ISO/IEC JTC1/SC2/**WG2 N4967**
2018-06-01

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type:	**Working Group Document**
Title:	**Updated proposal on Gongche characters for Kunqu Opera**
Source:	**Eiso Chan(陈永聪), Jerry You(游程宇),** **Wang Xiaolei(王晓磊), Hou Le(侯乐)**
Status:	**Individual Contribution to WG2 #67, London, UK**
Action:	**For consideration by JTC1/SC2/WG2 and UTC**
Date:	**2018-06-01**
References:	**L2/17-087, L2/17-103, L2/18-189, IRGN2299, IRGN2290, IRGN2293**

We submitted the proposal called "*Proposal to encode characters for Gongche Notation in SMP*" (aka L2/17-087) to UTC last year. That document had been discussed in UTC #151, please see 151-A79 and 151-A80 in the minutes (aka L2/17-103).

> **E.5 Proposal to encode characters for Gongche Notation in SMP [Chan, et al, L2/17-087]**
>
> **[151-A79] Action Item for** John Jenkins: Add the seven Gongche characters in L2/17-087 with a "tiny slash downwards" to UAX #45.
>
> **[151-A80] Action Item for** Ken Lunde: Respond to submitter of L2/17-087 with UTC disposition.

For 151-A79, John has added these seven characters to UAX #45 as UTC-03160 through UTC-03166, and they will be included in the Version 11. Please also see Section 4 in this document. And then, Eiso submitted the document called "*Request to discuss how to handle seven unencoded Gongche characters for Kunqu Opera*" (aka IRGN2299) to IRG. IRGN2299 has been discussed in IRG #50 in Beijing, China. After the discussion, IRG thought these seven Gongche characters were not suitable to encode as the CJK Unified Ideographs. Please see IRG M50.8 in IRGN2290 (aka L2/18-189) and Section 13 in IRGN2293.

> **[IRGN2290]**
>
> **Recommendation IRG M50.8: Encoding of Gongche Characters (IRGN2299, IRGN2293)**
>
> **Unanimous**
>
> IRG reviewed IRGN2299 and considers Gongche symbols are for music notations and should not be included in CJK unified ideograph blocks (IRGN2293).

[IRGN2293]

13. Encoding of Gongche Characters

Reference:

- IRGN2299: Request to discuss how to handle seven unencoded Gongche characters for Kunqu Opera

The editors acknowledged the needs of these characters. The editor considered these characters to be symbols for musical notations and should not be in the CJK Unified Ideographs blocks.

This document provides the explanations and evidences of seven unencoded Gongche characters. As we know, the differences between these seven characters and the common ideographs are obvious, but there are so many Gongche characters which has been included in ISO/IEC 10646 and Unicode as CJK Unified Ideographs.

I request WG2 and UTC to encode these seven Gongche characters separately as non-CJKUI and create a new block called **Gongche**, if IRG's recommendation M50.8 been accepted.

On the other hand, there are also seven characters which are "合", "四", "一", "上", "尺", "工", "凡" with double tiny slashes downward shown in Figs. 3 & 4. Mr. Wang Xiaolei, Ms. Hou Le and I have checked so many books on Gongche notation. We can confirm those characters are uncommon, so there is no need to encode them at present.

Those seven characters which are "合", "四", "一", "上", "尺", "工", "凡" with double tiny slashes downward are theoretical Gongche characters. If any expert concerns with those ones, we think the better way is to make the **Gongche** block contain 32 characters.

1. Introduction

Gongche notation (工尺谱/工尺譜) is a traditional musical notation method, once popular in ancient China, and now it's also used in Chinese local operas, especially Kunqu Opera (昆曲/崑曲), Peking Opera (京剧/京劇), Yueju Opera (粤剧/粵劇) and so on. It uses the characters like Chinese ideographs to represent musical notes. It was named after two of the Chinese ideographs that were used to represent musical notes, namely "工" gōng and "尺" chě.

Many characters for Gongche Notation have the same appearances with Chinese ideographs exactly, so the vast majority of them could be used isomorphic ideographs to indicate. Hong Kong SARG once submitted two ideographs which are just used in the lyrics of Yueju Opera as UNCs in IRGN1405R. Please see Fig. 1. And UTC has submitted other two ideographs which are used in the lyrics of Kunqu Opera and traditional Gongche Notation to WS2017. Please see Fig. 2.

Fig. 1 H-87D2 & H-87D6 in IRGN1405R

2	伬	H-87D2	0365.211	彳	4	⿰彳尺	Chinese character used in the lyrics of Cantonese opera (粵曲曲譜字符). Sample is shown in Annex III. Cantonese opera is regarded as a cultural heritage of the People's Republic of China (seehttp://www.gov.cn/zwgk/2006-06/02/content_297946.htm).
3	伍	H-87D6	0365.211	彳	4	⿰彳五	- Ditto -

Fig. 2 UTC-02976(01240) & UTC-02977(01242) in WS2017

60.0 彳 ⿰彳上 3 6						仩 UTC-02976
60.0 彳 ⿰彳凡 3 6						仉 UTC-02977

However, some characters' appearances do not conform to the general principles of writing Chinese ideographs. The end strokes of "合", "四", "一", "上", "尺", "工", "凡" are extended by a tiny slash downward.

According to Fig. 3, when these 7 characters and UTC-02976, UTC-02977 would be included in ISO/IEC 10646 and Unicode, we can get the table for Gongche Notation as below.

Table 1 Characters for Gongche Notation

N/A	N/A	N/A	N/A	合	四	一
1̤	2̤	3̤	4̤	5̤	6̤	7̤
上	尺	工	凡	合	四	一
1̣	2̣	3̣	4̣	5̣	6̣	7̣
上	尺	工	凡	六	五	乙
1	2	3	4	5	6	7
仩	伬	仜	仉	伬	伍	亿
1̇	2̇	3̇	4̇	5̇	6̇	7̇
仩	伬	仜	仉	伬	伍	亿
1̈	2̈	3̈	4̈	5̈	6̈	7̈

After the discussion in IRG #50, it looks the better way to encode them is to treat them as the non-CJKUIs. If WG2 and UTC agree with the recommendation IRG M50.8, the new block should be created and include these seven Gongche characters in it. Please see or review Sections 2 and 3.

The other way is still to encode them as CJKUI, but IRG didn't think it was really a good way for this issue. Please see Sections 4 and 5.

2. Proposed Characters as Non-Ideographs

The ordering of these seven Gongche characters is based on the original natural ordering of the musical scale.

(合=5̇, 四=6̇, 一=7̇, 上=1̇, 尺=2̇, 工=3̇, 凡=4̇)

There is a character like kana の(U+306E) and a character like kana の(U+306E) by a tiny slash downward in the evidences, but, in fact, they are same with 四(U+56DB) and 四(U+56DB) by a tiny slash downward. Kana の(U+306E) looks like cursive form of 四(U+56DB). Notice that the kana の is from 乃(U+4E43) not 四(U+56DB), and the corresponding hentaigana letter is 乃(U+1B099).

We can know the variant form for 凡(U+51E1) is 凡(U+51E2) by Unihan Database, and some evidences show the glyph is 凡(U+51E2) by a tiny slash downward. According to Fig. 3, 凡(U+51E2) by a tiny slash downward and 凡(U+51E1) by a tiny slash downward are the same character.

These two characters should be normalized to be "四" and "凡". Please see Table 4 as well.

一 can't be unified with 一(U+31D6) or 一(U+4E5B).

