

CJK NEWS

UNICODE STANDARD VERSION 3.1

The Unicode Standard, Version 3.1 has been released. The primary feature of Unicode 3.1 is the addition of 44,946 new encoded characters. Together with the 49,194 already existing characters in Unicode 3.0, that comes to a grand total of 94,140 encoded characters in Unicode 3.1.

The new characters cover several historic scripts, several sets of symbols, and a very large collection of additional Chinese-Japanese-Korean (CJK) ideographs. Unicode 3.1 also features new character properties, and assignments of property values for the much expanded repertoire of characters.

All errata and corrigenda to Unicode 3.0 and Unicode 3.0.1 are included in this specification, together with significant enhancements of the Unicode conformance clauses and additions to other sections of the standard.

The base documentation for Unicode 3.1 can be found at
<<http://www.unicode.org/unicode/reports/tr27/>>

Source: unicode@unicode.org

PROPOSED CHANGES IN CHINESE ROMANIZATION GUIDELINES

Cataloging Policy and Support Office, Library of Congress
Washington, D.C. 20540-4305

The Library of Congress is reconsidering its practice of applying the new pinyin Chinese romanization guidelines. The Cataloging Policy and Support Office (CPSO) requested comments from various American bodies. Comments were due by May 1, 2001.

1) Adding an umlaut over the letter u in the syllables lue and nue

The pinyin romanization table does not place an umlaut over the letter u in the syllables lue and nue. An umlaut is used on two Chinese syllables (lue, nue) to distinguish them from other syllables that use the same letters but are pronounced differently (lu, nu). An umlaut was not added to the syllables lue and nue because their letters do not conflict with other syllables. It was also felt that the absence of diacritic marks was convenient for users, and would lead to fewer typographical errors in transcription.

It is proposed that an umlaut be added to these syllables (lue, nue) so that they will conform to the Chinese guidelines, as put forth in *Han yu pin yin fang an*. The proposed change would cause the syllables to conform to those in the ISO standard for Chinese. Proponents say the change should be made for the sake of convenient data storage and transmission, and information exchange.

Other major institutions that use romanized Chinese data do not employ the umlaut over these syllables consistently. The US Board on Geographic Names uses the umlaut over these syllables. Romanized Chinese in the National Library of China bibliographic records does not include the umlaut. Some of the National Library of Australia bibliographic records employ the umlaut, some do not.

The options under consideration are:

- 1) Retain the current policy.
- 2) Use an umlaut over the letter u in the syllables lue and nue in the future. Add the umlaut to these syllables on authority and bibliographic records on an "encountered" basis.
- 3) Use an umlaut over the letter u in the syllables lue and nue in the future. Locate these unattached syllables on authority and bibliographic records and add the umlaut to them. Add the umlaut to joined syllables on an "encountered" basis.

2) Use of apostrophe in joined syllables

The pinyin romanization guidelines call for an apostrophe to be added 1) before syllables beginning with joined letters a, e, or o, and 2) between connected syllables that end with the letter n and begin with the letter g. It was felt that adding an apostrophe between the joined letters n and g when they represented the juncture of syllables was easy to apply and made it convenient for a reader to distinguish where the break occurs.

It is proposed that the apostrophe only be added before syllables beginning with joined letters a, e or o. The second application, noted above, does not conform to the Chinese guidelines as put forth in *Han yu pin yin fang an*. The proposed change would conform to the ISO standard for Chinese. Proponents say the change should be made for the sake of convenient data storage and transmission, and information exchange. They reason that the current practice is not necessary; that cases of confusion rarely occur.

Other major institutions generally apply the apostrophe in a manner consistent with the guidelines of *Han yu pin yin fang an*. The US Board on Geographic Names applies the Chinese guidelines consistently. The National Library of China does not join romanized syllables, and so does not apply this *Han yu pin yin fang an* guideline. The National Library of Australia generally applies the Chinese guidelines for the apostrophe. Bibliographic records from the Bodleian Library, Oxford University show inconsistent application of the Chinese guidelines.

