A Study of the Problems in the Development of the Standardized Character Encoding of Nushu

suzuki toshiya[†1]

Nushu is a writing system which was (is) used by women living around Jiangyong county (江永県) of Yongzhou city (永州市), Hunan Province (湖南省) in China. On 2007, Zhang Liming submitted her initial proposal to include 449 Nushu characters into ISO/IEC 10646. Since her initial proposal, ISO/IEC JTC1/SC2/WG2 has been discussing the appropriate character set of Nushu and the character naming convention, the ordering rule. In spite of the interval during 2010-2011, the proposal draft of the first amendment (PDAM1) to ISO/IEC 10646 4th edition had included the Nushu character set. But in the PDAM1 ballot, many serious concerns are given by Japanese experts. As a result, Nushu ad-Hoc group under ISO/IEC JTC1/SC2/WG2 decided to postpone the standardization to the next amendment. In this report, the given concerns and follow-up works are summarized.

1. Introduction

1.1 Introduction of Nushu

Nushu (女書) is a writing system which was (is) used by women living around Jiangyong county (江永県) of Yongzhou city (永州市), Hunan Province (湖南省) in China. The name "Nushu" is Chinese word meaning "women's writing". The features of the scripts are recognized as follows;

- Each character presents a syllable, although multiple characters for same syllable are used, and a character can present different syllables.
- Language written by Nushu is dialectic Chinese; the major sound system in Nushu material is Chenguan Tuhua (城関土話), although the major spoken language of Nushu users is different. Anyway, written and spoken languages by Nushu users are known as the members of Jiangyong Tuhua (江永土話) family with 7 tones, they are significantly different from Mandarin Chinese with 4 tones.
- Nushu is mainly used for the private (one-to-one) communications between the family (grand mother, mother, and daughter) or the close friends of women. Few public (one-to-many) documents written by Nushu like flyers, posters and billboards are known. No record of wide education or training of Nushu is known. Also no library archiving old Nushu documents is known.

Figure 1: Sample of Nushu Text[15]

The origin of Nushu writing system is still unclear. The earliest official record reporting a strange script used by women in the region was in the era of Republic of China. As the studies of other minority or historic scripts in China, many hypotheses moving the origin of Nushu to far older era are proposed (e.g. Xie Zhiming (謝志民) insisted Nushu has the history longer than 3000 years). Although there are many short-lived hypotheses insisting non-Han cultural background of Nushu, the continual studies of Nushu materials are based on the assumption that the most Nushu characters are transformed from Hanzi in post-Kaishu era. Even the Nushu dictionaries compiled by the scholars insisting non-Hanzi origin, each Nushu characters are described by corresponding Hanzis. Also the character set in Nushu document is often smaller in comparison with the corresponding Chinese text; although per-character transliteration from Nushu to Chinese text, many phonetic loaning is found.

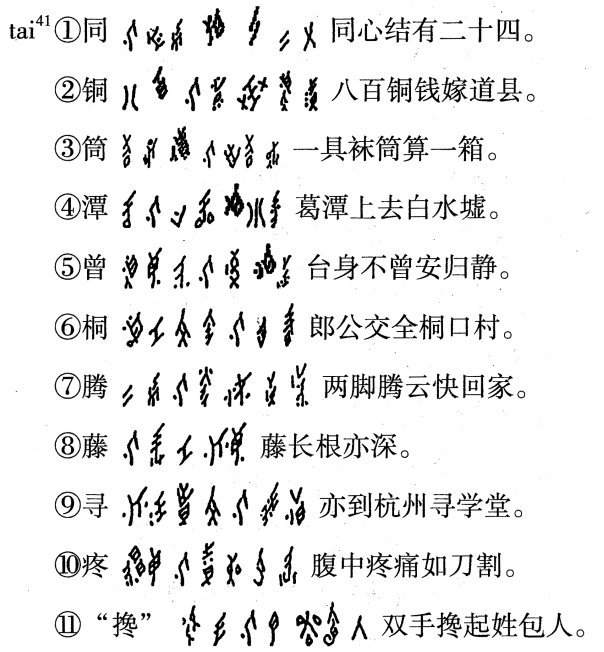


Figure 2: Syllabicalization or Phonetic Loaning in Nushu[12]

The motivation to develop Nushu instead of using Hanzi is supposed to be the solution to the inequality of the education for women (in the region, women were isolated from the education of writing). The hypothesis could explain why few public documents including Nushu and no systematic training of Nushu are known. Because of the lack of public communications by

†1 Hiroshima University, Faculty of Integrated Science and Arts

Nushu documents, the identity of the Nushu character is not stabilized yet. A character that user A uses to describe the syllable X could be used to describe completely different syllable Y by user B. If there is a communication between user A and user B, such disunions will be solved. But the lack of the wide communication made such disunions preserved.

