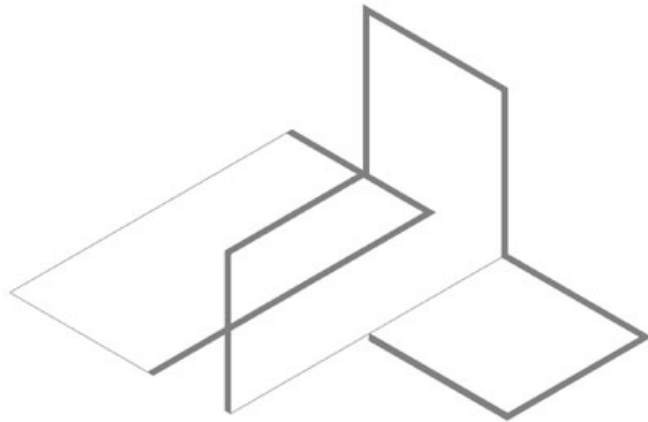


# Proposal for encoding Nüshu in the SMP of the UCS



中 國 女 書 文 化 研 究 中 心  
nüshu culture research center of china

**Source :China**

**Date:** 2007-9-17

Email :Zhaoml@mail.tsinghua.edu.cn  
kalisiqi@gmail.com



# Introduction

---

A kind of rural women-only Character Set

- Area
- Character shape
- Nature
- Character number
- Other features
- value



Xiaoshui Valley of Jiangyong, Hunan Province, China.

Two character systems



# area

---

- Nüshu originates from Chinese characters.
  - a variant form of Kaishu (a regular Chinese script).
- It appears to be italics, like Chinese character "多".
- spelling order is from top-right to bottom-left.

There are three kinds of situations for its origin from Chinese character.

[see the relationship with Chinese characters](#)

# Relationship with Chinese character

1. Borrow the whole character body, and change it from squared-shape to italic

e.g. 甲 (甲) 如 (如) 主 (主) 良 (良) 内 (内)

- ✓ some of them are reversed from Chinese character and become italic.

e.g. 义 (义) 己 (己) 春 (春) 吞 (吞) 天 (天)

- ✓ some are changed italic from simplified character

e.g. 后 (后) 个 (个) 来 (来) 归 (归) 听 (听) 双 (双)

# Relationship with Chinese character

2. some changes occur to the main body of Chinese character, but it is obvious to keep its Chinese shape.

e.g.  (借)  (落)  (白)  (苦)  (母)  (春)  (昔)  (左)

■ Some changes look more obvious than the above, but still kept Chinese trail.

e.g.  (羊)  (光)  (焦)  (明)  (去)  (别)  (心) .

3. sound signs and character-cluster: Some clusters contain more than 10 characters.

e.g.

 (声) → 、、、、



# Nature

---

- syllable character.
  - i.e. a character represents an independent syllable that records a group of homophones or characters with similar pronunciation in their local dialect.



# Character number

---

- As a mature syllable word system, less than 400 basic characters are enough to record and represent more than 400 syllables (about 1200 with tones included ) in the local Chenguan dialect.
- due to individual habit, every Nüshu elderly women prefer certain characters to others. So allograph comes. But the number of all signs is no more than 600.

There will be a further argumentation about the basic characters and allograph.





# Other features

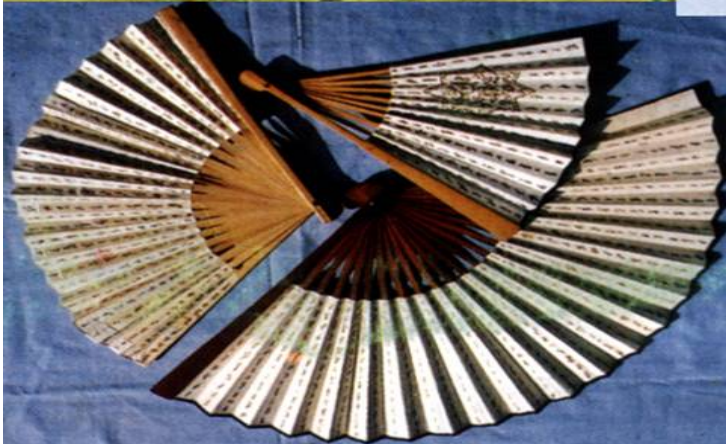
---

- free、 open、 natural and positive.
- Seven-character songbook is the main form besides few five-character poems.  
Autobiographies, letters, folk songs, monody, narration and translation constitute its most contents
- cloth covers, scrip, fans, cloth-handkerchiefs and flowered belt are available to write on.