Table 2 Proposed Gongche Characters

UCS	Representative Glyphs	Proposed Character Name	Evidence
U+19BF0		GONGCHE CHARACTER HE BY A TINY SLASH DOWNWARD	Fig. 3, 4, 28
U+19BF1		GONGCHE CHARACTER SI BY A TINY SLASH DOWNWARD	Figs. 3, 4, 15, 22, 23
U+19BF2		GONGCHE CHARACTER YI BY A TINY SLASH DOWNWARD	Figs. 3, 4, 17, 21
U+19BF3		GONGCHE CHARACTER SHANG BY A TINY SLASH DOWNWARD	Figs. 3, 4, 5, 7, 9, 11, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 27, 29
U+19BF4		GONGCHE CHARACTER CHE BY A TINY SLASH DOWNWARD	Figs. 3, 4, 5, 7, 8, 9, 11, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 27, 28, 29
U+19BF5		GONGCHE CHARACTER GONG BY A TINY SLASH DOWNWARD	Figs. 3 through 29
U+19BF6		GONGCHE CHARACTER FAN BY A TINY SLASH DOWNWARD	Figs. 3, 4, 7, 8, 11, 12, 13, 17, 21, 24, 28, 29

3. Unicode Properties as Non-Ideographs

```
19BF0;GONGCHE CHARACTER HE BY A TINY SLASH
DOWNWARD;Lo;0;L;;;;;N;;;;;

19BF1;GONGCHE CHARACTER SI BY A TINY SLASH
DOWNWARD;Lo;0;L;;;;;N;;;;;

19BF2;GONGCHE CHARACTER YI BY A TINY SLASH
DOWNWARD;Lo;0;L;;;;;N;;;;;

19BF3;GONGCHE CHARACTER SHANG BY A TINY SLASH
DOWNWARD;Lo;0;L;;;;;N;;;;;

19BF4;GONGCHE CHARACTER CHE BY A TINY SLASH
DOWNWARD;Lo;0;L;;;;;N;;;;;

19BF5;GONGCHE CHARACTER GONG BY A TINY SLASH
DOWNWARD;Lo;0;L;;;;;N;;;;;

19BF6;GONGCHE CHARACTER FAN BY A TINY SLASH
DOWNWARD;Lo;0;L;;;;;N;;;;;

Script=Common
```

4. Proposed Characters as Ideographs

The following is the information for these seven characters in UAX #45.

UTC-03160;X;U+5408;30.3;0174.241;合;UTCDoc L2/17-087

UTC-03161;X;U+56DB;31.2;0216.221;四;UTCDoc L2/17-087

UTC-03162;X;U+4E00;1.0;0075.011;一;UTCDoc L2/17-087

UTC-03163;X;U+4E0A;1.2;0076.071;上;UTCDoc L2/17-087

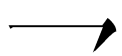
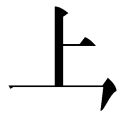
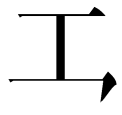
UTC-03164;X;U+5C3A;44.1;0299.321;尺;UTCDoc L2/17-087

UTC-03165;X;U+5DE5;48.0;0325.011;工;UTCDoc L2/17-087

UTC-03166;X;U+51E1;16.1;0134.021;凡;UTCDoc L2/17-087

And the following table shows their references, glyphs, IDSes and so on.

Table 3 Attributes for UTC-03160 through UTC-03166

U Ref.	Glyph	IDS	RS	FS	T/S	Var.	Evidence
UTC-03160		合	30.3	3	0	N/A	Fig. 3, 4, 28
UTC-03161		四	31.2	3	0	N/A	Figs. 3, 4, 15, 22, 23
UTC-03162		一	1.0	1	0	N/A	Figs. 3, 4, 17, 21
UTC-03163		上	1.2	2	0	N/A	Figs. 3, 4, 5, 7, 9, 11, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 27, 29
UTC-03164		尺	44.1	4	0	N/A	Figs. 3, 4, 5, 7, 8, 9, 11, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 27, 28, 29
UTC-03165		工	48.0	1	0	N/A	Figs. 3 through 29
UTC-03166		凡	16.1	4	0	N/A	Figs. 3, 4, 7, 8, 11, 12, 13, 17, 21, 24, 28, 29

5. Unicode Properties as Ideographs

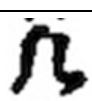
The properties for these seven characters are showed as below if they been treated as CJKUIs. “3xxxx” means the code point(s) in the future extension.

```
3xxxx;CJK UNIFIED IDEOGRAPH-3xxxx;Lo;0;L;;;;;N;;;;;

Script=Han
```

6. Gongche Characters Mapping

Table 4 Unified Gongche Characters Mapping

UCS	Glyph 1	Glyph 2	Zhao	Chen	WLi	WLiu	Yu	Gu	Zheng	BK
19BF0					absent	absent	absent	absent		absent
19BF1					absent		absent		absent	absent
19BF2					absent			absent	absent	absent
19BF3									absent	
19BF4										
19BF5										
19BF6										

Zhao is for 赵晓楠, Chen is for 陈根方, WLi is for 王锡纯 & 李秀云, WLiu is for 王季烈 & 刘富梁, Yu is for 俞振飞, Gu is for 顾铁华, Zheng is for 郑剑西, BK is for 《中国百科大词典》. Please see the full name of these books in Section 8.

7. Evidences

Fig. 3

赵晓楠: 《工尺谱常识与视唱》, 北京:人民音乐出版社 (Beijing: People's Music Publishing House), 2014, ISBN 978-7-103-04558-9, P. 3

工尺谱字	上	尺	工	凡	合	四	一
简谱音高	1	2	3	4	5	6	7
工尺谱字	上	尺	工	凡	合	四	一
简谱音高	1	2	3	4	5	6	7
工尺谱字	上	尺	工	凡	六	五	乙
简谱音高	1	2	3	4	5	6	7
工尺谱字	仩	尺	工	凡	仜	伍	亿
简谱音高	1	2	3	4	5	6	7
工尺谱字	仩	尺	工	凡	仜	伍	亿
简谱音高	1	2	3	4	5	6	7

Fig. 4

陈根方: 《中国工尺谱的数字实现研究》 (*Research on Digitization of Chinese Gongche Notation*), 上海大学博士学位论文 (Shanghai University Doctoral Dissertation), 2011, P. 6

降低八度的方法就是在每个基本谱字的最后一个笔划上加向左下方向的尾巴笔画，降低一个八度加一个尾巴笔画，降低两个八度加两个尾巴笔画，尾巴笔画都和谱字的最后一个笔画相连。如表 1.3 所示。