The academic studies on Nushu is since the report of Gong Zhebing (宮哲兵) on 1983[10], therefore the background motivation to use Nushu was already lost in the region. Most Nushu users were already aged over 70 when Gong reported Nushu, and the successors are not known at that time. By the increase of academic interest of this script since the end of 1990s, the local government tried to use Nushu as a promotable local culture, and some schools are tentatively opened. However, the training seems to be the calligraphic duplication of the existing works, not for their own communication. Therefore, the general and basic character set of Nushu for daily communication is not studied yet.

1.2 Brief History of Nushu Standardization in ISO

Although the basic character set or the authorized typefaces are not stabilized, the standardization of Nushu character encoding in the international standard was started. Zhao Liming (趙麗明) submitted her initial proposal to standardize the Nushu character set including 449 characters[1]. Zhao's character identification method is based on the statistical survey for the huge collection of her Nushu materials, Zhonghua Nushu Heji (中国女書合集)[13]. As most Nushu decipherments do, Nushu Henji makes the per-character transliterated text from Nushu text to Chinese text (see Figure 3). In next, the statistics of the corresponding Hanzi for each Nushu glyphs are investigated (see Figure 4). Zhao chosen the most dominant identity of a Nushu character by the most frequently used pair of the phonetic value and corresponding Hanzi (see Figure 5Figure 6). Just after the submission of Nushu character list to ISO/IEC 10646, the character list is used as a part of Nushu Duben (女書読本)[15].

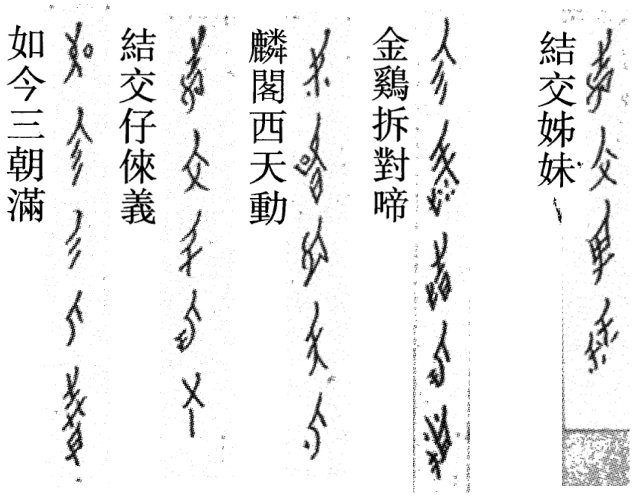


Figure 3: Sample of Transliteration in Nushu Heji[13]

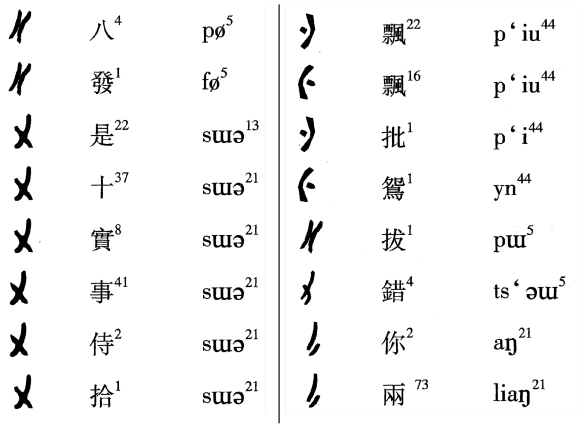


Figure 4: Nushu character list in Nushu Heji[13]

（八）	pø^{5}	八 326
	pəŋ^{44}	拔 1
（十）	sɯə^{33}	十 802 事 309 实 34 侍 2
		拾 1 莳 1
	sɯə^{13}	是 23
	ɕi^{33}	誓 8
（卜/飄）	p'iu^{44}	飄 60-3 漂 2
	p'i^{44}	批 14 拔 6
	vu^{13}	雨 11 武 1 舞 1
	u^{5}	屋 9-3
	p'ai^{44}	喷 6
	p'əŋ^{21}	片 5 骗 1
	piu^{44}	标 1
	piou^{42}	嫖 1
	y^{13}	宇 1 羽 1

Figure 5: Nushu character list in WG2 N3598[3]

Two-stroke characters

1B102	𛄂	NUSHU CHARACTER NA33 → 4E8C 二
1B103	𛄃	NUSHU CHARACTER TSHA5
1B104	𛄄	NUSHU CHARACTER IE21 → 1B173 𛅳 nushu character ie21-a
1B105	𛄅	NUSHU CHARACTER POE5
1B106	𛄆	NUSHU CHARACTER SWE33
1B107	𛄇	NUSHU CHARACTER PHIU44
1B108	𛄈	NUSHU CHARACTER U5

Three-stroke characters

1B109	𛄉	NUSHU CHARACTER CYA35 → 6C34 水
1B10A	𛄊	NUSHU CHARACTER TIE42
1B10B	𛄋	NUSHU CHARACTER NJIE33

Figure 6: Character List in ISO/IEC 10646:2014/Amd.1 (WG2 N4484)[6]