# Poems on the fan



# Writing material



2007-9-17

Nushu

11



# Process description

---

- the collection and collation of Nüshu characters
- the judgment and collation of Nüshu basic and variant characters
- Code object
- order rules



# the collection and collation

---

- *Chinese Nüshu works collection, 2005 :*  
more than 90% original ecological data.
- *Jiangyong dialect research, 1993*
- We identified every Nüshu character of the book by finding its corresponding Chinese character, represented its sound and set a character table respectively in the order of strokes and phonetics.



# Chinese Nüshu works collection





# Basic characters and allograph

---

- Basis a: *character-cell theory*
  - a. with the same origin
  - b. with the same configuration methods
  - c. with the same pronunciation
  - d. with no distinction in the meaning. (substitutable )
- Basis b: *character frequency*
- Characters complied with the above features but with different graphs are regarded as allograph. The character with the highest frequency is called the **basic** and others as **allograph**.

referring to '*Nüshu basic characters and their origin verification*', 2006



# Code object

---

- characters with different origin are regarded as different ones and given respective code.

e.g. 我 (𠂇 𠂇) (occasionally 𠂇, 𠂇 which comes from  
五) ; 一 (一, 𠂇)





# Code object

---

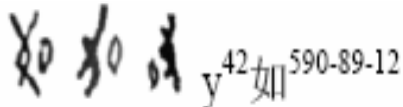
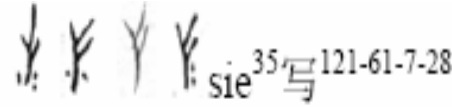
- characters with tiny stroke distinction (such as the distinction between linked stroke and broken one) are seen as one graph, and the one that has the highest frequency is representative and coded.

E.g.  niang<sup>42</sup> 娘<sup>1586-334-17</sup>,  k'w<sup>5</sup> 刻<sup>132-3</sup>



# Code object

---

- characters with some distinction of minor part such as its position, quantity and shape but the same main structure and configuration contour are regarded different, and coded separately. But the character with lower frequency is regarded as allograph.
- E.g.  

# Example

如 如 如

如 590-89-12


y <sup>42</sup>	如余餘 儒虞娛	如 105 如 10 如 8	如 131 余 10 如 54 如 2 儒 1	如 213 余 2 如 3 如 1	如 75 余 3 如 11 如 11 如 4	如 141 余 3 如 11
-----------------	------------	----------------------	----------------------------------	-------------------------	---------------------------------	-------------------




# Code object

---

- sometimes characters with similar component and configuration are spelled confusedly. They are coded as allograph.

e.g.  tou<sup>33</sup>度<sup>38-37</sup>, tu<sup>33</sup>独<sup>76-71</sup>

- later-created :not in accordance with Nüshu spelling rules. They are later-created subjectively and not coded. 



# Order rules

---

- ◆ The basic characters are coded first and the allograph following.

449 characters are coded all-together.

basic characters: 385, 1B000..1B180

allographs: 64, 1B181..1B1C1

Order standard: strokes and sound.



# name rules

---

- we can use the Chinese naming rule to name Nüshu character. Its name is the corresponding code sequence number: Nüshu-1BXXX.



# Supplementary Note

---

- the font of Nüshu characters is made by *Weifang Beida Jade Bird Huaguang Imagesetter Co., Ltd* which is famous and professional in China. And the codes of characters begin from 1B000.



# Prospects

---

- a new category to various characters;
- Has great academic value to history, linguistics, folklore and literature;
- of great impetus to the research of female culture and information processing;
- made good use of to modern mind cure. and has a universal meaning in the world.











---

- THANK YOU!

- WELCOME YOUR QUESTION!