表 1.3 基本谱字和降低一个和二个八度后的谱字

基本谱字	合	四	一	上	尺	工	凡
降低一个八度谱字	合	四	一	上	尺	工	凡
降低二个八度谱字	合	四	一	上	尺	工	凡

Figs. 5 through 6

王锡纯，李秀云：《遏云阁曲谱·牡丹亭》，遏云阁，1870, Folio 1

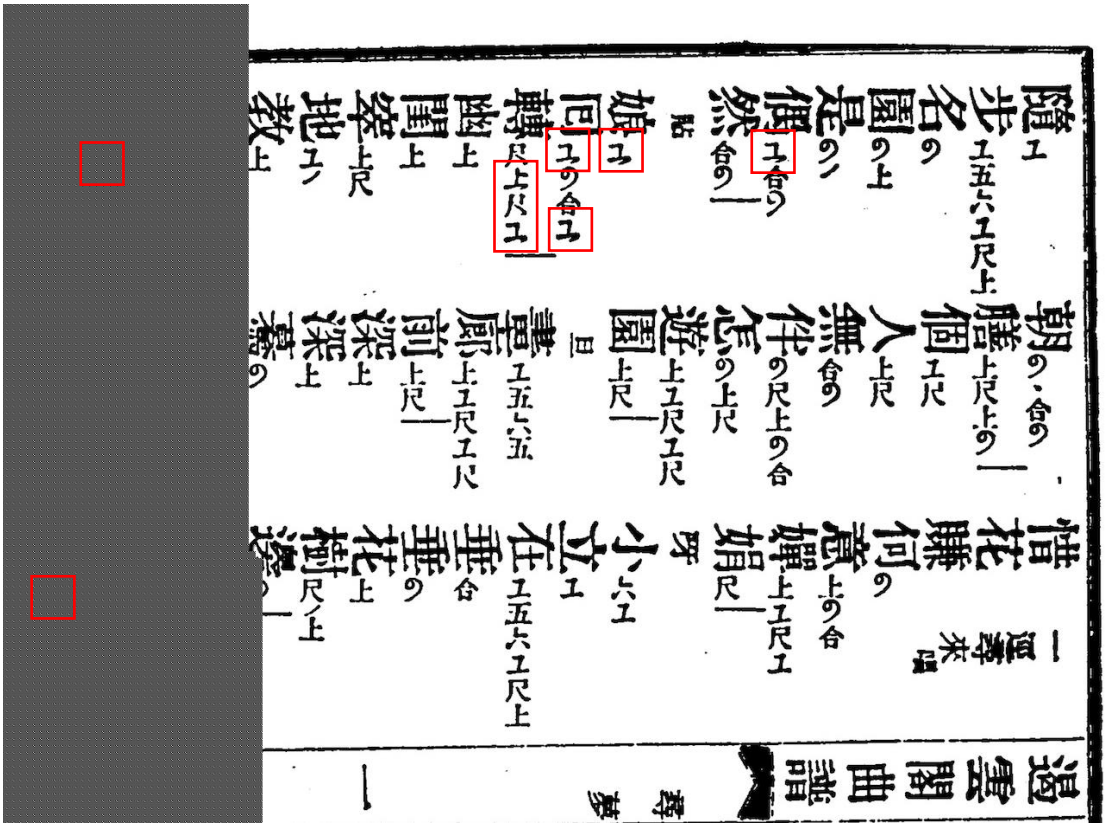


Fig. 19

俞振飞: 《粟庐曲谱·牡丹亭》, 上海: 上海辞书出版社, 1953, 2011 reprint, Folio 141