2. Problem in Proposed Character Set

2.1 Stability of the Transliteration

The identity of the Nushu character in Zhao's study is tightly combined with corresponding Hanzi used in the transliteration. Although no errata to Nushu Heji are published, it seems that Zhao is updating the transliteration text internally, as shown in Figure 7. The identity of the character is not stabilized. Of course, some dictionaries referring Nushu Heji but identifying the character by their own criteria propose the corrections of the transliteration text (e.g. Xie Zhming's "Nuzi Zidian" p. 344)

WG2 N3337 (2007)	WG2 N3598 (2009)
（多） lɘɯ^{44} 多 542 知 14 闹 2 lɘɯ^{33} 落 126 洛 3 ø^{44} 衣 93 luoɯ^{44} 单 11	（多） lɘɯ^{44} 多 542 lɘɯ^{33} 落 126 洛 3 luoɯ^{44} 单 11 丹 11 nɘɯ^{33} 闹 2 ts'ɯə^{21} 刺 2
(not found)	（衣） ø^{44} 衣 93

Figure 7: Example of Incompatible Update from the Proposal in 2007 to that in 2009[7]

It should be noted that most Nushu text is written in the rhyming poetry format; the phonetic value is restricted by the surrounding context, but the meaning is often interchanged with other characters with same phonetic value. In Nuhan Zidian (compiled before Nushu Heji), the number of corresponding Hanzi to a Nushu character representing the syllable TAI41 is 11 (see Figure 2), but the statistic survey in Nushu Duben lists for TAI42 is only 2 (see Figure 8).

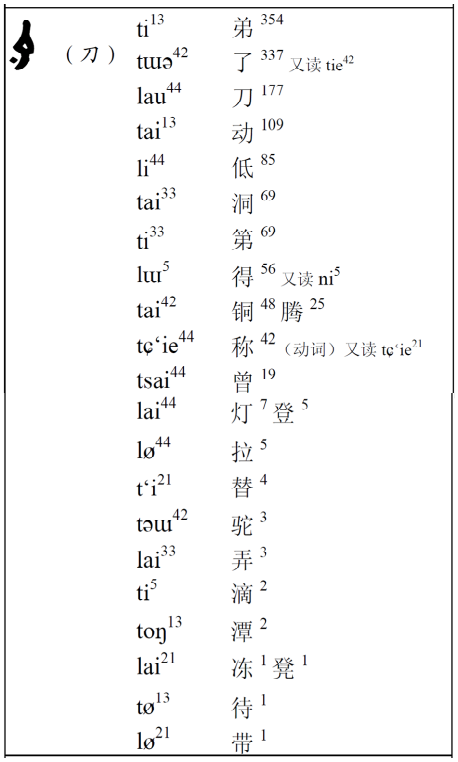


Figure 8: Corresponding Hanzis in Zhao's research corresponding to the Nushu glyph in Figure 2[12]

2.2 Unclear Priorities to non-Hanzi looking Glyphs

Among the published collection of Nushu materials, Nushu Heji is the largest collection (ca. 220,000 Nushu characters from 650 documents). The statistic survey of Nushu Heji was published as "Nushu Yongzi Bijiao"[14] on 2006. It took the per-author statistics as Figure 9.

		传本佚名	高銀仙	义年华	阳焕宣	何艳新
mau^{42}	毛	毛 5	毛 9 毛 7	13	毛 4	毛 8 毛 2
ȵioŋ^{44}	英	英 1	英 44	英 5	英 13	英 47

Figure 9: Per-Author Statistics in Nushu Yongzi Bijiao[3]

However, when per-author statistics is simplified to a pan-author statistics (like Figure 5), Zhao put the priorities to the glyphs that are difficult to remind the corresponding Hanzi, and it caused many inconsistencies. In the case of "毛" entry in Figure 9, the glyph looking like as a skewed "毛" (found in 高銀仙, 義年華 and 何艷新) was excluded because "this glyph is not traditional Nushu glyph, because 陽煥宣 did not use", although skewed "毛" glyph frequency is higher than the traditional glyph (30 versus 18). As a result, the skewed "毛" does not appear in the frequency table at all. On the other hand, the glyph looking like skewed "英" showing similar contrast (only 3 authors use it) is included in the proposed character set.

The statistics included in the proposals to ISO/IEC 10646 show the remarkably low frequency to the daily-used characters as Figure 10. It is supposed that the filtering of Hanzi-looking glyphs caused such result.

