Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC 1/SC 2/WG 2 IRG N1706

Date: 2010-09-30

Source:	Toshiya Suzuki
Title:	Consolidated P&P of Old Hanzi
Actions required:	For consideration at IRG #35 at Macao
Distribution:	Old Hanzi experts
Medium :	Electronic

This document is the collection of the existing IRG documents submitted by Old Hanzi group or by the experts before the official adoption of Old Hanzi project. The documents related with the policy & processes are selected. The list of the sources and selected glyphs are excluded.

By this consolidated documents, we can track the background reasons of each rules, and check the lost rules (the rules that had ever been agreed but not applied now), the conflict or mismatches between the reason of the rule (why it was introduced) and the application of the rule (how it is applied).

Old Hanzi experts are requested to review the historical progress of Old Hanzi P&P, and send their comments to the project secretary, Selena. The deadline of the review comments is scheduled at 2010-Oct-22nd (see IRG N1703). The consolidated review comments will be discussed in IRG#35 at Macao, China.

IRG N1014 Draft Agreement on Old Hanzi Encoding

Source: LI Guoying <bjlgy@hotmail.com> and Tom BISHOP <tbishop@wenlin.com>
Status: Expert Contribution
Date: 2003-11-19
Action: Consideration by IRG

On November 17, 2003, in room 353 of Guilin Park Hotel, the IRG Old Hanzi Encoding Interest Group held a discussion, led by Prof. LI Guoying, Vice Dean of Beijing Normal University. Also present were (in alphabetical order): Tom BISHOP of Wenlin Institute, USA; Kyeong Won LEE of Hanyang University, South Korea; LEE Kyoo Kap of Yonsei University, South Korea; LIU Zhijin of East China Normal University; LONG Yuchen of Macao; NGUYEN Quang Hong of Vietnam; WEI Li of The Commercial Press; YE Maofeng of China Electronic Technology Standardization Institute.

No formal vote was taken, but there appeared to be agreement on the following.

(1) There is a pressing need for encoding Old (pre-Qin Dynasty) Hanzi to facilitate important historical and linguistic research, education, and publishing. The absence of a standard encoding is an obstacle to creating and exchanging Old Hanzi texts, fonts, and databases.

(2) Three types of Old Hanzi that particularly need encoding are Jiaguwen (oracle bone inscriptions), Jinwen (bronze inscriptions), and Xiaozhuan (small seal). While other types (such as Zhanguo Wenzi, writing from the Warring States period) should eventually be considered, the three types are very important and have been most extensively studied, and are therefore emphasized at this stage.

(3) There is generally no one-to-one correspondence between Old Hanzi and modern Hanzi, or between the main types of Old Hanzi, and the meanings and correspondences of many Old Hanzi are still unknown. Therefore, each type of Old Hanzi should be completely encoded in its own separate block (even when there appears to be a clear one-to-one correspondence with modern encoded Hanzi). This will enable the different types to be distinguished in plain text, and will facilitate research leading to mappings of correspondences.

(4) Scholars commonly arrange Old Hanzi (not only Xiaozhuan but also Jiaguwen and Jinwen) according to the 540 radicals of Shuowen Jiezi. It is therefore proposed to use this order for assigning codepoints for all types of Old Hanzi. Old Hanzi that don't correspond to Shuowen characters may be positioned either according to their theoretical Shuowen radicals (when plausible), or in a section for characters without determined radicals. The exact ordering may necessarily be somewhat arbitrary, but once standard encodings exist, it will be possible to implement useful indexes and look-up methods according to various organizational principals.

(5) For deciding between unification and separation of similar forms that are believed to be variants of each other, the unification principles for Old Hanzi should be established, resembling those for modern Hanzi, but with slightly more tendency toward separation, since the possible significances of some small differences are still uncertain, and Old Hanzi are primarily of interest to scholars who often need to make fine distinctions.

(6) Only authentic Old Hanzi should be encoded in the new Old Hanzi blocks. Sometimes new imitation Old Hanzi, corresponding to modern Hanzi, are created for decorative purposes, but such usage is unimportant for research into authentic Old Hanzi, and in any case is already possible using special fonts with the same encodings as modern Hanzi.