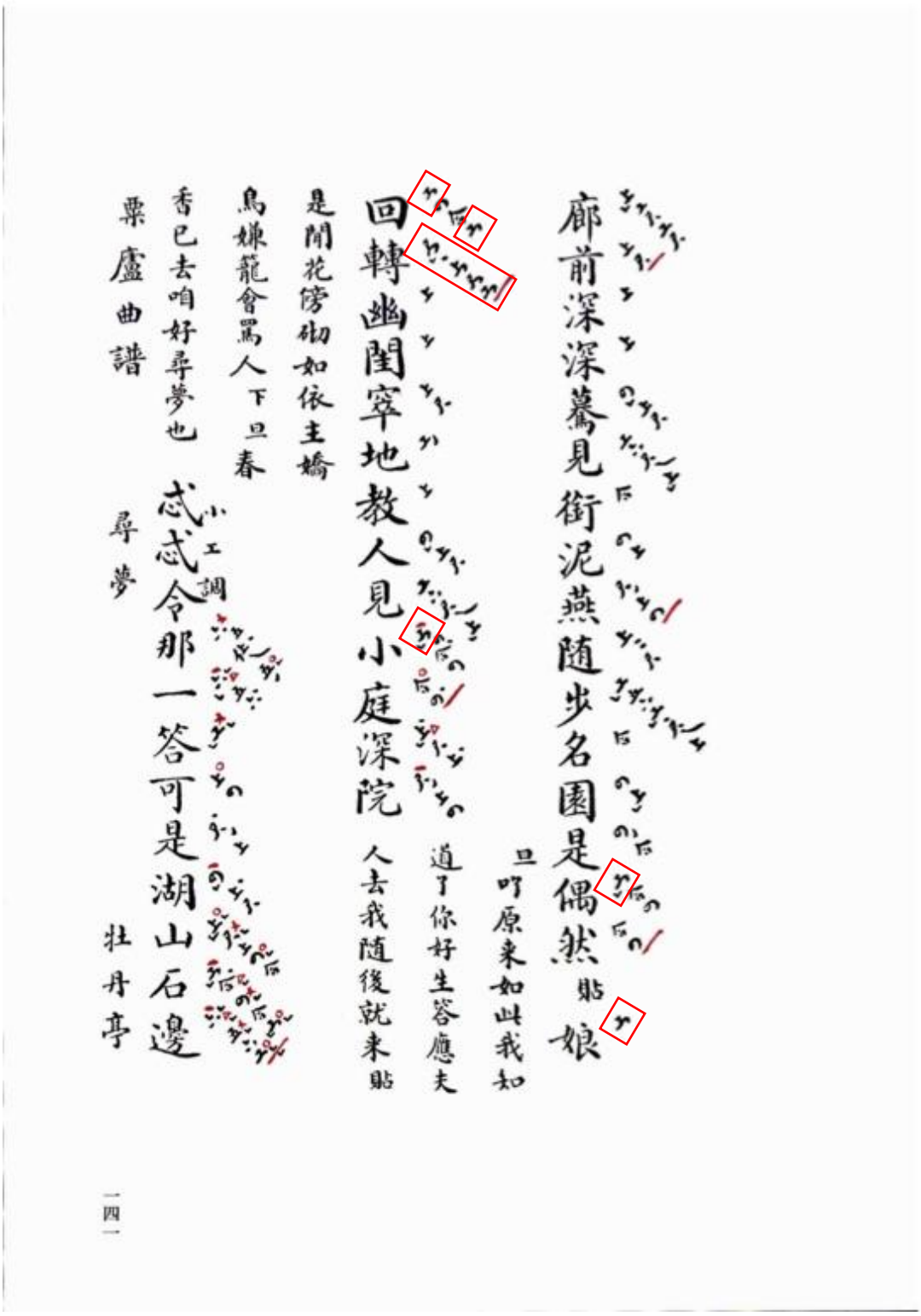


Fig. 20

俞振飞：《粟庐曲谱·西楼记》，上海：上海辞书出版社, 1953, 2011 reprint, Folio 262

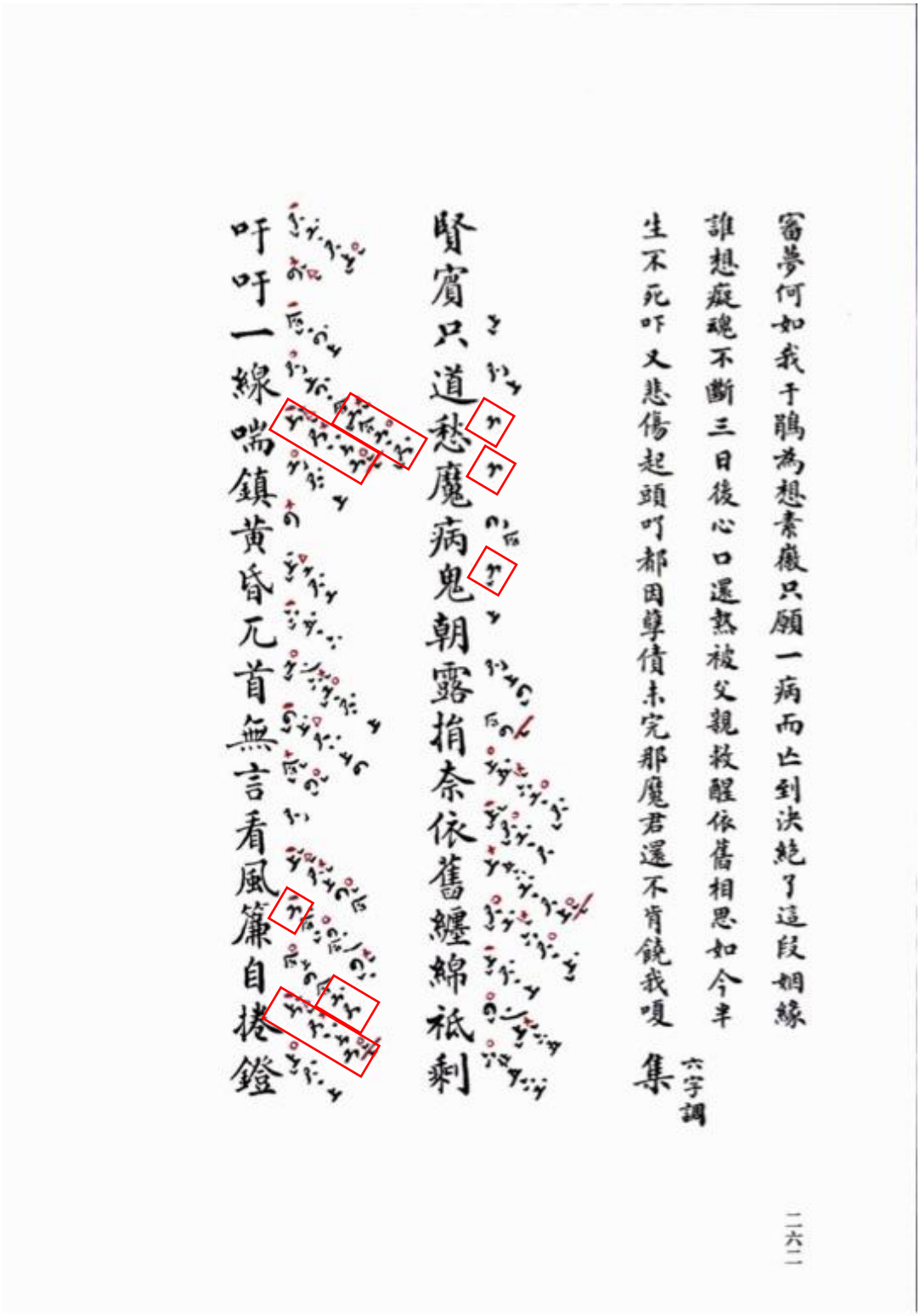


Fig. 21

俞振飞: 《粟庐曲谱·长生殿》, 上海: 上海辞书出版社, 1953, 2011 reprint, Folio 285

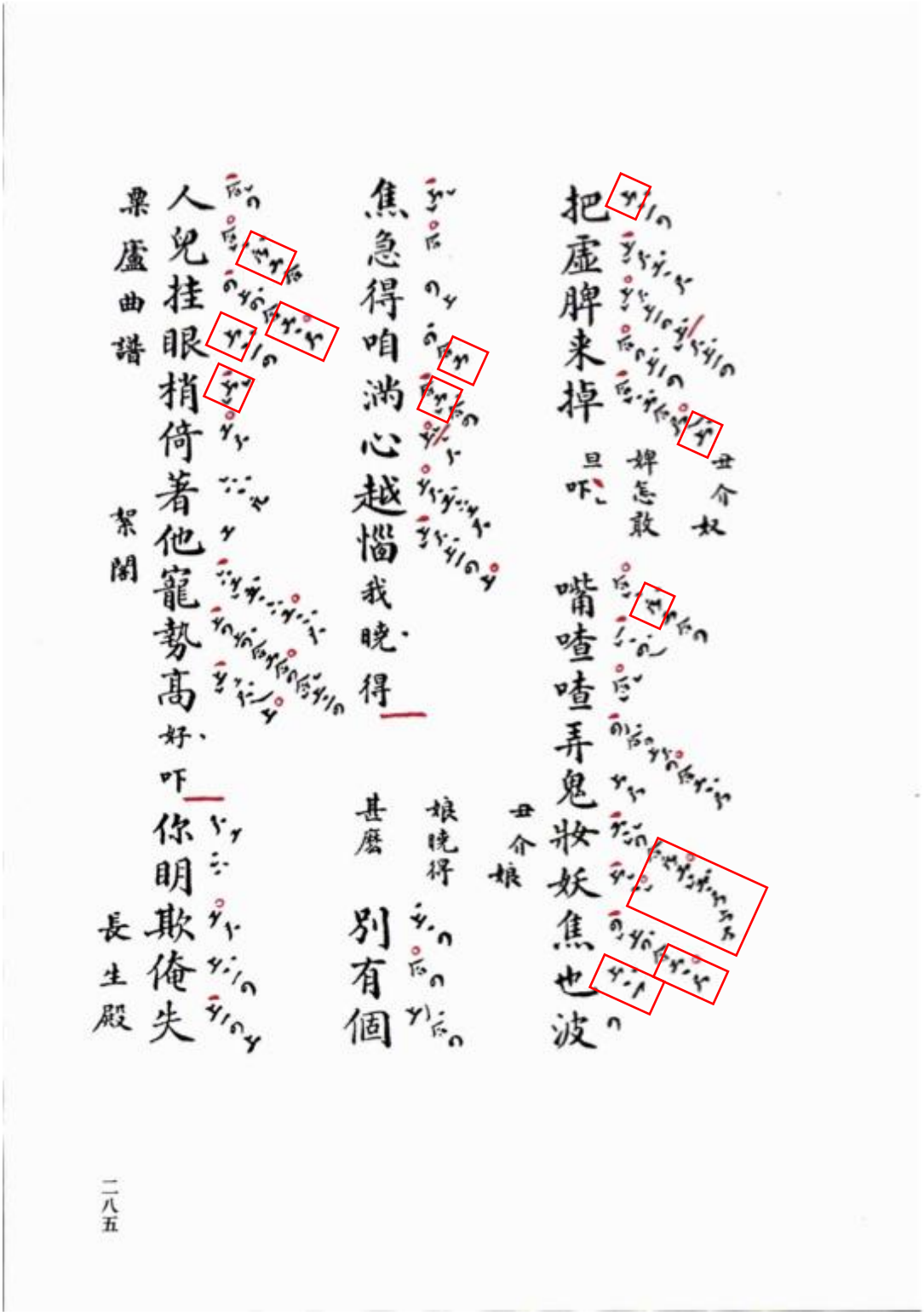


Fig. 22