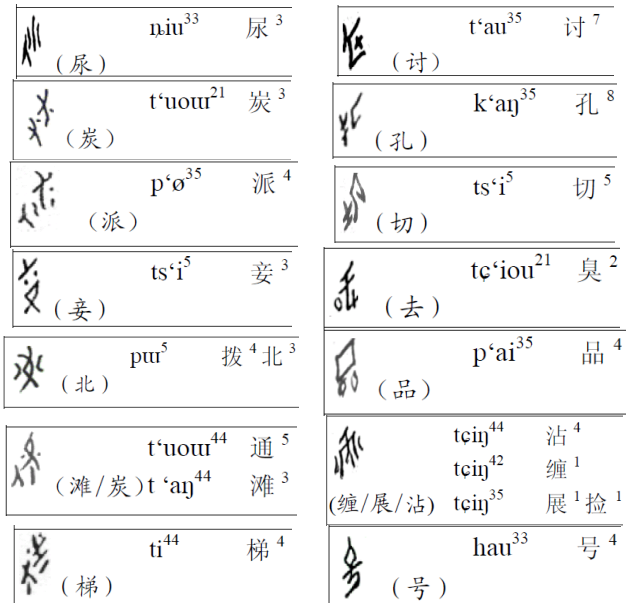


Figure 10: Rarely used character in the proposed Nushu character set[7]

2.3 Discussion in WG2 on the Proposal

Japanese experts express their concerns on the reliability of the proposed character set and their properties since 2009. The concern on the stability of the most dominant phonetic value in the proposed character names was accepted in ISO/IEC

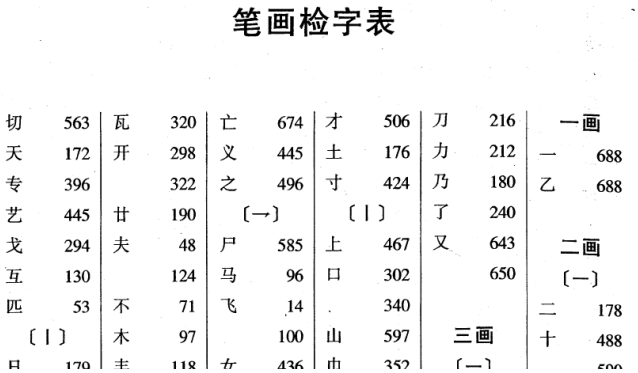
JTC1/SC2/WG2 62nd meeting on February 2014, WG2 decided to exclude the phonetic values from the character names. The inclusion of Nushu in ISO/IEC 10646 itself was either postponed from ISO/IEC 10646:2014/Amd.1 to Amd.2, but the inclusion itself is not reverted. Therefore, the exclusion of the unstable entities in the proposed character set, and the definition of the maintainable collation method are urgent task for the expert.

3. Existing Collation Methods for Nushu

As summarized in previous section, the definition of the maintainable collation method is the urgent issue for the experts. In this section, the collation methods in existing Nushu dictionaries are summarized.

3.1 Collation via Hanzi

For most scholars, the decipherment of Nushu documents is based on the per-glyph transliteration from Nushu to Hanzi. For such scholars, it is reasonable to identify Nushu glyphs by corresponding Hanzi. One of the earliest Nushu dictionary, Nushu Zidian (女書字典) by Zhou Shuoyi (周碩沂)[11], has phonetic index index (using consonant as major key, and vowel as minor key; the order of the consonant is same with PinYin for Mandarin Chinese) and corresponding Hanzi index (total stroke number and 1st/2nd stroke stype classification, see Figure 11).


笔画检字表

切	563	瓦	320	亡	674	才	506	刀	216	一画	
天	172	开	298	义	445	土	176	力	212	一	688
专	396		322	之	496	寸	424	乃	180	乙	688
艺	445	廿	190	〔一〕		〔丨〕		了	240		
戈	294	夫	48	尸	585	上	467	又	643	二画	
互	130		124	马	96	口	302		650	〔一〕	
匹	53	不	71	飞	14		340			二	178
〔丨〕		木	97		100	山	597	三画		十	488
日	179	未	118	女	436	巾	352	〔一〕			500

Figure 11: Hanzi Index in Nushu Zidian by Zhou[11]

The user could not know the representative Nushu glyph before opening the main text page. Although several Nushu glyphs could remind the corresponding Hanzi easily, most Nushu glyphs have quite different shapes from the semantically corresponding Hanzi. Therefore, even if the identification by Hanzi itself might be reasonably acceptable for some researchers, it is not useful to process Nushu document without Chinese transliteration.

3.2 Collation via Phonetic Values

For the scholars classifying Nushu as syllabic script, it is reasonable option to use their phonetic values as the most dominant keys to identify the glyphs. However, due to the different sound system in Chenguan Tuhua, the phonetic value classification is not stabilized in the society of Nushu scholars; Zhao finds 21 consonants, 37 vowels and 7 tones, Zhou finds 20 consonants, 34 vowels and 6 tones, and Xie finds 21 consonants, 32 vowels and 5 tones. As a result, the neutral unification of existing data is already difficult.

