

Problems in the Establishment of Nonunique Chinese Personal Headings with Special Reference to NACO Guidelines and Vendor-Supplied Authority Control

Hsi-chu Bolick

Current vendor software for authority control is found to generate negative results for nonunique Chinese headings in the local catalog. After authority control, nonunique Chinese names in the bibliographic records are found routinely altered to match authority headings established in the Library of Congress (LC) authority file that are unrelated to the Chinese script of the bibliographic record. Headings with standard diacritics in the Wade-Giles romanization scheme are the most problematic. The name headings that are negatively affected in the authority control services need to be corrected before the LC Pinyin conversion project takes place around the year 2000 to prevent headings from being distorted further in the Pinyin conversion process. In this paper, I examine the reasons why vendor software produces negative results for nonunique Chinese names, and suggest measures for Chinese-Japanese-Korean (CJK) libraries in North America and vendors who supply authority-control services that include CJK data to improve the situation. These include vendor software upgrades, modifications to CJK Name Authority Cooperative procedures, etc. The authority control service at the University of North Carolina at Chapel Hill is used to illustrate the problem.

In the fall of 1997, the Library of Congress (LC) announced to the East Asian library community its decision to adopt the Pinyin system for the romanization of Chinese. The planning for this adoption includes converting LC's current Chinese

bibliographic data from the present Wade-Giles transliteration system to the Pinyin system around the year 2000. The conversion will use an approach similar to the one the National Library of Australia took in converting its CJK (Chinese-

HSI-CHU BOLICK (bolick@email.unc.edu) is acting East Asian Bibliographer/Cataloger, Collection Development Department, Davis Library, University of North Carolina at Chapel Hill. The author wishes to thank Margaretta Yarborough, Joe Collin, and Karl Lo for reading an earlier version of this article. Manuscript received September 14, 1998; accepted for publication November 17, 1998.

Japanese-Korean) data in 1996 (Eastlib Discussion List 1997).

While the East Asian library community is planning for the adoption of the new standard, I will address some problems in the conversion of nonunique personal name headings that should be taken into consideration in planning for Pinyin conversion in order to ensure Chinese bibliographic data integrity.

The use of a vendor-supplied authority control service at the University of North Carolina at Chapel Hill (UNC), under the currently available software program, generates negative results for nonunique Chinese transliterated name headings in bibliographic records. Nonunique names in bibliographic records routinely are altered to the authority headings established in the LC authority file that are unrelated to the bibliographic records (see figure 1). Headings with standard diacritics in the Wade-Giles system are the most problematic. Although, in a technical sense, the retrieval of such headings is not affected, the presence of multiple authors under one romanized form presents two problems: retrieval sets high in recall but low in precision; and misrepresented names in the local catalog that, if not corrected manually, will be distorted further in the process of Pinyin conversion when it takes place in a few years.

NONUNIQUE CHINESE NAME HEADINGS

The term "nonunique name" in the LC authority file refers to the use of one established heading for multiple people who cannot be distinguished from one another by the form of their names alone. Nonunique name authority headings occur when people share identical names, and additional personal data (such as birth and death dates or other qualifiers) are not available at the time the heading is established. It is common to encounter this type of heading in LC's authority records for all languages.

Without qualifiers to the personal name, every Chinese name transliterated under the Wade-Giles scheme is subject to becoming a nonunique heading because of

at least three factors: the nonapplication of the tonal value in transliteration practice; the use of diacritics in the Wade-Giles transliteration scheme; and Name Authority Cooperative (NACO) authority practice and normalization.

THE NONAPPLICATION OF THE TONAL VALUE IN TRANSLITERATION PRACTICE

National standard Chinese is rich in homonyms, and is a tonal language. By applying one of the basic four different tones to a Chinese syllable, a variety of unique Chinese characters is generated. But in North American cataloging practice, according to the current version of the ALA-LC romanization tables, the tonal values are ignored in transliteration of Chinese scripts (Library of Congress 1991). This practice groups Chinese characters that are pronounced with the same sound but different tones under one transliterated form. For example, combinations of tonal numerical marks 1, 2, 3, 4 and the roman letters *ta* in *ta¹ ta² ta³ ta⁴* represent at least four different Chinese characters. In current LC practice, however, these four characters are transliterated into one '*ta*' in the authority and bibliographic records.