顾铁华: 《粟庐曲谱外编·浣纱记》, Self Printed, 2002, Folio 156

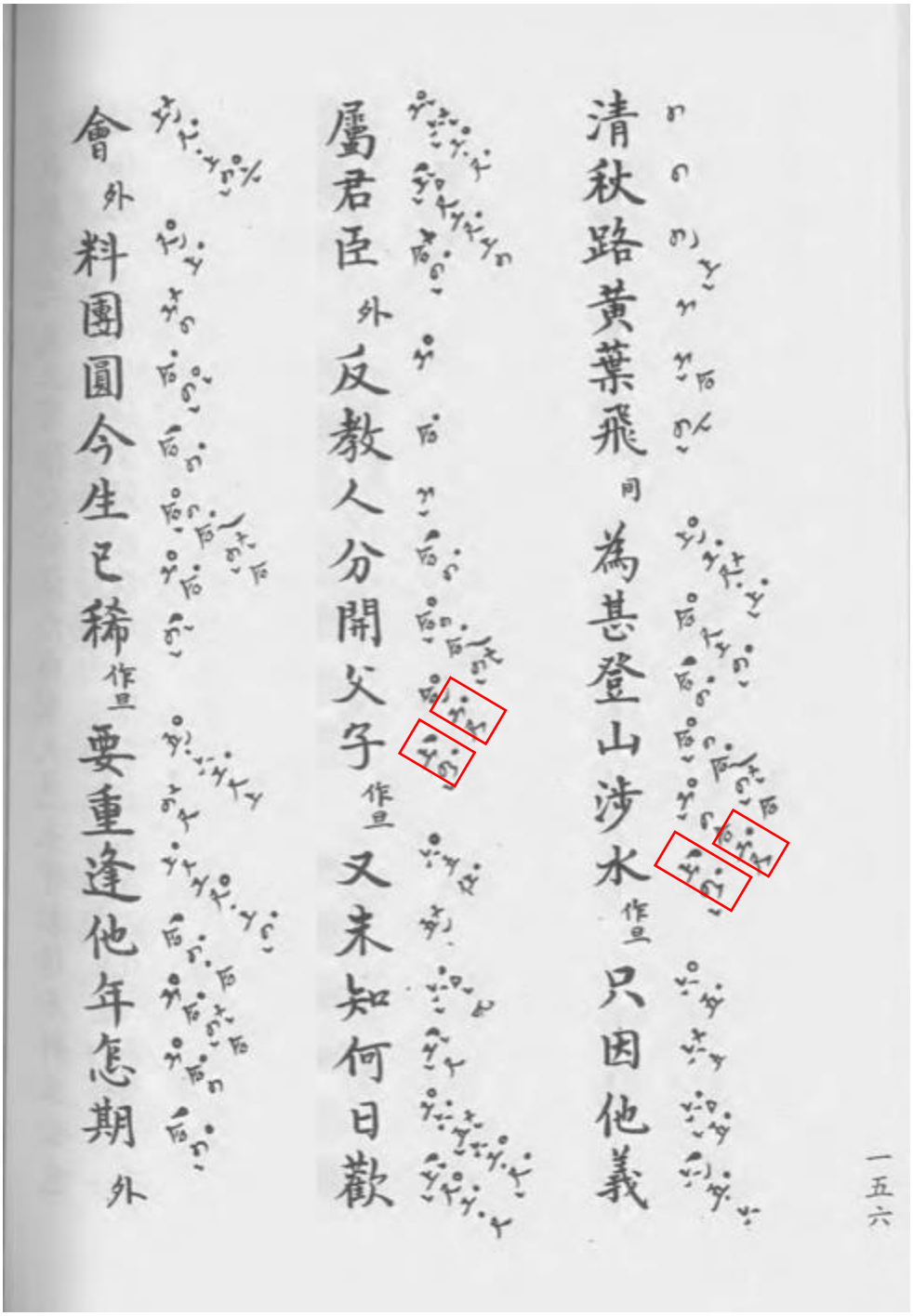


Fig. 23

顾铁华: 《粟庐曲谱外编·浣纱记》, Self Printed, 2002, Folio 157

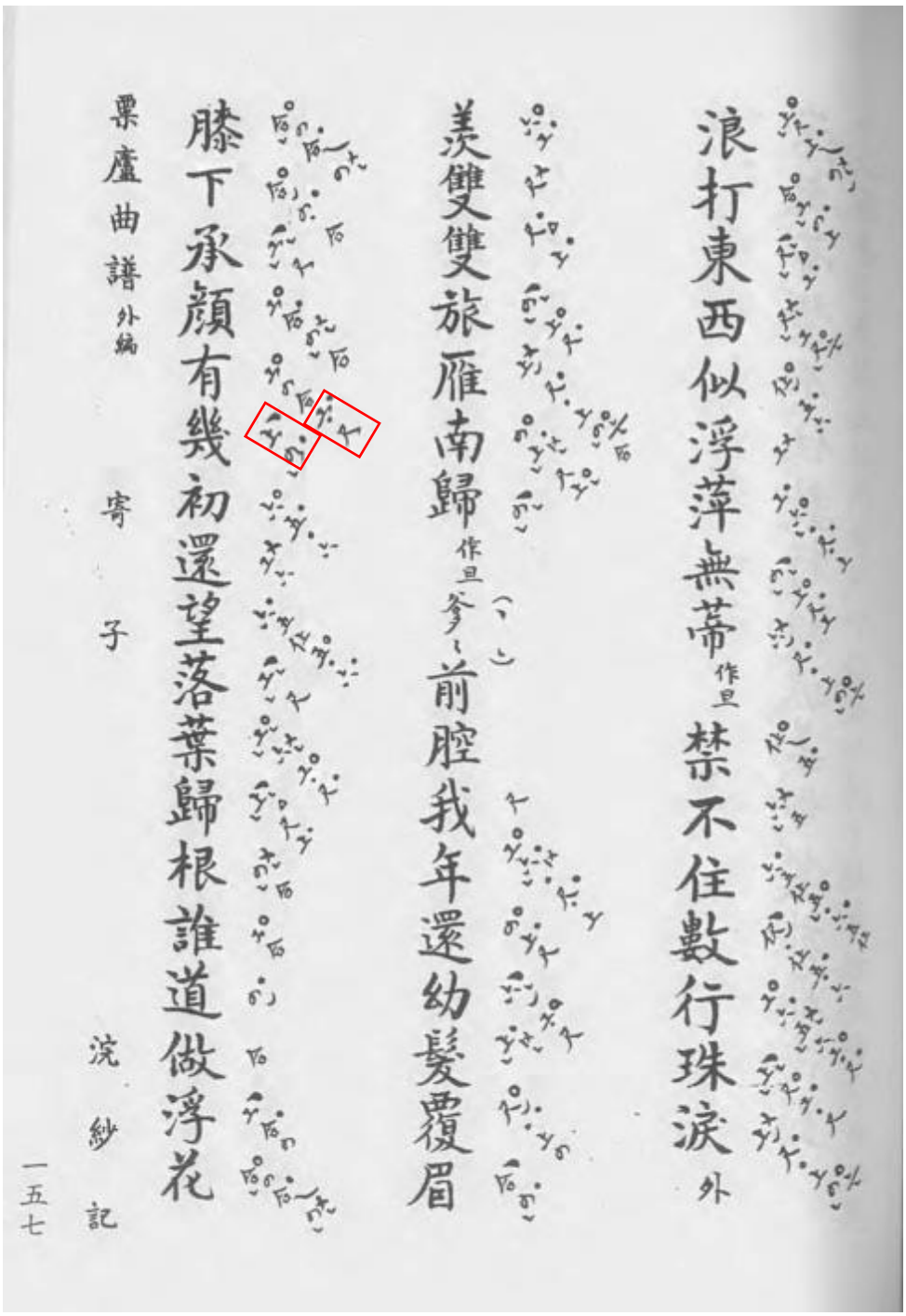


Fig. 24

顾铁华: 《粟庐曲谱外编·长生殿》, Self Printed, 2002, Folio 504