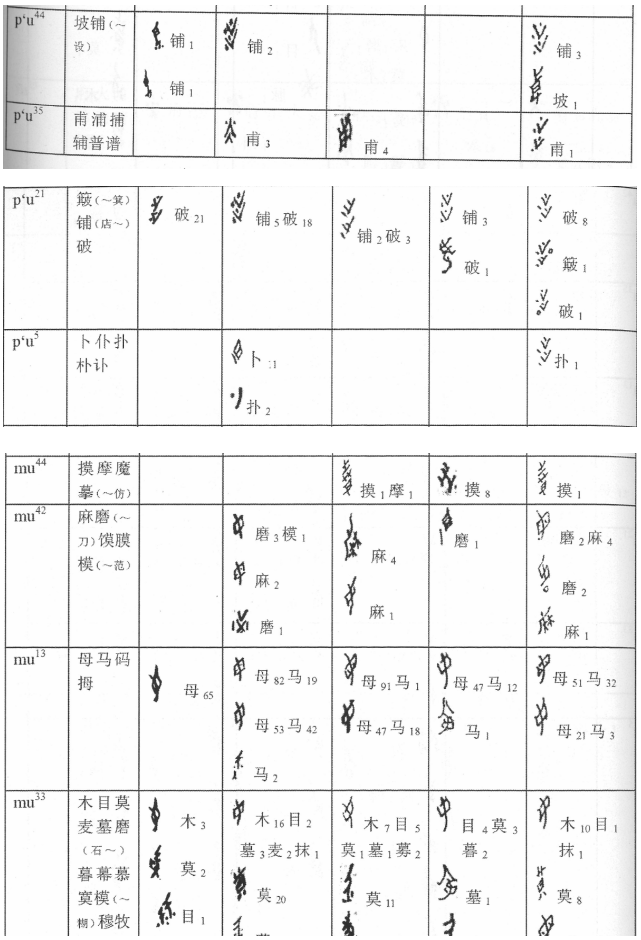


Figure 12: PHU and MU in Nushu Yongzi Bijiao by Zhao[14]

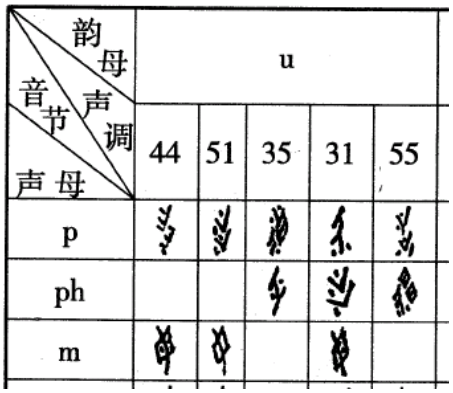


Figure 13: PU, PHU and MU in Nuzi Zidian by Xie[12]

For example, for the syllable PHU, Zhao finds PHU44, PHU35, PHU21, PHU5 (Figure 12), Xie finds PHU35, PHU31, PHU55 (Figure 13), and Zhou finds PHU44, PHU35, PHU21, PHU55 (Figure 14). The case of MU is worse; Zhao finds MU44, MU42, MU13, MU33, Xie finds MU44, MU51, MU31, Zhou finds MU41, MU21 and MU33. There is no crossing section at all. Of course, it is difficult to discuss the crossing section of the representative glyphs.

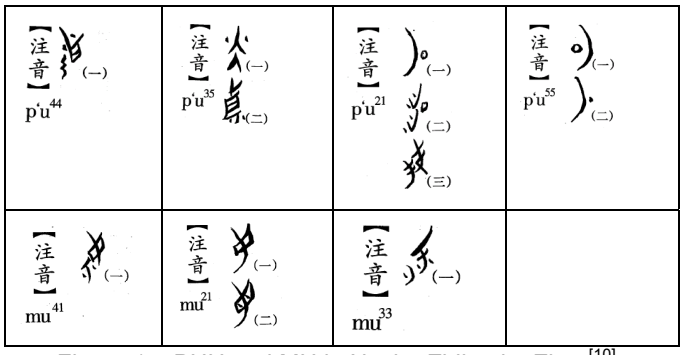


Figure 14: PHU and MU in Nushu Zidian by Zhou[10]

The number of the phonetic values used in Nushu documents is less than half of theoretically possible combinations. In Zhao's survey, 952 phonetic values are found for 5439 combinations. In Zhou's survey, 1094 phonetic values are found for 4080 combinations. In Xie's survey, 860 phonetic values are found for 3360 combinations. According to the survey of a few dialectic sound systems in Jiangyong Tuhua by Fangyan Diaocha Zibiao (方言調查字表), the distinctive syllables (including the difference of tones) in daily used words would be about from 500 to 1000.

Also the variety of the possible phonetic values for a single Nushu glyph is not easy to guess. According to Zhao's survey, the phonetic values for the glyph derived from "刀" is 21. The frequencies of the most frequent syllable TI13 (354) and of the 2nd frequent syllable TWE42/TIE42 (337) show no significant contrast. It means that using a phonetic value as the most dominant key introduces the difficulty; a user should guess the phonetic value for a given Nushu glyph, but the guessed phonetic values must be the most frequent. Considering that most dictionaries based on the phonetic values are word dictionary, not the character dictionary, the phonetic value might be useful but should not be used as the primal key in the collation, except for the users making Nushu document by the transliteration from Hanzi text to Nushu.