(7) The next steps to be taken include: (a) attempt to establish a consensus on the principles listed above; (b) prepare draft code charts for the three main types of Old Hanzi, along with detailed sources and references for each character; (c) make the draft code charts available for evaluation by the international community; (d) revise the draft code charts on the basis of this evaluation; (e) submit the revised code charts to IRG for approval.

(8) The following persons have volunteered to take responsibility for initial preparation and distribution of draft code charts: LI Guoying for Xiaozhuan and Jiaguwen; LIU Zhijin for Jinwen.

(End of Document)

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1087

Date: 2004-11-27

Doc. Type:	Member body contribution
Title:	A Statement on old Hanzi Encoding Principles
Source:	TCA
Status:	Input to IRG
Action:	For discussion on the old Hanzi interest group
Distribution:	IRG Members and Ideographic Experts
Reference:	IRG N1055
No. of pages:	9
Medium:	Electronic

1. Introduction

Etymologists divide the history of Chinese scripts development into contemporary stage and ancient stage by the delimiter of the Liban (the seal script change). Nowadays, the Internet provides a newly cyber space. Facing the transitions of the Chinese languages and characters, there exist platforms for contemporary scripts appeared after Liban. Traditional-Simplified and form-variant character exchange mechanisms are available globally across country boarders. However, in order to set up a platform for ancient script digital processing, an encoding system for ancient script message exchange is needed for information process. A working group for ancient script encoding was agreed to be formed according to the conclusion of the 2003 IRG meeting held in Chengdu. In that meeting, the consensus of proper attention to both ancient and contemporary scripts was reached. Also, the scope of old Hanzi is bounded to Pre-Qin scripts appeared before the Liban. The conclusion was approved by WG2 meeting held in Canada, June 2004. TCA thus, gathering computer experts and ancient script scholars, formed a working group for ancient script encoding. The working group aggressively starts to discuss ancient script encoding rules and plan the flow of pre-work. The progress will be reported to the incoming IRG meeting in Cheju, Korea, asking opinions for the further work.

2. Encoding Reasons

2.1 Old Hanzi should be classified based on the principles of the time they were used, the locations they were collected, the materials they were on, their form-variances, and properties of form, pronunciation, and meaning. This establishes the baselines for old Hanzi encoding classification.

2.1.1 Period: Shang Dynasty, Western Zhou Dynasty, Spring and Autumn Period, Warring Kingdoms, and Qin Dynasty.

2.1.2 Materials: Jia-gu-wen (inscription on animal bones and tortoise shells) (divided into "the Shang Dynasty Jia-gu-wen " and "the Western Zhou Dynasty Jia-gu-wen), Jinwen (bronze inscription) (including Zhou-wen), jade and stone writings (jade carving, stone carving and so on, including other carving), bamboo and silk writings (bamboo books(Jan-tu) and silk books. Qin-li is part of "Jan-tu"), small Juahn (majority are "Shuowen Jiezi" small Juahns, covering "Shuowen Sinfu" small Juahns), and miscellanea (including ancient pottery writings, seal writings, currency writings, lacquer ware writings, etc.).

2.1.3 Locations: In Warring Kingdoms period, there were five character affiliations, namely Qin (秦), Chu(楚), Jin(晉), Qi(齊), and Yan(燕).

2.1.4 Writing-variances: Variant constructs for a character need to be encoded separately. In the case of variant writing, the representative writing should be chosen. Other writing-variances should be listed aside with the representative. Characters with variant constructs may be listed as notes. A construct-variance replaces a component in a character with a different component or a component with different writing. The writing-variance involves some stroke(s) in a character having minor different drawing(s), however the basic structure of the character is retained. The evolution of old Hanzi took a long time. The decision involves the structure differences of basic strokes. For example, if a dot in a character became a horizontal bar, the character is deemed as a construct- variance. However, characters with different sizes of dot of different lengths of horizontal bar will make them writing-variances.

2.2 Ancient script encodings are divided into six categories and twenty two blocks based on classification principles stated above.