DIACRITICS IN THE WADE-GILES TRANSLITERATION SCHEME

Structurally, the Wade-Giles transliteration scheme doesn't employ a unique roman letter to correspond to each individual Chinese phonetic sound. Instead, it uses combinations of roman letters and diacritics to stand for some distinctive phonetic values. There are four diacritics—ayn, umlaut, circumflex, and breve—in the original Wade-Giles scheme (Lu 1995). The modified Wade-Giles transliteration scheme currently used by LC and East Asian libraries in the U.S. uses two diacritics, ayn (´) and umlaut (¨), to indicate aspiration and the middle vowel sound. For example, *p´t´k´* represent aspiration of *ptk*, while *ü* is the middle vowel sound to be distinguished from *u*. Chinese characters that are transliterated into identical roman letters but

LC Authority Record for Chang, Yèu

ARN: 768061
 Rec stat: c Entered: 19820520
 ▶ Type: z Upd status: a Enc lvl: n Source:
 Roman: ■ Ref status: a Mod rec: Name use: a
 Govt agn: ■ Auth status: a Subj: a Subj use: a
 Series: n Auth/ref: a Geo subd: n Ser use: b
 Ser num: n Name: b Subdiv tp: Rules: c ¶

▶ 1 010 n 82071276 ¶
 ▶ 2 040 DLC #c DLC #d DLC #d DLC-R ¶
 ▶ 3 005 19940428061253.4 ¶
 ▶ 4 100 10 Chang, Yu ¶
 ▶ 5 400 10 Zhang, Yu ¶
 ▶ 6 400 10 Zhang, You ¶
 ▶ 7 400 10 Chang, Kung-li ¶
 ▶ 8 500 10 Chou, Heng ¶
 ▶ 9 670 [Editor of Y'un Tai-ying lai hung ...] ¶
 ▶ 10 670 Y'un, T.Y. Y'un Tai-ying lai hung ch' u yen lu, 1981
 (a.e.) #b t.p. (Chang Y'u) ¶
 ▶ 11 670 [Author of Hsiao tang chia] ¶
 ▶ 12 670 His Hsiao tang chia, 1969: #b t.p. (Chang Y'u) ¶
 ▶ 13 670 [Author of Wen i ti jen wu chi ch' i t' a] ¶
 ▶ 14 670 His Wen i ti jen wu chi ch' i t' a, 1953: #b t.p. (Chan Y'u) ¶
 ▶ 15 670 [Author of Hsin k' an ching pan ...] ¶
 ▶ 16 670 His Hsin k' an ching pan p' i p' ing cheng pai Chiang
 chuan, 1643: #b caption (Chang Y'u) ¶
 ▶ 17 670 [Author of Feng hsia] ¶
 ▶ 18 670 His Feng hsia, 1975: #b t.p. (Chang Y'u) ¶
 ▶ 19 670 [Joint author of T' ai-tzu-t' an] ¶
 ▶ 20 670 Wei, M. T' ai-tzu-t' an, 1954 (a.e.) #b t.p. (Chang Yu) ¶
 ▶ 21 670 [Author of Huan ching mi li] ¶
 ▶ 22 670 [Author of Chieh chung] ¶
 ▶ 23 670 Her Chieh chung, 1985: #b t.p. (Chang Y'u) ¶
 ▶ 24 670 [Author of Chang Y' u hsiao shuo hs' uan] ¶
 ▶ 25 670 His Chang Y' u hsiao shuo hs' uan, 1985: #b t.p. (Chang
 Y'u) ¶
 ▶ 26 670 Her Huan ching mi li, 1983: #b t.p. (Chang Y'u) vita
 (Her real name is Chou Heng) ¶

Bibliographic Record in DRA Database *Before* Vendor-Supplied
Authority Control

DBCN: AOX-3204 Entered: 04/17/1996 Modified: 04/17/1996

Type: a Bib l: m Enc l: I Desc: a Ctry: hk Lang: chi Mod:
 Srce: d Ill: Audience: Form: Cont: Gvt: Cnf: 0 Fst: 0
 Ind: 0 Fic: 1 Bio: Dat tp: s Dates: 1995 Control:

005; ;a 19960410171011.0 \$
 010; ;o 32376017 \$
 020; ;a 9624474249 (pbk.) : \$c HK\$38.00 \$
 040; ;a SFR \$c SFR \$d JTZ \$d NOC \$
 066; ;c \$1 \$
 090; ;a PL2837.Y76 \$b C48 1995 \$
 092; ;a F \$
 100; 1 ;6 880-01 \$a Chang, Y'u. \$
 245; 10;6 880-02 \$a Chuang feng pan sha / \$c Chang Y'u chu. \$
 250; ;6 880-03 \$a Ti 2 pan. \$
 260; ;6 880-04 \$a Hsiang-kang : \$b Ch`in + yuan ch`u pan she,
 \$c 1995. \$
 300; ;a 207 p. ; \$c 17 cm. \$
 490; 1 ;6 880-05 \$a Chang Y'u hs' uan huan hsi lieh ; \$v 18 \$
 500; ;a Short stories. \$

Figure 1. Example of Authority Verification Process on a Nonunique Chinese Heading.

(Continued on next page)

```

800; 1 ;6 880-06 $a Chang, Y`u. $t Chang Y`u hs`uan huan hsi
      lieh ; $v 18. $
880; 1 ;6 100-01/$1 $a !=G!:e. $
880; 10;6 245-02/$1 $a !We!Lc!@6!2_ / $c !=G!:e!UN. $
880; ;6 250-03/$1 $a !Os2!JH. $
880; ;6 260-04/$1 $a !a)!H$: $b !49+!QZ!3K!JH!Na, $c 995. $
880; 1 ;6 490-05/$1 $a !=G!:e!K0!<h2!i!3W ; $v 18 $
880; 1 ;6 800-06/$1 $a !=G!:e. $t !=G!:e!K0!<h2!i!3W ; $v 18. $

```

Bibliographic Record in DRA Database *After* Vendor-Supplied
Authority Control

DBCN: AOX-3204 Entered: 04/17/1996 Modified: 1/06/1997

Type: a Bib 1: m Enc 1: I Desc: a Ctry: hk Lang: chi Mod:
 Srce: d Ill: Audience: Form: Cont: Gvt: Cnf: 0 Fst: 0
 Ind: 0 Fic: 1 Bio: Dat tp: s Dates: 1995 Control:

```

005; ;a 19960410171011.0 $
010; ;o 32376017 $
020; ;a 9624474249 (pbk.) : $c HK$38.00 $
040; ;a SFR $c SFR $d JTZ $d NOC $
066; ;c $1 $
090; ;a PL2837.Y76 $b C48 1995 $
092; ;a F $
100; 1 ;6 880-01 $a Chang, Yu. $
245; 10;6 880-02 $a Chuang feng pan sha / $c Chang Y`u chu. $
250; ;6 880-03 $a Ti 2 pan. $
260; ;6 880-04 $a Hsiang-kang : $b Ch`in + yuan ch`u pan she,
      $c 1995. $
300; ;a 207 p. ; $c 17 cm. $
490; 1 ;6 880-05 $a Chang Yu hs`uan huan hsi lieh ; $v 18 $
500; ;a Short stories. $
800; 1 ;6 880-06 $a Chang, Yu. $t Chang Y`u hs`uan huan hsi
      lieh ; $v 18. $
880; 1 ;6 100-01/$1 $a !=G!:e. $
880; 10;6 245-02/$1 $a !We!Lc!@6!2_ / $c !=G!:e!UN. $
880; ;6 250-03/$1 $a !Os2!JH. $
880; ;6 260-04/$1 $a !a)!H$: $b !49+!QZ!3K!JH!Na, $c
      1995. $
880; 1 ;6 490-05/$1 $a !=G!:e!K0!<h2!i!3W ; $v 18 $
880; 1 ;6 800-06/$1 $a !=G!:e. $t !=G!:e!K0!<h2!i!3W ; $v 18. $

```

Figure 1. Example of Authority Verification Process on a Nonuniqued Chinese Heading.

with different diacritics or without diacritics, such as *p`i* and *pi*, *t`i* and *ti*, *k`o* and *ko*, and so on, stand not only for different pronunciations within each pair but different meanings as well. This design in the Wade-Giles scheme generates a high volume of unrelated Chinese characters that have one identical roman form except for the distinguishing diacritical marks; for instance, *ch`un*, *chün*, *chun*; *ch`üan* and *chuan*; *yu* and *yü*, etc.

NACO AUTHORITY PRACTICE AND
NORMALIZATION

Authority records contributed to the LC authority file are created by LC and

NACO participating institutions. When creating authority records for the LC authority file, NACO catalogers must follow a set of LC guidelines to ensure the integrity and quality of the large shared file. Rules of normalization are one set of the guidelines that catalogers apply to avoid creating duplicate headings.