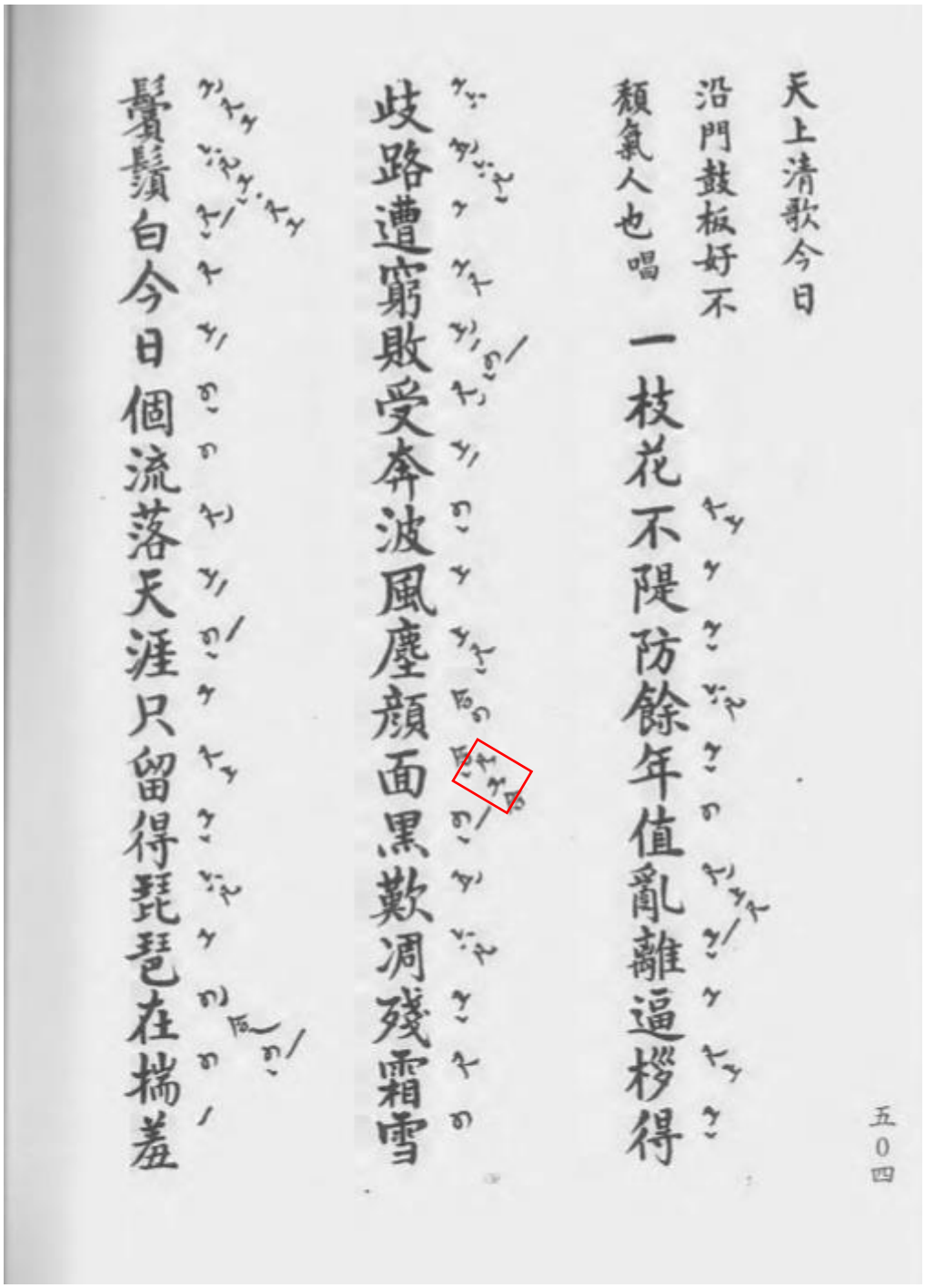


Fig. 25

顾铁华: 《粟庐曲谱外编·长生殿》, Self Printed, 2002, Folio 505

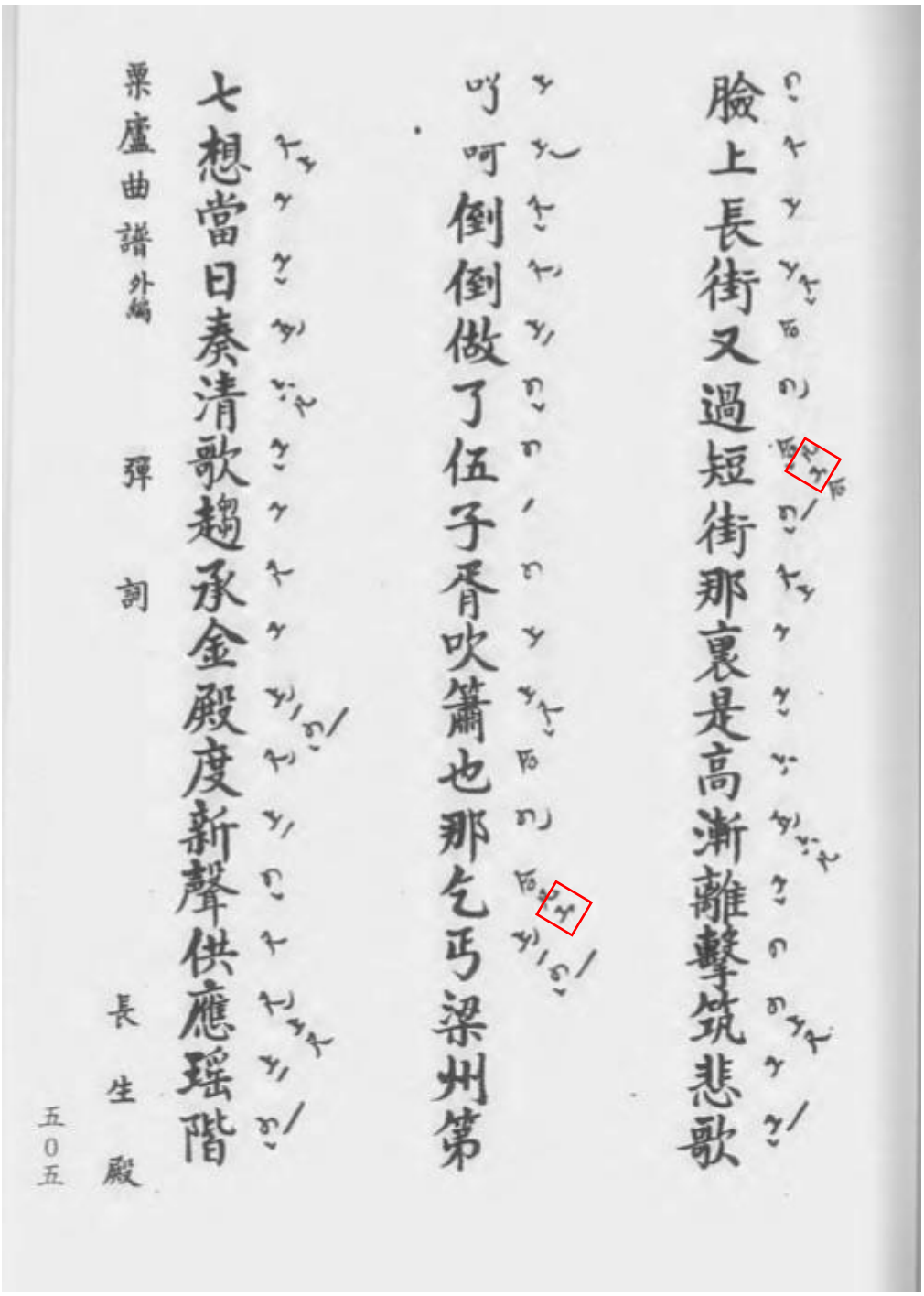


Fig. 26

顾铁华: 《粟庐曲谱外编·长生殿》, Self Printed, 2002, Folio 510

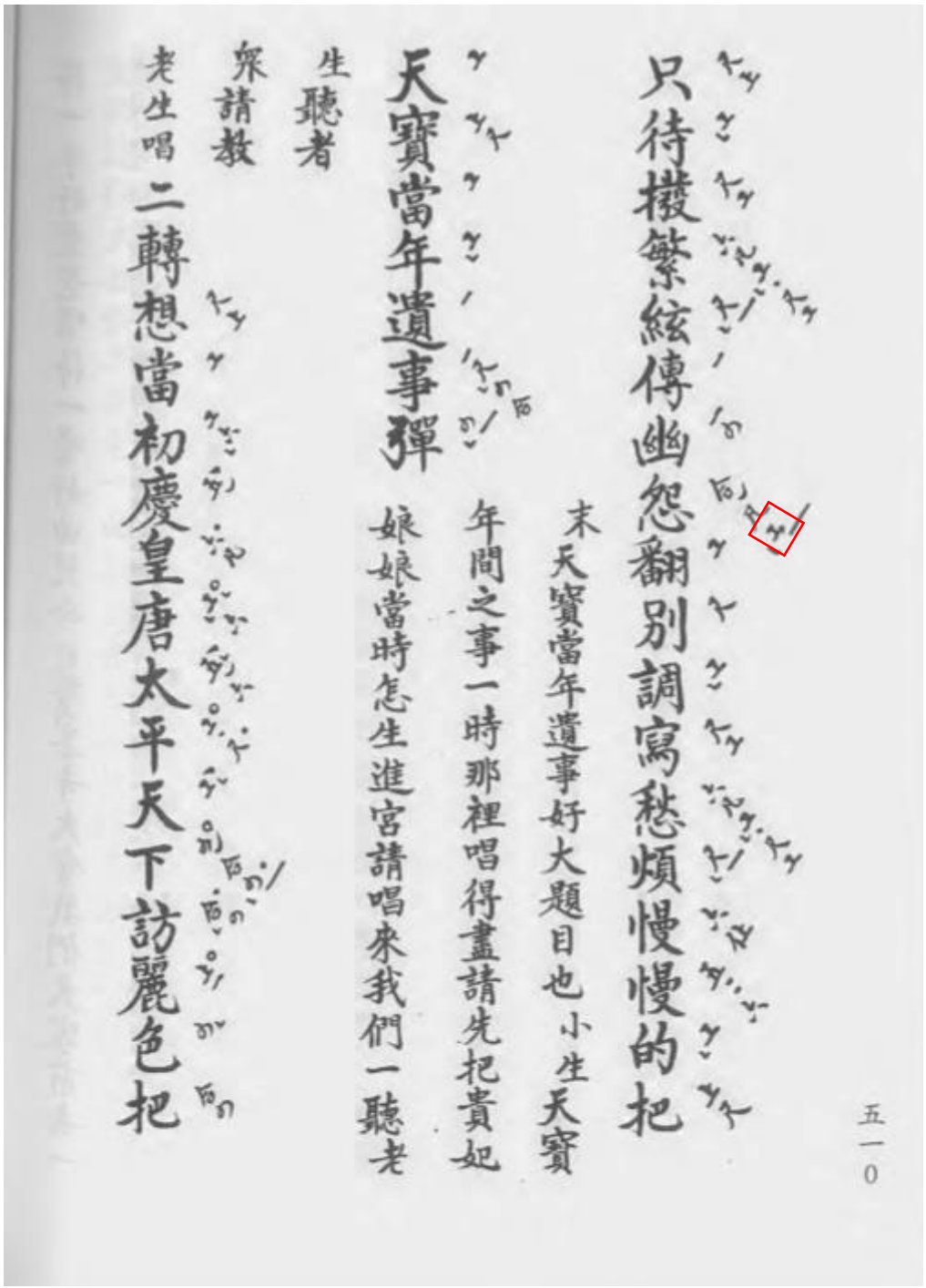


Fig. 27