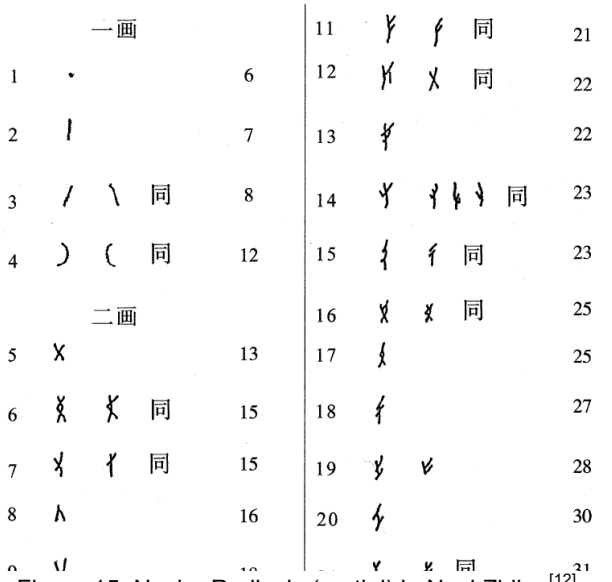


Figure 15: Nushu Radicals (partial) in Nuzi Zidian[12]

3.3 Collation via Radicals

Some Nushu characters are almost impossible to find a corresponding Hanzi and supposed to be independently created as pictograph. To classify the mixture of the Hanzi-derived glyph and Hanzi-independent glyph uniformly, some scholars proposed to use the radical system designed for Nushu. Xie proposed 64 radicals for Nushu.

The difficulty is how the appropriate radical could be determined. Because the phonetic loaning in Nushu is so wide that some scholars classify Nushu as the syllabic script, the radical for Nushu glyph could not be determined uniquely by their semantics. However, some radical classification rule in Nuzi Zidian seems to be influenced from Hanzi radical; when the Nushu glyphs are supposed to be derived from the Hanzi. Because Nushu radical set defined by Xie does not cover the frequently used Hanzi radicals in Hanzi-derived Nushu, the radicals determined by Nuzi Zidian is not intuitive for both of Hanzi users and

Nushu Glyph	Nushu Radical	Nushu Radical Name	Corresponding Hanzi	Hanzi Radical	Confusing Radical	Nushu Radical Name
	・部	点	守	宀	寸部	犁
	・部	点	官	宀	己部	己
	・部	点	肉	肉	分部	分
	丿部	竪	守	宀	守部	守
	丿部	竪	天	天	王部	王

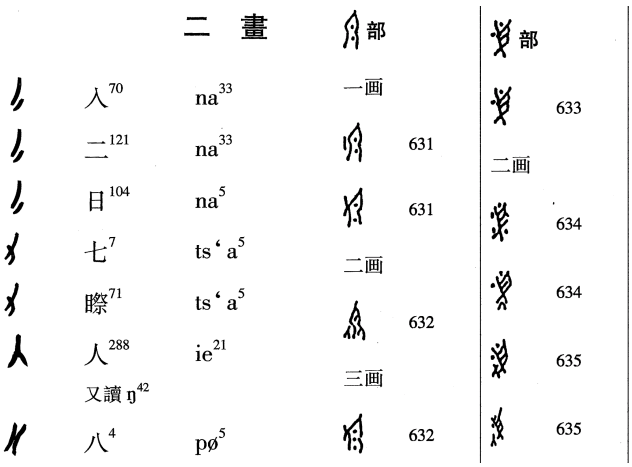
Figure 16: Example of Non-Intuitive Radical Classification in Nuzi Zidian[12]

3.4 Collation via Total Stroke Number

Another widely approved collation method is the classification by the total stroke number. As written in above, most Nushu characters have different sound values, so they are not stable to use as unique key. The easiest alternative property for Chinese scholars to categorize the characters in small groups would be the total stroke number. Zhao used the total stroke number as the primal key, the phonetic value as the second key, and corresponding Hanzi as the third key. Xie used the radical as the primal key, and the stroke count as the secondary key.

However, there was a long history between the commonalization of Hanzi writing and the stabilization of the stroke counting of Hanzi glyph; the regulation of the glyph shapes for the official documents in the civil service system. Before Jin dynasty, no Chinese dictionary ordered characters by their stroke counts. Because Nushu was mainly used for the communications with close friends or family sharing same experience, there was no serious motivation to regularize the glyph shapes; the glyphic differences by the number of dots, a circle or dot, the kink or curve contrast are often changed. Therefore, the selection of the representative glyph is difficult

problem.



Nushu character list in the appendix of Nushu Heji　　Nuzi Zidian Index

Figure 17: Examples of Stroke Count Collation in Nushu Heji and Nuzi Zidian

Orie Endo solved this problem by restricting the author. Zhao made the per-author statistic as the second step (her first step, the character list of Zhonghua Nushu Heji had not taken the author consideration), but her proposal to ISO/IEC 10646 mixes all statistics uniformly, so many inconsistencies are found in the selection of the representative glyphs.

4. Phonetic and Glyphic Collation

As summarized in previous section, the collation methods by corresponding Hanzi, phonetic value, total stroke count, and radicals have each advantages and disadvantages. It is impossible to determine the unique codepoint ordering in ISO/IEC 10646 by single method. The combinations of multiple methods should be discussed. It should be noted that the collation methods for ISO/IEC 10646 is identical with the index of the dictionary; it is acceptable for dictionary index to list same entity at the multiple position, but it is unacceptable for the collation method to determine the code position in ISO/IEC 10646.

The group size (the number of glyphs sharing same property value; same phonetic value, same total stroke count, same corresponding Hanzi, etc) for each collations could be summarized as follows.

- Semantics by corresponding Hanzi
  - Nushu Yongzi Bijiao (Zhao): 3129
  - Nushu Zidian (Zhou): 2663
- Syllables (consonant + vowel + tones)
  - Nushu Yongji Bijiao (Zhao): 952
  - Nushu Zidian (Zhou): 1094
  - Nuzi Zidian (Xie): 860
- Radical
  - Nuzi Zidian (Xie): 64
- Total Stroke Count
  - Nushu Duben (Zhao): 16
  - Nuzi Zidian (Xie): 20

The proposed character set by Zhao to ISO/IEC 10646 uses the total stroke count as the first collation key and the most frequently used phonetic value as the second collation key. As described in the section 3.2, the number of phonetic values found by Zhao's survey is 952. But the number of the characters in the proposed character set is 397. Considering that no simple rule is found for the phonetic values related to single Nushu glyph (see Figure 8), the phonetic collation may confuse the users in searching a glyph from the coded character set.

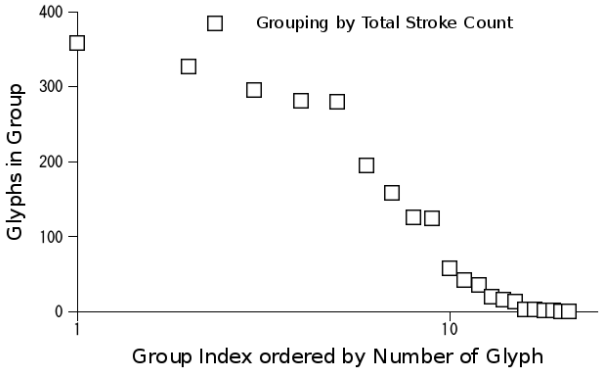


Figure 18: Glyph Group Size by Total Stroke Count

In addition, the number of the glyphs in same total stroke count is not sufficiently small for the users to take an overview. In the proposed character set by Zhao, the character group for the total stroke count 8 has 64 characters. This number reflects the number after the unification process, but the unification rule is not clarified in existing Zhao's research. If the users have to search the glyph in raw document before the unification process, how many variants they should consider as the candidates? In Figure 18, the glyph group size by the total stroke count is shown for Nuzi Zidian. In the largest glyph group (total stroke count is again 8), the group size is 358. It is almost impossible for the users to overview. It means that the clarification of the unification is essential.

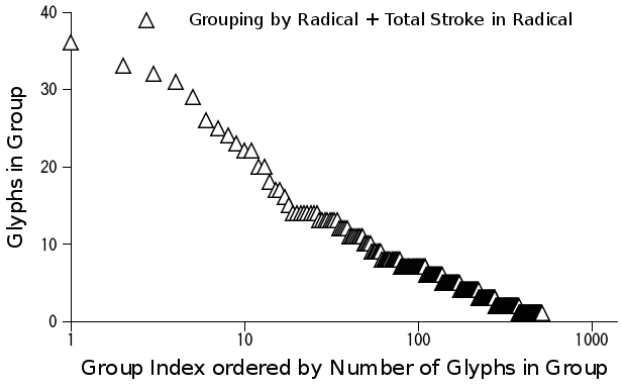


Figure 19: Glyph Group Size by Radical and Total Stroke Count Collation

Xie's collation method is the combination of the radical and the total stroke count. As shown in Figure 19, the combination could split 2328 glyph collection into 516 groups (excluding the

group without member), and all glyph group size is less than 40. If the clarification of the unification rule or the definition of the normalized glyph shape is difficult, the shape based classification by the radical and total stroke count would be considerable option

As most Hanzi dictionaries, Xie's Nuzi Zidian uses the radical as the first key and the stroke count as the second key. However, as Figure 16 shows the difficulty to assign the most reasonable radical to a given Nushu glyph. Also the collection of the radical is not stabilized yet. For example, Chen Qiguang analyzed the structure of Nushu glyphs derived from Hanzi and found several glyphic components corresponding to the source Hanzi. However, Nuhan Zidian by Chen uses the total stroke count as the first stroke. Considering that some of the glyphic components are not included in Xie's radical system, or identified differently, the set of the radical should not be strictly included in the part of ISO/IEC 10646.

Although the stroke counting for Nushu glyph is not always intuitive, the counting method is stabilized; all of Zhao Xie and Chen works define same method. Therefore, in the current standardization, the best property for the first collation key would be the total stroke count, and the second

5. Summary

In this report, the problems of Nushu standardization draft are described. The first problem is that Nushu is not sufficiently studied to define the standard character set; the definition of the entity to be coded is still arguable, and there is no authorized dictionary. There are several Nushu dictionaries, but most of them are compiled independently and by incompatible methodologies. Even the comparison of their contents is not studied sufficiently; it is difficult to define the entity independent with specific dictionary. The second problem is that the proposal is not based on the neutral or stable cross sections of existing studies, but on the studies by the scholars interested in ISO standardization.

However, as far as PRC national body proposed the standardization officially and the block to encode Nushu characters is already allocated, it would be difficult to form a consensus to cancel the standardization itself. To minimize the impact of the standardization based on the insufficient studies, the collation method should be carefully chosen to minimize the unneeded distinctions in ISO/IEC 10646. In this report, it is discussed that the collation should not include the non-glyphic properties. The stroke counting is already stabilized method, but it is insufficient to split Nushu glyphs into the groups with the size that the users can overview. There is a study to define the radical for Nushu, but the radical collection is not stabilized yet. The discussion of the radicals for the core set of Nushu glyphs is needed before the standardization

Acknowledgement

This study is supported by KAKENHI 24500116, 26330377.

Special thanks to Orie Endo, Liu Ying, Deborah Anderson for detailed discussion.

References

[1] China National Body: "Proposal for encoding Nushu in the SMP of the UCS", 2007-08-13, ISO/IEC JTC1/SC2/WG2 N3287

[2] China National Body: "Reply to WG2N3322 - UC Berkeley's comments on Nushu", 2007-09-16, ISO/IEC JTC1/SC2/WG2 N3337

[3] China National Body: "Proposal for encoding Nushu in the SMP of the UCS", 2009-03-18, ISO/IEC JTC1/SC2/WG2 N3598

[4] Orie Endo: "Comments and Questions on N3598", 2009-09-05, ISO/IEC JTC1/SC2/WG2 N3705

[5] China National Body: "Updated Proposal for encoding Nyushu in the SMP of the UCS", 2012-09-10, ISO/IEC JTC1/SC2/WG2 N4341

[6] ISO/IEC JTC1/SC2/WG2: "Proposed Draft of Amendment (PDAM) 1 to ISO/IEC 10646:2014", 2013-09-27, ISO/IEC JTC1/SC2/WG2 N4484

[7] suzuki toshiya: "Comments on Nushu in WG2 N4484", 2013-12-25, ISO/IEC JTC1/SC2/WG2 N4513

[8] suzuki toshiya, Orie Endo: "Additional Comments on N4484 /PDAM1", 2014-Jan-30, ISO/IEC JTC1/SC2/WG2 N4533.

[9] Nushu ad hoc committee: "Nushu ad hoc report", 2014-02-25, ISO/IEC JTC1/SC2/WG2 N4561

[10] 宮哲兵 (Gong Zhebing): "關於一種特殊文字的調查報告—湘南瑤山采風記", 中南民族学院学報, 1983, Volume 3, p.122-128.

[11] 周碩沂 (Zhou Shouyi): "女書字典 (Nushu Zidian)", 兵麓書社, 2002, ISBN 978-7-80665-234.

[12] 謝志民 (Xie Zhming): "中国女字字典 (Zhongguo Nuzi Zidian)", 民族出版社, 2009, ISBN 978-7-105-09963-4

[13] 趙麗明 (Zhao Liming): "中国女書合集 (Zhongguo Nushu Heji", 中華書局, 2005, ISBN 7-101-04405-0

[14] 趙麗明 (Zhao Liming): "女書用字比較 (Nushu Yongzi Bijiao)", 2006, 知識産権出版社, 2006, ISBN 978-7-8019-8261-2

[15] 趙麗明 (Zhao Liming): "女書読本 (Nushu Duben)", 2009, 湖南人民出版社, ISBN 978-7-5438-5282-2

[16] 遠藤織枝 (Orie Endo): "中国女文字研究", 明治書院, 2002, ISBN 462548300X

[17] 陳其光: "女漢字典 (Nuhan Zidian)", 2006, 中央民族大学出版社, ISBN 978-7810569866

[18] 陳其光: "女字の形、音、義、源", 論文集"消えゆく文字 中国女文字の世界", 2009, ISBN 978-4-88303-251-8, p.51-65.

[19] 孫琦: "女書文字リストの中の異体字", 論文集"消えゆく文字 中国女文字の世界", 2009, ISBN 978-4-88303-251-8, p.119-129.

[20] 陳力衛: "女書が衰退した原因について", 論文集"消えゆく文字 中国女文字の世界", 2009, ISBN 978-4-88303-251-8, p.73-84.