The options under consideration are:

- 1) Retain the current policy.
- 2) Cease application of an apostrophe between joined syllables when the first ends with the letter n and the second begins with the letter g. Remove this apostrophe on authority and bibliographic records on an "encountered" basis.

**MARBI PROPOSAL NO. 2001-09
(MAPPING OF EACC CHARACTERS TO UNICODE/UCS)**

The MARBI Proposal no. 2001-09 (Mapping of EACC characters to Unicode/UCS) is now available at URL:

<http://lcweb.loc.gov/marc/marbi/2001/2001-09.html>

This proposal is scheduled for MARBI discussion at its third session during the upcoming ALA Annual Meeting to be held in San Francisco, Monday, 18 June 2001

Hideyuki Morimoto

**RESEARCH LIBRARY GROUP'S CHINESE-LANGUAGE
RECORD CONVERSION**

RLG's conversion of Chinese-language records in the RLG Union Catalog from Wade Giles to Pinyin continues apace. Conversion is done "cluster by cluster", so those who have records in the RLG Union Catalog will already have some number already in Pinyin because they are in clusters that have already been converted. About half of all 1.3 million Chinese-language clusters have now been converted.

For those interested in monitoring progress, feel free to bookmark the RLG Web page, "Pinyin Project: Status of RLG Union Catalog Record Conversions" at: <http://www.rlg.org/eas/pinyinstatus.html>

Karen Smith-Yoshimura
RLG Pinyin Project Manager
E-mail: kss@notes.rlg.org

OCLC BEGINS CONVERSION OF CHINESE BIBLIOGRAPHIC RECORDS FROM WADE-GILES TO PINYIN

During the week of May 7, 2001, OCLC began its conversion of Chinese language bibliographic records in WorldCat, the OCLC Online Union Catalog, from the Wade-Giles transliteration scheme to pinyin. Working in close cooperation with the Library of Congress and the Research Libraries Group, OCLC has been planning and testing this conversion process for nearly two years.

On May 7, OCLC began converting all Chinese language bibliographic records (identified by fixed field Language Code "chi"), starting from the highest OCLC number and working backwards through WorldCat. We expect this phase to be completed by the end of August 2001. On May 8, the complete set of approximately 8900 converted Chinese language CONSER serial records were loaded into WorldCat. Once all Chinese language records are converted, OCLC will continue on to convert non-Chinese language records that contain identifiable Wade-Giles data. This phase of the conversion process is expected to be complete before the end of 2001.

October 1, 2000 was the mutually agreed upon date for the implementation of pinyin in United States bibliographic systems. OCLC would like to strongly remind users that, for the conversion to work correctly, all bibliographic records that have been created in pinyin or that have been manually converted to pinyin should contain the marker field 987. The presence of field 987 will assure that the record is not inadvertently converted again by the pinyin conversion software. For details on the use of field 987, see OCLC Technical Bulletin 240 "Pinyin Conversion Project" and "Further Pinyin Conversion Information for OCLC Users".

OCLC's conversion of LC authority records that represent headings used in Chinese language records was completed during October 2000. LC and volunteer Chinese language NACO participants completed follow-up manual conversions and cleanups during November 2000. OCLC is also offering an array of local data conversion options to both members and non-members.

You may contact OCLC's Midwest Region Marketing Manager Ms. Chris Mottayaw (e-mail: chris_mottayaw@oclc.org; phone: 800-848-5878 x6476; fax: 614-718-7444) for more information about the local pinyin conversion service options. Questions concerning the specifications for the bibliographic conversion should be addressed to Philip Melzer, Team Leader, Korean/Chinese Cataloging Team, Library of Congress at pmel@loc.gov.

source: <http://www.oclc.org/oclc/pinyin/progressreport_20010508.htm>