Normalization is a process that is designed to eliminate all but the essential characters of a heading for the purpose of comparison (Cataloger's Desktop 1998). Normalization programs do this by removing all diacritics and punctuation and converting all letters to upper case (see figure 2).

Under this normalization rule and

Catalog form:	Ile-de-Montréal (Québec)
Normalized form:	ILE DE MONTREAL QUEBEC

Figure 2. Example of Normalized Heading.

NACO guidelines, a new heading is created only when the normalized form does not match the normalized form of an already established heading in the LC authority file. If its normalized form is found in the authority file, the heading is considered established. When applying normalization to Chinese transliterated names, diacritics, which are essential for distinguishing Chinese syllables under the Wade-Giles system, are considered not essential and therefore are eliminated from the heading comparison process. If headings are found normalized to the same form, they are put together in one authority record (see figure 3).

Under NACO practice, the Chinese names in example 1 of figure 3 share one authority record. The authorized heading for these two names will either be *Cheng, Ch'en*, or *Ch'eng, Chen*, depending upon which name gets created first in the authority file. Likewise, the Chinese names in example 2 of figure 3 have one authority record, and the established heading takes the diacritics associated with the name first created in the authority file.

When dealing with a new nonunique name whose normalized form matches an established heading in the authority file, NACO catalogers do nothing to the established heading but revise the existing authority record in the note area (670 fields in the authority record) to include new information on additional names and their related citations. New reference headings for different names are added to the authority record when appropriate. If the existing authority record is coded as a unique name, it is changed to a nonunique name record when new names are found that share the same normalized form.

For instance, in the above examples, when the new name *Cheng, Ch'en* appears in the cataloging source and *Ch'eng, Chen* is found already estab-

lished as an authorized heading in the LC authority file, *Cheng, Ch'en* is only added to the note

area. *Ch'eng, Chen* remains as the established heading for this record. Cross-references related to *Cheng, Ch'en* can be added to the record as long as they do not normalize to the established form.

If unaware of this practice for nonunique names, one is likely to interpret *Ch'eng, Chen* as it appears in the authorized heading field of the authority record, as the authoritative form for *Cheng, Ch'en*. Likewise, one might take *Chang, Ch'üan*, if established first, as the established authority heading for *Ch'ang, Chuan* and *Chang, Chuan* (see figures 1 and 4).

CATALOGING FOR NONUNIQUE CHINESE NAMES

As a rule, catalogers choose LC-established headings over name forms found in the cataloging source for heading forms in bibliographic records to control quality and uniformity of access points; yet in creating bibliographic heading forms for Chinese nonunique names, this is not always the case. In the example of the heading *Ch'eng, Chen*, when the new name *Cheng, Ch'en* appears in the cataloging source, for which a nonunique authority heading is found established under *Ch'eng, Chen*, instead of using the established heading *Ch'eng, Chen*, the cataloger enters *Cheng, Ch'en* as the author's name in the bibliographic record to

- | | | | |
|----|----|------------------|---------------|
| 1. | a. | Catalog form: | Cheng, Ch'en |
| | | Normalized form: | CHENG, CHEN |
| | b. | Catalog form: | Ch'eng, Chen |
| | | Normalized form: | CHENG, CHEN |
| 2. | a. | Catalog form: | Chang, Ch'üan |
| | | Normalized form: | CHANG, CHUAN |
| | b. | Catalog form: | Ch'ang, Chuan |
| | | Normalized form: | CHANG, CHUAN |
| | c. | Catalog form: | Chang, Ch'üan |
| | | Normalized form: | CHANG, CHUAN |