<table>
<tr><td>Jiaguwen
(甲骨文)</td><td>1 Shang Jiaguwen
(商甲骨文)</td><td>2 Western Zhou Jiaguwen
(西周甲骨文)</td></tr>
<tr><td rowspan="5">Jinwen
(金文)</td><td>3 Shang Jinwen
(商金文)</td><td>4 Western Zhou Jinwen
(西周金文)</td></tr>
<tr><td>5 Spring and Autumn Period Jinwen
(春秋金文)</td><td>6 Warring Kingdoms Chu Jinwen
(戰國楚系金文)</td></tr>
<tr><td>7 Warring Kingdoms Jin Jinwen
(戰國晉系金文)</td><td>8 Warring Kingdoms Qi Jinwen
(戰國齊系金文)</td></tr>
<tr><td>9 Warring Kingdoms Yan Jinwen
(戰國燕系金文)</td><td>10 Warring Kingdoms Qin Jinwen (including Zhouwen)
(戰國秦系金文(含籀文))</td></tr>
</table>

Jade and stone writings (玉石文字)	11 Shang Jade and stone writings (商玉石文字)	12 Jade and stone writings (西周玉石文字)
	13 Spring and Autumn Period Jade and stone writings (春秋玉石文字)	14 Warring Kingdoms Chu Jade and stone writings (戰國楚系玉石文字)
	15 Warring Kingdoms Jin Jade and stone writings (戰國晉系玉石文字)	16 Warring Kingdoms Qi Jade and stone writings (戰國齊系玉石文字)
	17 Warring Kingdoms Yan Jade and stone writings (戰國燕系玉石文字)	18 Warring Kingdoms Qin Jade and stone writings (戰國秦系玉石文字)
Bamboo and silk writings (簡帛文字)	19 Warring Kingdoms Chu bamboo and silk writings (戰國楚系簡帛文字)	20 Warring Kingdoms Qin bamboo and silk writings (戰國秦系簡帛文字)
Small Juahn (小篆)	21 Small Juahn (小篆)	
Miscellanea	22 ancient pottery writings, seal writings, currency writings, lacquer ware writings, etc. (古陶文、璽印文字、貨幣文字、漆器文字)	

3. Encoding Process

3.1 Rules of character selection

3.1.1 Representative characters: Representative characters were chosen by experts. Rules for collecting representative characters are as follow: well-recognized characters are collected; uncertain or arguable characters are left out. Representative characters are manually drawn careful with aesthetic and uniform styles. Original scripts are scanned into graphic files with clear resolutions. Please see the appendix: Ancient Jinwen script encodings table.

3.1.2 Decision of construct-variances: Construct-variances are deemed as different ancient characters, thus they should be encoded separately. In the case of variant writing, the representative writing should be chosen. Other writing-variances should be listed aside with the representative. Characters with variant constructs may be listed as notes. Only the most five popular Construct-variances will be collected for each character.

3.1.3 Writing-variances should be affirmed by experts. Rules of their recognition should be logged character by character.

3.2 Encoding procedures

3.2.1 Fill in a row in the attribute table shown below for each character collected first. This table will be used as the foundation for encodings in each block. (Please see examples in the appendix.)

Running No.	Rep. Script	Original Shape	Source	Carrier	Period	Area	Material	Radical	Pronunciation	Glyph Determ.	Corresd. Modern Char	Notes

Data attributes:

3.2.1.1 Running number: used to identify individual script.

3.2.1.2 Representative script: Provides a reference script for comparison based on emulating drawing from the original material.

3.2.1.3 Original shape: Original scripts on archaeological finds or their rubbings.

3.2.1.4 Source (category): Records literature, page number, sequence number or index of the source of the script. (The sources are presented with abbreviations. The abbreviation table is given.)

3.2.1.5 Carrier：Name of the archaeological find or literature carried the script.

3.2.1.6 Period：The period the carrier of the script was used or created. There were five periods: Shang Dynasty, Western Zhou Dynasty, Spring and Autumn Period, Warring Kingdoms, and Qin Dynasty.

3.2.1.7 Area: The place the carrier was used or created.

3.2.1.8 Material: The material of the carrier of the script.

3.2.1.9 Radical: coded with three digits, corresponding to 540 radicals in "Shuowen Jiezi."

3.2.1.10 Pronunciation：mark the pronunciation of the script with phonetic symbols.

3.2.1.11 Glyph determination: Converts a script into a corresponding Kaishu glyph according to the rules of glyph determination.

3.2.1.12 Corresponding modern characters: Lists one or more possible corresponding modern character(s) of the script.

3.2.1.13 Notes: Related comment.

3.2.2 Organization and encoding sequence is as follow: Jiaguwen, Jinwen, jade and stone writings, jade and stone writings, bamboo and silk writings, small Juahn, and Miscellanea.

4. Ending words

4.1 Experiencing ancient script encoding workflow construction, the working group clarified attributes of shapes of ancient Chinese scripts, distinguished construct-variances from writing-variances, and confirmed recognition rules of shape variances. The working group encoded old Hanzi in order based on

attribute definitions. The working group constructed an ancient script encoding system to facilitate digital character exchange in Chinese character cultural circle.

4.2 The task of ancient script encoding is arduous; the questions related to it are complicated. If the task can be accomplished stage by stage, ancient literatures and archaeological found literatures will be able to be retrieved and printed. This will accelerate the speed of recording and organizing cultural heritages, and will accelerate publication and propagation of research results.

Appendix:

Shang Jinwen Script Encoding Table												
Running No,	Rep. script	Original shape	Carrier	Source	Period	Area	Material	Radical	Pronunciation	Glyph determ.	Corresd. modern char.	Notes
001			狽元作父戊卣[1]	《集成》5278.2-6	Shang		bronze	元	yuan	元	元	或釋為兀
002	 2		△天父癸簋	《集成》3340	Shang		bronze	天	tian	天	天	

Western Zhou Jinwen Script Encoding Table												
Running number	Rep. script	Original shape	Carrier	Source	Period	Area	Material	Radical	Pronunciation	Glyph determ.	Corresd. modern char.	Notes
001			我方鼎	《集成》2763.1-4	W. Zhou		bronze	一	yi	一	一	
002			曾鼎	《集成》2838-6	W. Zhou		bronze	元	yuan	元	元	
003			沇兒鐘	《集成》203.2-9	W. Zhou		bronze	元	yuan	元	元	
004			天作从尊	《集成》5688	W. Zhou		bronze	天	tian	天	天	
005			頌壺	《集成》9731	W. Zhou		bronze	天	tian	天	天	

[1] 器名舊作〈兀作父戊卣〉。
[2] 字形見於〈△天父癸簋〉。

Spring and Autumn Period Jinwen Script Encoding Table												
Running No.	Rep. script	Original shape	Carrier	Source	Period	Area	Material	Radical	Pronunciation	Glyph determ.	Corresd. modern char.	Notes
001			秦公簋	《集成》4315.3-9	Spring & Autumn	Qin	bronze	一	yi	一	一	
002			朐簋		Spring & Autumn		bronze	元	yuan	元	元	
003			秦公簋	《集成》4315.3-9	Spring & Autumn		bronze	元	yuan	元	元	
004	 [3]		少虞剑	《集成》11696	Spring & Autumn		bronze	元	yuan	元	元	
005	 [4]		虞公剑	《集成》11663	Spring & Autumn		bronze	元	yuan	元	元	
006			吳季子之子剑	《集成》11640	Spring & Autumn		bronze	元	yuan	元	元	
007			洹子孟姜壶	《集成》9730	Spring & Autumn		bronze	天	tian	天	天	
008				《集成》211	Spring & Autumn		bronze	天	tian	天	天	

[3] 器名舊稱〈吉日壬午劍〉。
[4] 器名舊稱〈斁公劍〉。

Warring Kingdoms Jin Jinwen Script Encoding Table

Running No.	Rep. script	Original shape	Carrier	Source	Period	Area	Material	Radical	Pronunciation	Glyph determ.	Corresd. modern char.	Notes
001			十一年皋落戈	《考古》1991年第5期	Warring Kingdoms	Jin	bronze	一	yi	一	一	
002			元年鄭令矛	《集成》11552	Warring Kingdoms	Jin	bronze	元	yuan	元	元	
003			中山王△鼎	《集成》211	Warring Kingdoms	Jin	bronze	天	tian	天	天	
004			中山侯恣鉞	《集成》11758	Warring Kingdoms	Jin	bronze	天	tian	天	天	

