemes for non-Roman scripts, approved by the Library of Congress and the American Library Association*. Tables compiled and edited by Randall K. Barry. Washington, Library of Congress, 1997. ISBN 0-8444-0940-5.

Bburx Ddie Su (= Bian Xiezhe). *Nuo-su bbur-ma shep jie zzit: Syp-chuo se nuo bbur-ma syt mu curx su niep sha zho ddop ma bbur-ma syt mu wo yuop hop, Bburx Ddie da Su. = Yi wen jian zi ben: Yi Han wen duizhao ban*. Chengdu: Sichuan minzu chubanshe, 1984.

Bburx Ddie Su. *Nip huo bbur-ma ssix jie: Nip huo bbur-ma ssi jie Bburx Ddie curx Su. = Yi Han zidian*. Chengdu: Sichuan minzu chubanshe, 1990. ISBN 7-5409-0128-4

Canadian Aboriginal Syllabic Encoding Committee. *Repertoire of unified Canadian aboriginal syllabics proposed for inclusion into ISO/IEC 10646: International standard universal multiple-octet coded character set*. [Canada]: CASEC [1994].

€uro: one currency for Europe.

<http://europa.eu.int/euro/html/entry.html>

Journal of the International Phonetic Association, 24.2 (1994): 95–98, and 25.1 (1995): 21.

Pullum, Geoffrey K. *Phonetic Symbol Guide* / Geoffrey K. Pullum and William A. Ladusaw. 2nd ed. Chicago, University of Chicago Press, 1996. ISBN 0-226-68535-7; 0-226-68536-5 (pbk.).

Pullum, Geoffrey K. *Remarks on the 1989 revision of the International Phonetic Alphabet*. Journal of the International Phonetic Association, 20:1 (1990), 33–40.

RFC 2152. *UTF-7: A mail-safe transformation format of Unicode*, by D. Goldsmith (Apple Computer, Inc.) and M. Davis (Taligent, Inc.). May 1997.

<http://www.rfc-editor.org> (and search for the RFC) or <ftp://ftp.isi.edu/in-notes/rfc2152.txt>

RFC 2279. *UTF-8: A transformation format of ISO 10646*, by F. Yergeau (Alis Technologies). January 1998.

<http://www.rfc-editor.org> (and search for the RFC) or <ftp://ftp.isi.edu/in-notes/rfc2279.txt>

Samaranayake V. K. *A Standard Code for Information Interchange in Sinhalese*, by V. K. Samaranayake and S. T. Nandasara. 1990. (ISO/IEC JTC1/SC2/WG2 N673).

---

## R.4 Selected Resources

Alexander, J. T. *A Dictionary of the Cherokee Indian Language*. [Sperry, Oklahoma?] Published by the author, 1971.

Allworth, Edward. *Nationalities of the Soviet East: Publications and Writing Systems*. New York, Columbia University Press, 1971. ISBN 0-231-03274-9.

*Alphabete und Schriftzeichen des Morgen- und Abendlandes*. 2. überarb. u. erw. Aufl. Berlin, Bundesdruckerei, 1969.

*The Alphabet Makers: A Presentation from the Museum of the Alphabet.* Waxhaw, North Carolina. 2nd ed. Huntington Beach, California, Summer Institute of Linguistics, 1991. ISBN 0-938978-13-6.

Armbruster, Carl Hubert. *Initia Amharica: An Introduction to Spoken Amharic.* Cambridge, Cambridge University Press, 1908–1920.

Bergsträsser, Gotthelf. *Introduction to the Semitic Languages: Text Specimens and Grammatical Sketches.* Translated ... with an appendix on the scripts by Peter T. Daniels. [2d. ed.] Winona Lake, Eisenbrauns, 1995. ISBN 0-931464-10-2.

Translation of *Einführung in die semitischen Sprachen*, 1928.

*A Bibliography on Writing and Written Language*, edited by Konrad Ehlich, Florian Coulmas, Gabriele Graefen. Berlin, New York, Mouton de Gruyter, 1996. (Trends in linguistics. Studies and monographs, 89.) ISBN 3-11-010158-0.

Contains references to about 27,500 publications covering mainly 1930–1992.

*The Book of a Thousand Tongues*, by Eugene A. Nida. Rev. ed. London, United Bible Society, 1972.