Figure 3. Normalization in Chinese Name Headings.

```

ARN: 681573
  Rec stat: c      Entered: 19820208
  ▶ Type: z      Upd status: a      Enc lvl: n      Source:
  Roman: █      Ref status: a      Mod rec:      Name use: a
  Govt agn: █   Auth status: a      Subj: a      Subj use: a
  Series: n     Auth/ref: a      Geo subd: n   Ser use: b
  Ser num: n    Name: a      Subdiv tp: █  Rules: c ¶

  ▶ 1 010      n 81136731 ¶
  ▶ 2 040      DLC #c DLC ¶
  ▶ 3 005      19840912155140.5 ¶
  ▶ 4 100 10   Y`u, Yu-jen, #d 1879-1964. ¶
  ▶ 5 400 10   Y`u, Yu-jen, #d 1878-1964 #w nnaa ¶
  ▶ 6 400 10   Yu, Youren, #d 1879-1964 ¶
  ▶ 7 400 10   U, Y`ujin, #d 1879-1964 ¶
  ▶ 8 670      His Yu-jen shih ts`un, 1956. ¶
  ▶ 9 670      Liu, Y.T. Min kuò Y`u Yu-jen hsien sheng nien p`u,
  1981: #b p. 1 (Y`u Yu-jen, b. 1879) ¶
  ▶ 10 670     Biog. dict. of Repub. China, 1971: #b v. 4, p. 74
  (Y`u Yu-jen, 4-11-1879-11-10-1964) ¶

```

Figure 4. The Unique Name Authority Record (See *Name: a* in the Fixed Field).

reflect the Chinese characters.

This practice is necessary in cataloging Chinese materials for the purpose of appropriately matching roman data to the corresponding vernacular data in the Machine-Readable Cataloging (MARC) record. The current negative authority control result found on nonunique Chinese names occurs because available vendor software does not take into account this special cataloging practice for nonunique Chinese names.

VENDOR-SUPPLIED AUTHORITY CONTROL SERVICES

In 1995, three vendors, Library Technology Inc. (LTI), Blackwell North America (BNA) (now OCLC Authority Control Services), and the Washington Libraries Network (WLN), were invited to UNC to demonstrate their authority services. During their presentations, LTI, BNA, and WLN displayed similar approaches and software designs in processing authority control for local files.

LTI was chosen to provide vendor authority control for UNC. Its procedure for authority control is as follows (LTI 1998):

First, headings are "normalized" for spacing and punctuation to increase the probability of a link between a catalog record heading and an authority record heading. Next, the normalized headings are compared against a comprehensive index of

authorized and variant-form headings generated from authority records. When a match occurs, a link is created between the authority record heading and the catalog record heading. If the catalog record heading matches a *see* reference in an authority record, the catalog record heading is replaced by the heading in the LC authority record.

Existing authority control software lacks special instructions that would account for the special cataloging practice for nonunique Chinese names. All nonunique Chinese names in bibliographic records that are normalized in the same way as authority headings (but that differ from them in diacritics) are treated as incorrect forms, and thus are changed to authority headings. As a result, after vendor authority control processing, one finds the roman data doesn't match the vernacular data for nonunique names in the bibliographic record.

Taking again the example of *Ch'eng, Chen*, under the current vendor software process, because *Ch'eng, Chen* is the established heading, all occurrences of *Cheng, Chen* will be changed to the authority heading *Ch'eng, Chen* in the catalog after the authority control service.

LTI, BNA, and WLN all agreed that they have this problem in their services. It appears that on the current market there are no available commercial vendors who offer authority control services that take this situation into consideration.

WADE-GILES, PINYIN SCHEMES, AND NORMALIZATION

Normalization, the current industry standard for database management—which by design aims for effective data storage and management—requires that the relationship between the related elements (the primary key and the data) is one-to-one or many-to-one, but never one-to-many or many-to-many (Dutka 1989). That is to say between personal names and authority records, normalization functions effectively when one authority record relates to one unique personal name, or one authority record stands for one unique person with multiple names, but not vice versa when multiple unique personal names are represented by one authority record.

The application of the Wade-Giles scheme in current cataloging and NACO practice for nonunique names shows a breakdown with the principle of normalization. Diacritics used in the Wade-Giles scheme to represent the uniqueness of sounds and words are considered not essential in the existing normalization rules and thus are excluded from the process. This explains why procedures that occur after the normalization process in authority-control services cannot process nonunique Chinese transliterated names properly.

The Wade-Giles scheme will soon be dropped for cataloging work, and the negative effect associated with it under normalization will be eased substantially when the new Pinyin scheme is adopted. Nevertheless, it will not disappear. Compared to the Wade-Giles scheme, the Pinyin scheme reduces the use of diacritics down to the umlaut for distinguishing certain unique Chinese characters. For instance, the Chinese character “green” is transliterated in Pinyin as *lǜ*, while the Chinese character “furnace, stove” is transliterated in Pinyin as *lú*. Under current normalization, *lǜ* and *lú* again appear as the same form. Therefore without software upgrades or modification in the transliteration scheme, the nonunique name syndrome in the authority control process will continue to happen with Pinyin data.

PINYIN CONVERSION

Before the year 2000 (when LC will start converting its database to the Pinyin scheme), libraries preparing to follow LC in converting their local files to Pinyin not only should undertake correction of all nonunique Chinese names that are affected negatively by the authority control process as described above but also should work on correcting diacritics in the bibliographic records to prevent data distortion in the conversion process.

The National Library of Australia (NLA) in 1996 successfully converted its CJK data from the Wade-Giles scheme to the Pinyin system with computer technology. After closely observing the process and the conversion result at NLA, LC announced that its Pinyin conversion project would be carried out using the same approach as NLA's.

Describing the CJK conversion project in detail, Groom (1997) wrote that one key element in designing the software program for Australia's Pinyin conversion project was the setting up of a conversion table that listed and correlated all valid Wade-Giles words with their Pinyin counterparts. With the conversion table in place, the conversion program was thus able to identify and subsequently convert all Wade-Giles words to their Pinyin counterparts when run against the local bibliographic file.

The conversion table was written precisely based on the Wade-Giles rules to map each Chinese character to its Pinyin counterpart, taking into account the exact use of Wade-Giles diacritics and their position in the syllables, etc. The conversion program carries out the conversion only when the bibliographic data match correctly to the conversion table. This is to say local bibliographic records submitted for conversion are required to have the same Wade-Giles forms and diacritics as the conversion table to guarantee accurate data conversion. Any spelling errors in roman forms or misused diacritics will cause incorrect conversion or no conversion at all.

Groom (1997) reported that conversion errors were found when records con-

tained unexpected diacritics. In some cases, diacritics were not input as prescribed in the Wade-Giles scheme, with the result that no conversion took place because a diacritic discrepancy occurred between the catalog entries and the conversion table. In other cases, diacritics were omitted in the NLA catalog entry, resulting in an erroneous conversion to another heading.

SUGGESTIONS

The inadequacy of vendor-supplied authority control of Chinese name headings demonstrates an immediate need for the East Asian library community to find solutions for the incompatibility between the normalization rules governing the setup of Chinese personal headings in the NACO program and the design of the standard transliteration scheme used in North America for transcribing Chinese. Through examination it appears that a direct method for resolving such incompatibility is simply to make all Chinese personal headings unique in the LC authority file. Use of unique headings not only allows Chinese names to be processed correctly under normalization rules but is effective for authority control. As a means of doing this and solving the current authority control problem, several approaches are possible: improving and developing new system software; completing NACO headings with dates; providing clear instructions regarding the use of Roman forms for authorized headings; and eliminating diacritics from the Pinyin scheme.

IMPROVING AND DEVELOPING NEW SYSTEM SOFTWARE

Before a permanent way is found to deal effectively with romanized Chinese names in the LC authority file, in the short term we need to request that a new algorithm be developed immediately in vendor software that can halt the incorrect replacement of nonunique Chinese names in the bibliographic records with authority heading forms. One way of updating the current vendor software for

this purpose is to incorporate the NAM element in the fixed field of the USMARC Authority Format record into the new algorithm so that the software will perform global replacement of catalog forms only to those headings that are linked to *unique* name records. Headings that are linked to nonunique authority records should be marked and stored in a file for further processing. The review of unprocessed nonunique names should not be a heavy burden for catalog maintenance because most of these names deliberately are kept in those forms by human decision at the time of cataloging. Thus, the concern over the error rate for those names is small.

In the longer term, software must be developed that incorporates vernacular data in the normalization process. Adding vernaculars to authority records has been advocated by the East Asian library community to solve problems of differentiating Chinese names given the high volume of homonyms found in transliterated data in authority records (Yu 1996; Morimoto 1996).

To increase the uniqueness of Chinese names in the LC authority file, the new software should not only have the capability of storing and displaying vernacular data, but should also include a mechanism that takes vernacular data into consideration in the normalization process. That is, an additional procedure should be devised that would involve comparing vernacular data for headings that are identical when normalized. Headings that would otherwise be conflated could be kept separate when the vernacular data do not match. Authority control software would follow the same design for heading comparison and correction. This extra procedure would substantially improve effectiveness in controlling Chinese names and prevent false conversions during the authority control process.

A more cooperative effort must be made through the NACO program to improve Chinese name heading quality. Before any new updates are developed in the software, NACO catalogers and the NACO program can reduce or prevent

the normalization impact to Chinese names in the meantime by making the heading unique in ways such as: completing NACO headings with dates; providing clear instructions regarding the use of roman forms for authorized headings; and modifying the Pinyin scheme to eliminate the use of diacritics.

COMPLETING NACO HEADINGS WITH DATES

Adding birth and death dates to Chinese romanized names is very effective for establishing unique headings in the LC authority file. All NACO catalogers should be encouraged or required to provide such data to Chinese headings when available. To make this a common and practical practice in NACO work, not only should the NACO program be expanded to include more East Asian catalogers to share the burden and expense of such authority work, but a global effort should be made to solicit information from libraries, publishers that can provide such information through Internet, or author introductions in published books (Eastlib Discussion List 1998).

PROVIDING CLEAR INSTRUCTIONS REGARDING THE USE OF ROMAN FORMS FOR AUTHORIZED HEADINGS

Instructions must be provided about whether to use roman forms that appear in the source or author-supplied roman forms as authoritative forms instead of using systematic transliterated forms from the Chinese scripts. The *Anglo-American Cataloguing Rules*, 2d ed., 1988 revision rule 22.3C2, and its alternative rule, which direct the setup of Chinese headings in NACO practice are insufficient for catalogers to handle the increasingly common situation where, in the cataloging source, the romanized form of a name and its Chinese script coexist on the title page.

Due to a lack of guidelines in the NACO procedures for such circumstances, the choice of forms for these names varies among catalogers. Some headings are established under 22.3C2 and its alternative rule, using the most

popular form of authors as known by readers and in reference sources. For some headings lacking reference sources or guidance in the NACO procedure, it is difficult to comply fully with chapter 22 to establish unique Chinese headings, and catalogers refer to Rule 1.0 for guidance, which prescribes transcribing the statement of responsibility of a work in the language and script in which it appears. As a result, the former rule usually brings the roman forms already supplied by the source as authority headings, while the latter rule results in headings that are systematically transliterated from the original scripts.

Roman forms systematically transliterated from Chinese scripts, as discussed earlier, frequently result in a high volume of homonyms, thus producing nonunique headings. If NACO guidelines clearly prescribe 22.3C2 and its alternative to take precedence over rule 1.0E in establishing authority headings in such situations, more unique name headings will result.

The author or publisher-supplied roman forms appearing in the source not only indicate the authors' choice of forms in their names, but are usually formulated less systematically and predictably than those generated according to the transliteration scheme. Therefore, they are less likely to become nonunique names in the authority file than names formulated by the transliteration scheme. Such practices also reduce the confusion of public catalog users who recognize authors' names in their roman forms from publications rather than the forms from the prescribed transliteration scheme. Figure 5 displays the difference between names transliterated with the Wade-Giles scheme and those supplied by authors or publishers.

Under the Wade-Giles scheme, *Tung Chien-hua* represents hundreds of different Chinese characters for names, but the combination of *Tung Chee Hua* together, so far, only represents one personal name. Likewise, *Ng Chi-sum*, formulated under the Cantonese pronunciation, is drastically different from *Wu Chih-sen*, formulated under the standard Mandarin sound on which most transliteration schemes for Chinese are based.

```

ARN: 1292081
  Rec stat: c      Entered: 19850314
▶ Type: z      Upd status: a      Enc lvl: n      Source:
  Roman: ■      Ref status: a      Mod rec:      Name use: a
  Govt agn: ■    Auth status: a      Subj: a      Subj use: a
  Series: n      Auth/ref: a      Geo subd: n    Ser use: b
  Ser num: n     Name: b      Subdiv tp: ■   Rules: c ¶
▶ 1 010      n 85047803 ¶
▶ 2 040      DLC #c DLC #d DLC #d DLC-R ¶
▶ 3 005      19901016100402.2 ¶
▶ 4 100 10   Ch'en, Ch'en ¶
▶ 5 400 10   Chen, Zhen ¶
▶ 6 670      [Joint author of Wen tzu] ¶
▶ 7 670      Yeh, T.H. Wen tzu, 1983: #b t.p. (Ch'en Ch'en) ¶
▶ 8 670      [Joint editor of Chien ming Sung shih] ¶
▶ 9 670      Chien ming Sung shih, 1985: #b t.p. (Ch'en Chen) ¶
▶ 10 670     [Editor of Min tsu ti ch' u ts' ai cheng kai lun] ¶
▶ 11 670     Min tsu ti ch' u ts' ai cheng kai lun, 1989: #b t.p.
      (Ch'en, Ch'en ¶)

```

Figure 5. Nonunique Name Authority Record (See *Name: b* in Fixed Field, and All 670 Fields).

ELIMINATING DIACRITICS FROM THE PINYIN SCHEME

The Pinyin scheme will no doubt work much better in normalization for Chinese name headings than the current Wade-Giles scheme. Not only does Pinyin allow each sound to be associated with a unique roman symbol, but it also uses fewer diacritics. The use of diacritics to distinguish a pair of vowels or aspiration in Chinese has made Wade-Giles very ineffective in normalization. Thus, the only diacritic introduced for use in the incoming Pinyin scheme—the umlaut—would need to be eliminated and replaced with an alternative expression in the scheme. Repeating the roman letter that the umlaut attaches to is one alternative to representing the sound with umlaut. For instance, *li* would be presented as *luu*. In the pair of words *lü* (meaning “green” in Chinese) and *lu* (meaning “stove” in Chinese) can thus be distinguished as *luu* and *lu* in normalized form. This practice is common in publications and in other transliteration schemes used in Taiwan, Japan, and in the Macintosh Chinese language kit, for example. To allow the access of Pinyin headings where standard umlauts are used, Pinyin entries with standard umlauts should be entered as cross-references in authority records.

Data processing for nonroman languages, especially Chinese, in areas of library services is by no means easy. But

with careful design in software to incorporate Chinese vernacular data, adjustment to transliteration schemes, and cooperation among NACO catalogers to set up unique personal headings, I am confident that Chinese names can be processed as effectively as names in other languages in the online environment.

WORKS CITED

- Anglo-American cataloguing rules*. 1988. Prepared under the direction of the Joint Steering Committee for Revision of AACR, a committee of the American Library Association, the Australian Committee on Cataloguing, the British Library, the Canadian Committee on Cataloguing, the Library Association, the Library of Congress. 2d ed., 1988 revision, ed. Michael Gorman and Paul Winkler. Chicago: American Library Association.
- Cataloger's desktop. 1998. Descriptive cataloging manual. Z1: Name and series authority records. Washington, D.C.: Library of Congress, Cataloging Distribution Service (issue 3).
- Dutka, Alan F. 1989. *Fundamentals of data normalization*. Reading, Mass.: Addison-Wesley.
- Eastlib Discussion List. 1997. Library of Congress' announcement of its intention to convert to Pinyin romanization . . . Sept. 29.
- . 1998. Survey of CJK NACO participants. June 16.
- Groom, Linda. 1997. Converting Wade-Giles cataloging to Pinyin: The development and implementation of a conversion program for the Australian National CJK Services. *Library resources & technical services* 41: 269–70.

Library of Congress. 1991. *ALA-LC romanization tables*. Washington, D.C.: Library of Congress, Cataloging Distribution Service.

Library Technologies, Inc. 1998. Authority control. Online. Available: <http://www.librarytech.com/a-aint-b.htm>.

Lu, Suping. 1995. A study on the Chinese romanization standards in libraries. *Cata-*

loging & classification *quarterly* 21, no. 1: 81-96.

Morimoto, Hideyuki. 1996. Memorandum to RLG East Asian Studies Librarians' Forum (unpublished).

Yu, Abraham. 1996. Enhancement and linkage of the online name authority file with the bibliographic database. *CEAL bulletin*, no. 106: 1-6.

Elegant Solutions for Preservation



A DIVISION OF LIBRARY BINDING SERVICE



- Pamphlet Binder _____
- Music Binder _____
- Archival Folders _____
- Manuscript Folders _____
- Hinge Board Covers _____
- Academy Folders _____
- Newspaper/Map Folders _____
- Bound Four Flap Enclosure _____
- Archival Binders _____
- Polypropylene Sheet & Photo Protectors _____
- Archival Boards _____

ARCHIVAL PRODUCTS

A Division of Library Binding Service

**Call
for a complete
catalog**

P.O. Box 1413
Des Moines, Iowa 50305-1413
800.526.5640
Fax 800.262.4091
archival@ix.netcom.com
<http://www.archival.com>