顾铁华: 《粟庐曲谱外编·长生殿》, Self Printed, 2002, Folio 511

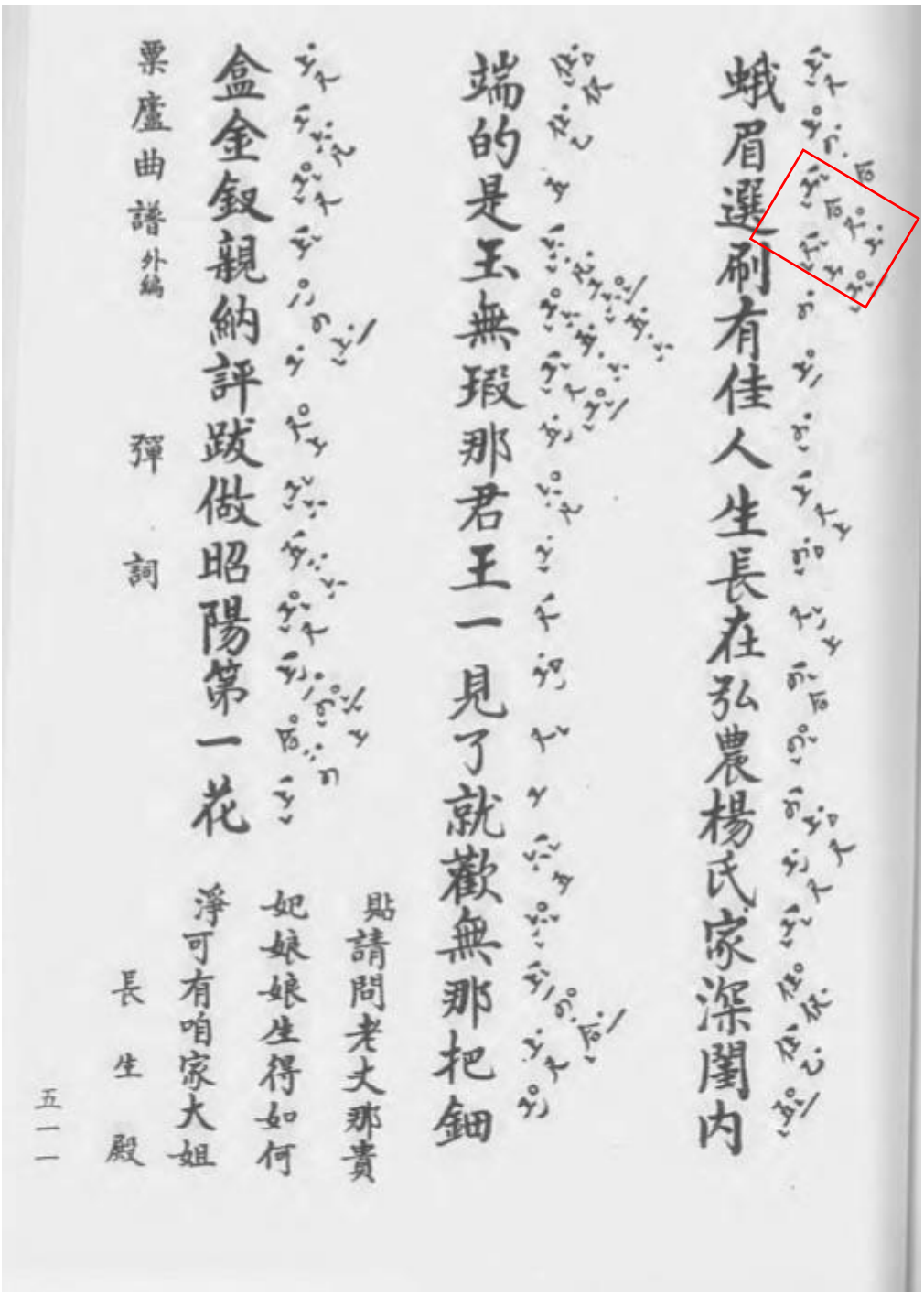


Fig. 28

郑剑西: 《二黄寻声谱》, 北京: 中国戏剧出版社, 2015, PP. 38-39

*This figure was provided by 新水令 who was a user on Zhihu.com.

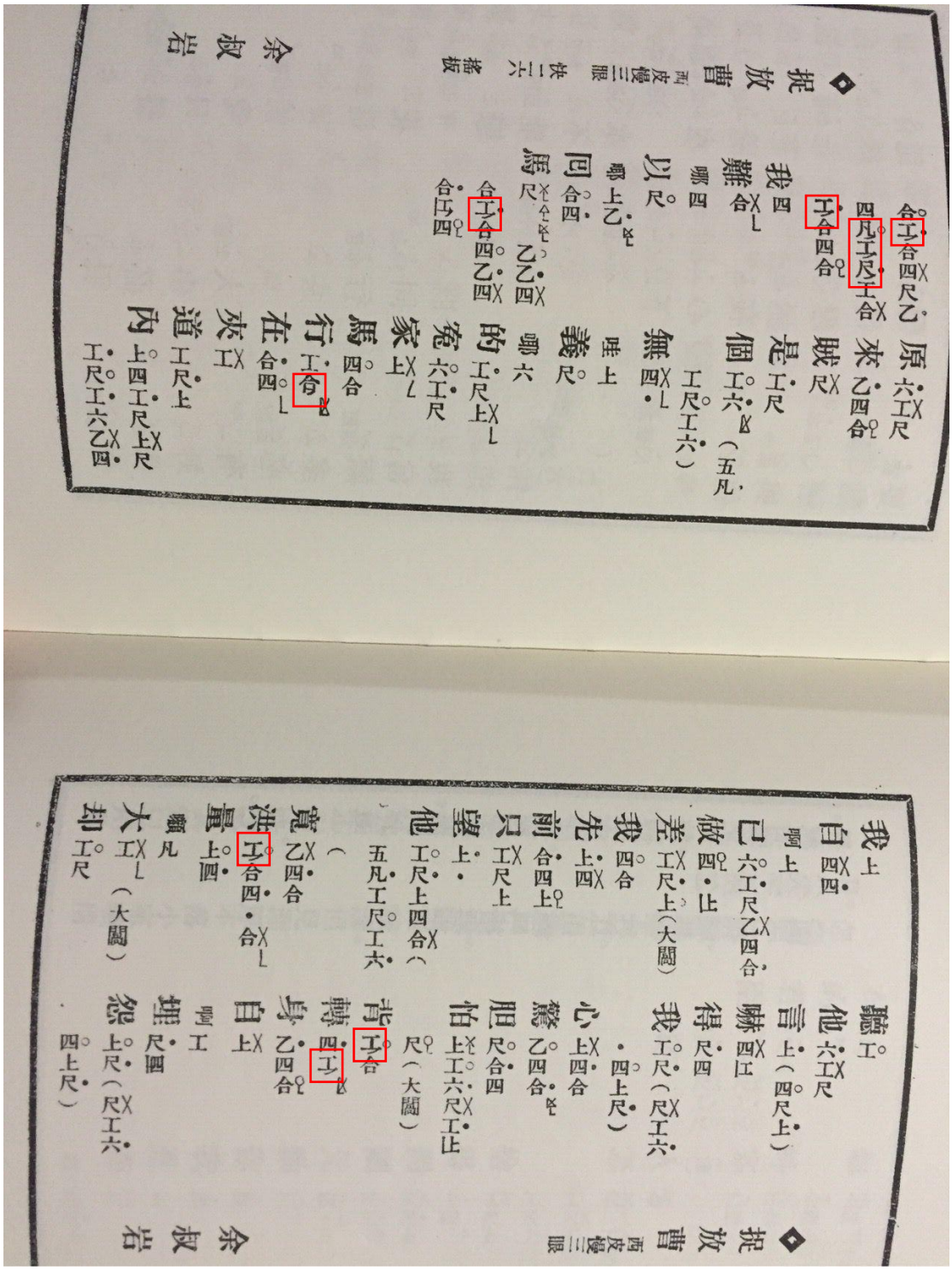
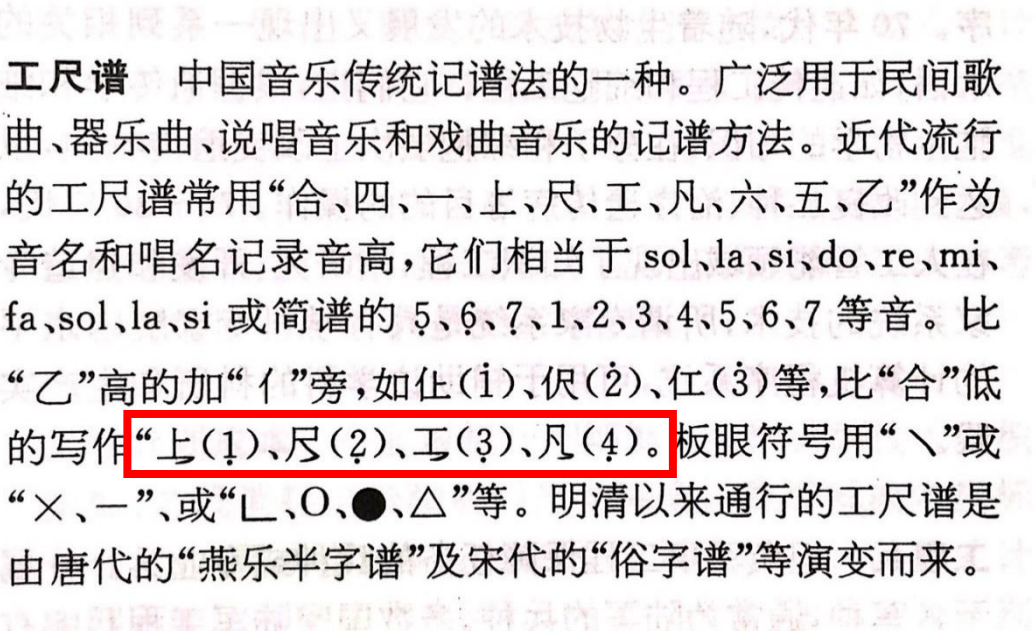