Warring Kingdoms Yan Jinwen Script Encoding Table

Running No.	Rep. script	Original shape	Carrier	Source	Period	Area	Material	Radical	Pronunciation	Glyph determ.	Corresd. modern char.	Notes
001			重金扁壺	《集成》9617-8	Warring Kingdoms	Yan	bronze	一	yi	一	一	

Warring Kingdoms Chu Jinwen Script Encoding Table

Running No.	Rep. script	Original shape	Carrier	Source	Period	Area	Material	Radical	Pronunciation	Glyph determ.	Corresd. modern char.	Notes
001			鄂君啓舟節	《集成》12113	Warring Kingdoms	Chu	bronze	能	yi	△	一能	字從能翼省聲

Warring Kingdoms Qin Jinwen Script Encoding Table												
Running No.	Rep. script	Original shape	Carrier	Source	Period	Area	Material	Radical	Pronunciation	Glyph determ.	Corresd. modern char.	Notes
001	 5		卅一年相邦冉戈	《集成》12342	Qin	Qin	bronze	一	yi	一	一	

Warring Kingdoms Qi Jinwen Script Encoding Table												
Running No.	Rep. script	Original shape	Carrier	Source	Period	Area	Material	Radical	Pronunciation	Glyph determ.	Corresd. modern char.	Notes
001			陳逆簋	《集成》4630	Warring Kingdoms	Jin	bronze	元	yuan	元		

[5] 原形參王輝編著《秦銅器銘文編年集釋》圖四十七，頁 57。舊稱〈廿一年相邦冉戈 〉

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1102

Date: 2004-12-01

Title:	**Resolution from the Old Hanzi Expert Group**
Source:	**Old Hanzi Expert Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**5**
Medium:	**Electronic**

The members of the Old Hanzi Expert Group has discussed the issues of Old Hanzi for preparing valid encoding model at the Oriental Hotel, in Jeju of Korea in Dec 1, 2004. In the meeting, the members reviewed the documents IRGN999 and IRGN1052 first, and then discussed the working procedure. The followings are the conclusion:

1. Old Hanzi should be considered and classified as different scripts from CJK Ideographs.

There is a very closed relationship between Old Hanzi and modern ideographs. Basically, Old Hanzi is the ancient type of modern ideographs. But there are some differences between them. The basic handing unit of Old Hanzi is line and that of modern ideographs is stroke. In addition, there is not one-to-one mapping between Old Hanzi and modern ideographs. Based on the reasons above, Old

Hanzi should be encoded separately from CJK Ideographs.

2. Each type of Old Hanzi, due to their unique characteristics, should be classified as separate script.

There is generally no one-to-one mapping between Old Hanzi and modern ideograph, or between the main types of Old Hanzi. The meanings and correspondences of many Old Hanzi are still unknown. Therefore, each type of Old Hanzi should be encoded in its own block (even when there is a clear one-to-one mapping with modern ideograph). This will enable the different types to be distinguished in plain text, and will facilitate research leading to mappings.

3. Types of Old Hanzi：

For the time being, Old Hanzi Expert Group has identified the following first three types of Old Hanzi as the work scope:

(1) Oracle bone inscription (甲骨文)

(2) Bronze inscription(金文)

(3) Small Seal(小篆)

(4) Others

Note: There is still no common understanding about scripts during Warring States (in the 4th type above). It will be studied later.

4. Principles and rules of character selection

(1) Principle of historical authenticity:

Only authentic Old Hanzi should be encoded in the Old Hanzi blocks. Sometimes new Imitation Old Hanzi, corresponding to modern ideographs, are created for decorative purposes. These Imitation Old Hanzi are unimportant for research of Old Hanzi, and in any case they can be represented accurately using special fonts with the same encodings as modern ideographs.

(2) Principle of holding significant nuance:

For deciding between unification and separation of similar forms that are believed to be variants of each other, the unification principles for Old Hanzi should be established. Comparing to the unification principles of CJK unified ideographs, the unification principles would slightly tend more towards separation, since some small differences could have significance in future scholarly research. It is important to point out that Old Hanzi are primarily of interest to scholars who often need to make fine distinctions.