First ed. by Eric M. North, 1938.

Bringhurst, Robert. *The Elements of Typographic Style.* 2d ed. Point Roberts, Washington, Hartley & Marks, 1996. ISBN 0-88179-133-4; 0-88179-132-6 (pbk.).

Campbell, George L. *Compendium of the World's Languages.* London, Routledge, 1990. ISBN 0-415-06937-6 (set); 0-415-06978-5 (v.1); 0-415-06979-3 (v.2).

Clews, John. *Language Automation Worldwide: The Development of Character Set Standards.* Harrogate (North Yorkshire), SESAME Computer Projects, 1988. ISBN 1-87009-501-4.

Comrie, Bernard, ed. *The World's Major Languages.* Oxford, Oxford University Press, 1987. ISBN 0-19-520521-9; 0-19-506511-5 (pbk.).

Comrie, Bernard, ed. *The Languages of the Soviet Union.* Cambridge, Cambridge University Press, 1981. ISBN 0-521-23230-9; 0-521-9877-6 (pbk.).

Coulmas, Florian. *The Blackwell Encyclopedia of Writing Systems.* Cambridge, Massachusetts, Blackwell, 1996. ISBN 0-631-19446-0.

Coulmas, Florian. *The Writing Systems of the World.* Oxford, New York, Blackwell, 1989.

Daniels, Peter T. *The World's Writing Systems.* (See *World's Writing Systems.*)

DeFrancis, John. *Visible Speech: The Diverse Oneness of Writing Systems.* Honolulu, University of Hawaii Press, 1989. ISBN 0-8248-1207-7.

Diringer, David. *The Alphabet: A Key to the History of Mankind.* 3d ed., completely rev. with the assistance of Reinhold Regensburger. New York, Funk and Wagnalls, 1968.

Also published: London, Hutchinson. ISBN 0-906764-0-8.

Diringer, David. *Writing.* London, Thames and Hudson, 1962.

Also published: New York, Praeger.

Dixon, Robert M. W. *The Languages of Australia.* Cambridge, Cambridge University Press, 1980. ISBN 0-521-22329-6.

Drucker, Johanna. *The Alphabetic Labyrinth: The Letters in History and Imagination.* London, Thames & Hudson, 1999. ISBN 0-500-28068-1.

Endo, Shotoku. *Hayawakari Chugoku Kantai-ji.* Tokyo, Kokusho Kankokai, Showa 61 [1986].

- Esling, John. *Computer Coding of the IPA: Supplementary Report*. Journal of the International Phonetic Association, 20:1 (1990), 22–26.
- Everson, Michael. *Michael Everson's Cool Bibliography of Typography and Scripts*.  
<http://www.egt.ie/standards/iso15924/document/scriptbib.html>
- Extensible Markup Language (XML), 1.0*. (W3C Recommendation 10 February 1998). (REC-xml-19980210). Editors: Tim Bray, Jean Paoli, [and] C. M. Sperberg-McQueen.  
<http://www.w3.org/TR/1998/REC-xml-19980210>
- Faulmann, Carl. *Schriftzeichen und Alphabete aller Zeiten und Völker*. Reprint of *Das Buch Der Schrift ...*, 1880. Augsburg, Augustus Verlag, 1990. ISBN 3-8043-0142-8.
- Flanagan, David. *Java in a Nutshell*. 2d. ed. Sebastopol, California, O'Reilly, 1997. ISBN 1-565-92262-X.
- Friedrich, Johannes. *Geschichte der Schrift*. Heidelberg, C. Winter, 1966.
- Friesen, Otto von. *Runorna*. Stockholm, A. Bonnier [1933]. (*Nordisk kultur*, 6).
- Garneau, Denis. *Keys to Sort and Search for Culturally-Expected Results*. [s.l.] IBM, 1990. (IBM document number GG24-3516, June 1, 1990).
- Gaur, Albertine. *A History of Writing*. Rev. ed. New York, Cross River Press, 1992. ISBN 1-558-59358-6.  
 Also published: London, British Library. ISBN 0-7123-0270-0.
- Geiger, Wilhelm. *Maldivian Linguistic Studies*. New Delhi, Asian Educational Services, 1996. ISBN 81-206-1201-9.  
 Originally published: Colombo, H. C. Cottle, Govt. Printer, 1919.
- Gelb, Ignace J. *A Study of Writing*. Rev. ed. Chicago, University of Chicago Press, 1963. ISBN 0-226-28605-3; 0-226-28606-1 (pbk.).
- Giliarevskii, Rudzhero Sergeevich. *Languages Identification Guide*, by Rudzhero S. Gilyarevsky and Vladimir S. Grivnin. Moscow, Nauka, 1970.
- The Gospel in Many Tongues: Specimens of 875 Languages ...* London, British and Foreign Bible Society, 1965.
- Gunasekara, Abraham Mendis. *A Comprehensive Grammar of the Sinhalese Language*. New Delhi, Asian Education Services, 1986.  
 Reprint of 1891 edition.
- Haarmann, Harald. *Universalgeschichte der Schrift*. Frankfurt, Campus Verlag, 1990. ISBN 3-593-34346-0.
- Habein, Yaeko Sato. *The History of the Japanese Written Language*. Tokyo, University of Tokyo Press, 1984. ISBN 0-86008-347-0; 4-13-087047-5.
- Haugen, Einar Ingvald. *The Scandinavian Languages: an Introduction to their History*. London, Faber, 1976. ISBN 0-571-10423-1.  
 Also published: Cambridge, Massachusetts, Harvard University Press, 1976. ISBN 0-674-79002-2.
- Healey, John F. *The Early Alphabet*. Berkeley, University of California Press; [London] British Museum, 1990. (*Reading the Past*, 9). ISBN 0-7309-6.  
 British Museum Publications edition has ISBN 0-7141-8073-4.
- Holmes, Ruth Bradley. *Beginning Cherokee*, by Ruth Bradley Holmes and Betty Sharp Smith. 2d ed. Norman, University of Oklahoma Press, 1977. ISBN 0-8061-1464-9.

Huang, K'o-tung. *An Introduction to Chinese, Japanese and Korean Computing*, by Jack K. T. Huang and Timothy D. Huang. Singapore, Teaneck, New Jersey, World Scientific, 1989. ISBN 9971-50-664-5.

IBM. *National Language Support Reference Manual*. 4th ed. North York, Ontario, IBM Canada Ltd., National Language Technical Center, 1994. (National Language Design Guide, 2) (SE09-8002-03).

Ifrah, Georges. *From One to Zero, a Universal History of Numbers*. New York, Penguin, 1987. ISBN 0-14-009919-0.

Translation of *Histoire universelle des chiffres*. Also published: New York, Viking, 1985. ISBN 0-670-37395-8.

International Organization for Standardization. *Information Technology—An Operational Model for Characters and Glyphs*. Geneva, 1998. (ISO/IEC TR 15285:1998).

Isaev, Magomet Izmailovich. *Sto tridtsat' ravnopravnykh; o iazykakh narodov SSSR*. Moskva, Nauka, 1970.

Jensen, Hans. *Sign, Symbol and Script: An Account of Man's Efforts to Write*. New York, Putnam, 1969.

Translation of *Die Schrift in Vergangenheit und Gegenwart* (Berlin, Deutscher Verlag, 1969). Also published: London, Allen & Unwin. ISBN 0-04-400021-9.

Katzner, Kenneth. *The Languages of the World*. New ed. London, Routledge, 1995. ISBN 0-415-11809-3.

Kefarnissy, Paul. *Grammaire de la langue Araméenne syriaque*. Beyrouth, 1962.

Knuth, Donald E. *The T<sub>E</sub>Xbook*. 21st. printing, rev. Reading, Massachusetts, Addison-Wesley, 1994. ISBN 0-201-134489.

Ksar, Michael Y. "Untying Tongues," ISO Bulletin (June 1993), 2–8.

Launhardt, Johannes. *Guide to Learning the Oromo (Galla) Language*. Addis Ababa, Launhardt [1973?].

Leslau, Wolf. *Amharic Textbook*. Weisbaden, Harrassowitz; Berkeley, University of California Press, 1968.

Lunde, Ken. *CJKV Information Processing*. Beijing [etc.], O'Reilly, 1999. ISBN 1-565-92224-7.

Malherbe, Michel. *Les Langages de l'Humanité: Une Encyclopédie des 3000 Langues Parlées dans le Monde*. Paris, Laffont, 1995. ISBN 2-221-05947-6.

Maniku, Hassan Ahmed. *Say It in Maldivian (Dhivehi)*, [by] H. A. Maniku [and] J. B. Disanayaka. Colombo, Lake House Investments, 1990.

McManus, Damian. *A Guide to Ogam*. Maynooth, An Sagart, 1991. (*Maynooth monographs*, 4) ISBN 1-87068-417-6.

Muller, Siegfried H. *The World's Living Languages: Basic Facts of Their Structure, Kinship, Location, and Number of Speakers*. New York, Ungar, 1964.

Musaev, Kenesbai Musaevich. *Alfavit' iazykov narodov SSSR*. Moskva, Nauka, 1965.

Musset, Lucien. *Introduction à la Runologie*. Paris, Aubier-Montaigne, 1965.

Naik, Bapurao S. *Typography of Devanagari*. 1st ed., rev. Bombay, Directorate of Languages, Govt. of Maharashtra, 1971.



Nakanishi, Akira. *Writing Systems of the World: Alphabets, Syllabaries, Pictograms*. Rutland, Vermont, Tuttle, 1980. ISBN 0-8048-1293-4; 0-8048-1654-9 (pbk.).

Revised translation of *Sekai no moji*.

*New Echota Letters: Contributions of Samuel A. Worcester to the Cherokee Phoenix*, edited by Jack Frederick Kilpatrick and Anna Gritts Kilpatrick. Dallas, Southern Methodist University Press, [s.d.].

Includes reprint of an article by S. A. Worcester, which appeared in the *Cherokee Phoenix*, Feb. 21, 1828.

Nida, Eugene A. *The Book of a Thousand Tongues*. (See *Book of a Thousand Tongues*.)

Nöldeke, Theodor. *Compendious Syriac Grammar*. With a table of characters by Julius Euting. Translated ... from the 2d and improved German ed., by James A. Crichton. London, Williams & Norgate, 1904.

Reprinted: Tel Aviv, Zion Pub. Co. [1970].

*Nonroman Scripts in the Bibliographic Environment*. In *Encyclopedia of Library and Information Science*, 56 (supplement 19), 260–283. New York, Dekker, 1995.

O'Donnell, Sandra Martin. *Programming for the World, a Guide to Internationalization*. Englewood Cliffs, New Jersey, PTR Prentice-Hall, 1994. ISBN 0-13-722190-8.

Page, Raymond Ian. *Runes*. Berkeley, University of California Press; [London] British Museum, 1987. (*Reading the Past*). ISBN 0-520-06114-4.

British Museum Publications edition has ISBN 0-7141-8065-3.

Pavlenko, Nikolai Andreevich. *Istoria Pis'ma*. 2. izd. Minsk, Vysshiaia Shkola, 1987.

Ramsey, S. Robert. *The Languages of China*. 2d printing with revisions. Princeton, Princeton University Press, 1989. ISBN 0-691-01468-X.

Robinson, Andrew. *The Story of Writing*. London, Thames and Hudson, 1995. ISBN 0-500-01665-8.

Robinson, Theodore Henry. *Paradigms and Exercises in Syriac Grammar*. 4th ed. rev. by L. H. Brockington. Oxford, Clarendon Press; New York, Oxford University Press, 1962. ISBN 0-19-815416-X, 0-19-815458-5 (pbk.).

Roop, D. Haigh. *An Introduction to the Burmese Writing System*. [Honolulu], Center for Southeast Asian Studies, University of Hawaii at Manoa, 1997. (*Southeast Asia paper*, 11).

Originally published: New Haven, Yale University Press, 1972. (*Yale linguistic series*). ISBN 0-300-01528-3.

Ruhlen, Merritt. *A Guide to the World's Languages, Volume 1: Classification, with a Postscript on Recent Developments*. Stanford, California, Stanford University Press, 1991. ISBN 0-8047-1894-6 (v. 1).

Also published: London, Arnold. ISBN 0-340-56186-6 (v.1).

Sampson, Geoffrey. *Writing Systems: A Linguistic Introduction*. Stanford, California, Stanford University Press, 1985. ISBN 0-8047-1254-9.

Also published: London, Hutchinson. ISBN 0-09-156980-X; 0-09-173051-1 (pbk.).

Senner, Wayne M. *The Origins of Writing*. Lincoln, University of Nebraska Press, 1989. ISBN 0-8032-4202-6; 0-8032-9167-1 (pbk.).

Shepherd, Walter. *Shepherd's Glossary of Graphic Signs And Symbols*. Compiled and classified for ready reference by Walter Shepherd. New York, Dover, 1971. ISBN 0-486-20700-5.

Also published: London, Dent. ISBN 0-460-03818-4.

- Shinmura, Izuru. *Kojien* / Shinmura Izuru hen. Dai 4-han. Tokyo, Iwanami Shoten, 1991.
- Stevens, John. *Sacred Calligraphy of the East*. 3d ed., rev. and expanded. Boston, Shambala, 1995. ISBN 1-570-62122-5.
- Suarez, Jorge A. *The Mesoamerican Indian Languages*. Cambridge, Cambridge University Press, 1983. ISBN 0-521-22834-4; 0-521-29669-2 (pbk.).
- Universal Multiple-Octet Coded Character Set Coexistence and Migration*. High Wycombe, X/Open Company, 1994. (X/Open Technical Study, E401.) ISBN 1-85912-031-8.
- von Ostermann, Georg F. *Manual of Foreign Languages*. 4th ed., revised and enlarged New York, Central Book Company, 1952.
- Wemyss, Stanley. *The Languages of the World, Ancient and Modern: The Alphabets, Ideographs, and Other Written Characters of the World in Sound and Symbol*. Philadelphia, 1950.
- The World's Writing Systems*. Edited by Peter T. Daniels and William Bright. New York, Oxford University Press, 1995. ISBN 0-19-507993-0.
- W3C Recommendation*. See *Extensible Markup Language (XML)*.

This PDF file is an excerpt from *The Unicode Standard, Version 3.0*, issued by the Unicode Consortium and published by Addison-Wesley. The material has been modified slightly for this online edition, however the PDF files have not been modified to reflect the corrections found on the Updates and Errata page (see <http://www.unicode.org/unicode/uni2errata/UnicodeErrata.html>). More recent versions of the Unicode standard exist (see <http://www.unicode.org/unicode/standard/versions/>).

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and Addison-Wesley was aware of a trademark claim, the designations have been printed in initial capital letters. However, not all words in initial capital letters are trademark designations.

The authors and publisher have taken care in preparation of this book, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

The *Unicode Character Database* and other files are provided as-is by Unicode®, Inc. No claims are made as to fitness for any particular purpose. No warranties of any kind are expressed or implied. The recipient agrees to determine applicability of information provided.

*Dai Kan-Wa Jiten* used as the source of reference Kanji codes was written by Tetsuji Morohashi and published by Taishukan Shoten.

ISBN 0-201-61633-5

Copyright © 1991-2000 by Unicode, Inc.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publisher or Unicode, Inc.

This book is set in Minion, designed by Rob Slimbach at Adobe Systems, Inc. It was typeset using FrameMaker 5.5 running under Windows NT. ASMUS, Inc. created custom software for chart layout. The Han radical-stroke index was typeset by Apple Computer, Inc. The following companies and organizations supplied fonts:

Apple Computer, Inc.  
Atelier Fluxus Virus  
Beijing Zhong Yi (Zheng Code) Electronics Company  
DecoType, Inc.  
IBM Corporation  
Monotype Typography, Inc.  
Microsoft Corporation  
Peking University Founder Group Corporation  
Production First Software

Additional fonts were supplied by individuals as listed in the *Acknowledgments*.

The Unicode® Consortium is a registered trademark, and Unicode™ is a trademark of Unicode, Inc. The Unicode logo is a trademark of Unicode, Inc., and may be registered in some jurisdictions.

All other company and product names are trademarks or registered trademarks of the company or manufacturer, respectively.

The publisher offers discounts on this book when ordered in quantity for special sales. For more information please contact:

Corporate, Government, and Special Sales  
Addison Wesley Longman, Inc.  
One Jacob Way  
Reading, Massachusetts 01867

Visit A-W on the Web: <http://www.awl.com/cseng/>

First printing, January 2000.