Fig. 29

王伯恭: 《中国百科大辞典》, 北京: 中国大百科全书出版社, 1999, ISBN 7-5000-6212-5/Z·113, P. 1711



工尺谱 中国音乐传统记谱法的一种。广泛用于民间歌曲、器乐曲、说唱音乐和戏曲音乐的记谱方法。近代流行的工尺谱常用“合、四、一、上、尺、工、凡、六、五、乙”作为音名和唱名记录音高，它们相当于 sol、la、si、do、re、mi、fa、sol、la、si 或简谱的 5、6、7、1、2、3、4、5、6、7 等音。比“乙”高的加“亻”旁，如仩(i)、伬(2)、仜(3)等，比“合”低的写作“上(1)、尺(2)、工(3)、凡(4)。板眼符号用“丶”或“×、—”、或“乚、O、●、△”等。明清以来通行的工尺谱是由唐代的“燕乐半字谱”及宋代的“俗字谱”等演变而来。

Notice that China once submitted 丄 & 凡 as G_BK100001 and G_BK100010 to CJK C1 and CJK D, but, as we know, they hadn't been included in the final version of CJK Ext. E.

Fig. 30 IRGN907 CJK C1 V 1.0, P. 1

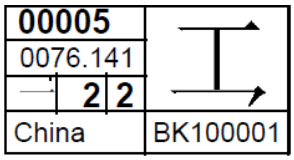


Fig. 31 IRGN907 CJK C1 V 1.0, P. 5

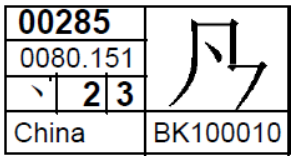


Fig. 32 IRGN1272 CJK D V 1.0, P. 1

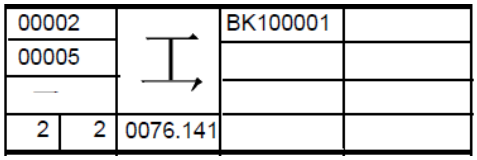
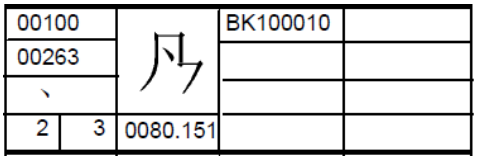


Fig. 33 IRGN1272 CJK D V 1.0, P. 3



8. Bibliography

王锡纯(Wáng Xīchún), 李秀云(Lǐ Xiùyún): 《遏云阁曲谱》(*Èyúngé Opern*), 遏云阁(Èyúngé), 1870

王季烈(Wáng Jìliè), 刘富梁(Liú Fùliáng): 《集成曲谱》(*Integratation Opern*), 上海: 商务印书馆(Shanghai: Commercial Press), 1925

俞振飞(Yú Zhènfēi): 《粟庐曲谱》(*Sùlú Opern*), 上海: 上海辞书出版社(Shanghai: Shanghai Lexicographical Publishing House), 1953, reprinted in 2011

王伯恭(Wáng Bógōng): 《中国百科大辞典》(*Encyclopedia Dictionary of China*), 北京: 中国大百科全书出版社(Beijing: Encyclopedia of China Publishing House), 1999, ISBN 7-5000-6212-5/Z·113

顾铁华(Gù Tiěhuá): 《粟庐曲谱外编》(*Additional Works of Sùlú Opern*), Self Printed, 2002

陈根方(Chén Gēnfāng): 《中国工尺谱的数字实现研究》 (*Research on Digitization of Chinese Gongche Notation*), 上海大学博士学位论文(Shanghai University Doctoral Dissertation), 2011

赵晓楠(Zhào Xiǎonán): 《工尺谱常识与视唱》(*Common Sense and Sightsinging of Gongche Notation*), 北京:人民音乐出版社(Beijing: People's Music Publishing House), 2014, ISBN 978-7-103-04558-9

郑剑西(Zhèng Jiànxī): 《二黄寻声谱》(*Èrhuáng Opern*), 北京: 中国戏剧出版社(Beijing: China Theatre Press), 2015

9. Acknowledgements

Mr. Jerry You(游程宇) and Mr. John Jenkins(井作恒) provide the fonts.

Dr. Ken Lunde(小林劍), Mr. Andrew West(魏安), Mr. Eric Liu(刘庆), Mr. Liang Hai(梁海), Dr. Jiao Lei(焦磊) and Miss Prajñavṛttā(པྲཛྙཱཝྲྀཏྟཱ) give some useful comments.

(End of Document)

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html **.**
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *Updated proposal on Gongche characters for Kunqu Opera*
2. Requester's name: *Eiso CHAN, Jerry YOU, WANG Xiaolei and HOU Le*
3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*
4. Submission date: *2018-06-01*
5. Requester's reference (if applicable):
6. Choose one of the following:
   This is a complete proposal: *YES*
   (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
   a. This proposal is for a new script (set of characters): *YES*
      Proposed name of script: *Gongche*
   b. The proposal is for addition of character(s) to an existing block: *NO*
      Name of the existing block:
2. Number of characters in proposal: *7*
3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary		B.1-Specialized (small collection)	X	B.2-Specialized (large collection)	
C-Major extinct		D-Attested extinct		E-Minor extinct	
F-Archaic Hieroglyphic or Ideographic				G-Obscure or questionable usage symbols	

4. Is a repertoire including character names provided? *YES*
   a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? *YES*
   b. Are the character shapes attached in a legible form suitable for review? *YES*
5. Fonts related:
   a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? *Jerry YOU*
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): *Jerry YOU, chengyu517@hotmail.com*
6. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *YES*
   b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *YES*
7. Special encoding issues:
   Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? *YES*

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

[1] Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? *NO*
   If YES explain
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? *NO*
   If YES, with whom?
   If YES, available relevant documents:
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? *NO*
   Reference:
4. The context of use for the proposed characters (type of use; common or rare) *Rare*
   Reference:
5. Are the proposed characters in current use by the user community? YES
   If YES, where? Reference:
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? *NO*
   If YES, is a rationale provided?
   If YES, reference:
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? *YES*
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? *NO*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? *NO*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? *NO*
    If YES, is a rationale for its inclusion provided?
    If YES, reference:
11. Does the proposal include use of combining characters and/or use of composite sequences? *NO*
    If YES, is a rationale for such use provided?
    If YES, reference:
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
    If YES, reference:
12. Does the proposal contain characters with any special properties such as control function or similar semantics? *NO*
    If YES, describe in detail (include attachment if necessary)
13. Does the proposal contain any Ideographic compatibility characters? *NO*
    If YES, are the equivalent corresponding unified ideographic characters identified?
    If YES, reference: