

# ISO/IEC JTC1/SC2/WG2/IRG N1534

## Universal Multiple-Octet Coded Character Set International Organization for Standardization

Doc Type: ISO/IEC JTC1/SC2/WG2/IRG

Title: A Preliminary Review of the UTC Sources Used for Extension D

Source: John H. Jenkins, Unicode Consortium

Status: Liaison Contribution

Action: For consideration by the IRG

Date: 2008-11-12

This document discusses the nature of the main sources used by the UTC for the characters in our Extension D proposal.

Two of the characters in our submission are derived from direct email communication. We will provide the evidence for these with our complete evidence submission.

The remaining characters are derived from dictionaries. We will in this document give information on the nature of dictionaries and why we consider these to be good sources of evidence of actual use. For each dictionary, we will give full bibliographic information, the short name we use for the dictionary, and the abbreviation we use when we note it as a redundant source. This is followed by a discussion of the nature of the dictionary.

For all of these dictionaries, we have examined the entire contents of the dictionary to provide mapping tables which are part of the UniHan database, and described in UAX 38 (see <<http://www.unicode.org/reports/tr38/>>). The Extension D proposal represents characters from these dictionaries not yet encoded. Detailed information on the meaning of the source indices can be found in UAX38.

Cheung Kwan-hin and Robert S. Bauer, *The Representation of Cantonese with Chinese Characters*, Journal of Chinese Linguistics, Monograph Series Number

18, 2002.

Source name: Cheung-Bauer

Source abbreviation: CB

This monograph is a survey done by two Hong Kong linguists of extant printed sources for written Cantonese, including newspapers and other current publications. As such, it is excellent (although indirect) evidence for current use.

Bernard F. Meyer and Theodore F. Wempe, *Student's Cantonese-English Dictionary*, Maryknoll, New York: Catholic Foreign Mission Society of America, 1947.

Source name: Meyer-Wempe

Source abbreviation: MW

This is the final edition of a dictionary which was the dominant Cantonese-English dictionary for the bulk of the 20th century. This particular edition remained in print for at least forty years after its original publication.

Sidney Lau, *A Practical Cantonese-English Dictionary*, Hong Kong: Government Printer, 1977.

Source name: Lau

Source abbreviation: La

This dictionary is still available and accompanies a popular Cantonese language course, also by Sidney Lau.

Roy T. Cowles, *A Pocket Dictionary of Cantonese*, Hong Kong: University Press, 1999.

Source name: Cowles

Source abbreviation: Co

This is a late edition of a Cantonese-English dictionary in print for most of the 20th century.

Robert H. Mathews, *Chinese-English Dictionary*, Cambridge: Harvard University Press, 1975.

Source name: Mathews  
Source abbreviation: Ma

This is a large Chinese-English dictionary aimed largely at students of Mandarin. The 1975 edition is the final edition; it actually was in print for most of the 20th century.

Courtenay H. Fenn, *The Five Thousand Dictionary*, Cambridge, Mass.: Harvard University Press, 1979.

Source name: Fenn  
Source abbreviation: Fe

This is a dictionary smaller than Mathews but with a similar history. It was popular among American students of Chinese for much of the 20th century.

John Defrancis, *ABC Chinese-English Comprehensive Dictionary: Alphabetically Based Computerized*, Honolulu: University of Hawaii Press (2003).

Source name: ABC2  
Source abbreviation: AB

This is the most recent edition of a major contemporary Chinese-English dictionary in wide circulation.

辭海, single volume, traditional Chinese edition, Hong Kong: Zhonghua Bookstore, 1983 (reprint of the 1947 edition), ISBN 962-231-005-2.

Source name: Cihai  
Source abbreviation: CH

The 辭海 is a well-known and important Chinese language dictionary. This is a late reprint of one major edition.

《新校正切宋本廣韻》。台灣黎明文化事業公司 出版，林尹校訂1976 年出版。

Source name: SBGY  
Source abbreviation: SB

This is a major medieval Chinese dictionary.

*Xiandai Hanyu Cidian*, Hong Kong, Commercial Press (2002).

Source name: XHC

Source abbreviation: XH

This is a popular, modern Chinese language dictionary which has gone through numerous editions over the decades.