(3) Principle of glyph integrity:

Only complete glyphs will be selected, excluding those broken with missing parts.

(4) Rules of character selection

An Old Hanzi may have different glyphs due to their differences in the time period of use, place of origin and aesthetic purposes. All such different glyphs of an Old Hanzi would be selected.

5. Rules of collation

Scholars commonly arrange Old Hanzi (not only Small Seal but also Oracle and Bronze inscription) according to the 540 radicals of Shuowen Jiezi(《說文解字》). It is therefore proposed to use the 540 radicals as collating scheme for all types of Old Hanzi. Detailed collating rules (in their order of precedence) are listed in the following:

(1) Old Hanzi found in Shuowen Jiezi will be sorted as they were in Shuowen Jiezi (540 radicals).

(2) Old Hanzi not found in Shuowen Jiezi will be assigned a radical for sorting in according to the Shuowen Jiezi (540 radicals).

(3) For Old Hanzi which cannot be handled by the above 2 rules, they will be sorted according to their attributes (namely: the time period of use, place of origin and aesthetic purposes).

6. Format of submission

Serial No.	Rep. Script/ Glyph	Original Shape/ Glyph	Source	Period/ Epoch	Area/ Terrain	Material	Radical	Glyph Determ.	Corresp. Modern Char	Notes
1										
2										

For example:

7. The work plan

Member parties interested are requested to submit their Old Hanzi samples to the Old Hanzi Technical Secretary, in the format shown above, before 31 January 2005. The samples should include three types, namely Oracle bone inscription(甲骨文), Bronze inscription(金文) and Small Seal(小篆) of Old Hanzi, with ten sample glyphs for each type.

Members attended the meeting were Li Guoying, Wang Tiekun, Yin Jianghong, Dai Hong, Wang Xiaoming, Shi Jianqiao, Zhang Liwei, Wu Jian, Bill Ling, Heo Chul, Julie S. C. Chuang, Judy Liu, Hsu Hsueh Jen, Selena Wei, WU Lieh-neng, Lee Jae Hoon, Lee Kyoo Kap, Lee Kyeong Won, Yasuhiro Anan, ATSUJI Tetsuji, Chan Kim Kun (Arnie), Ng Sio U, Richard S. Cook.

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N

Date: 2005-05-17

Doc. Type:	Member body contribution
Title:	Changes to the IRG N1102
Source:	TCA
Status:	Input to IRG
Action:	For Discussion in old Hanzi expert group
Distribution:	IRG Members and old Hanzi Experts
Reference:	IRG N1102
No. of pages:	2
Medium:	Electronic

In order to improve the working procedure for the old Hanzi, the following proposed changes should be made to the document IRG N1102.

Propose 1：

Text：

4. Principles and rules of character selection

(3) Principle of glyph integrity:

Only complete glyphs will be selected, excluding those broken with missing parts.

Suggest change to：

Only complete glyphs and some broken part can be discriminate will be selected, excluding those broken with missing parts .

In all such instances

			《合》22124
			《合》22072
			《合》14841

Propose 2：

Text：

5. Rules of collation

(2) Old Hanzi not found in Shuowen Jiezi will be assigned a radical for sorting in according to the Shuowen Jiezi (540 radicals).

Suggest change to ：

(2) Old Hanzi not found in Shuowen Jiezi should be judged it's radical and assigned for sorting in according to the Shuowen Jiezi (540 radicals).

Propose 3：

Text：

5. Rules of collation

(3) For Old Hanzi which cannot be handled by the above 2 rules, they will be sorted according to their attributes (namely: the time period of use, place of origin and aesthetic purposes).

Suggest change to ：

(3) Those old Han-Zi which can not be sorted and classified by the two rules above should be put in order by its structures and categories and then be attached to the relating radicals in Shuowen Jiezi

Propose 4：

4. Principles and rules of character selection

"(1) Principle of historical authenticity" and "(4) Rules of character selection" might be conflicted.

Text：

(1) Principle of historical authenticity:

Only authentic Old Hanzi should be encoded in the Old Hanzi blocks. Sometimes new Imitation Old Hanzi, corresponding to modern ideographs, are created for decorative purposes. These Imitation Old Hanzi are unimportant for research of Old Hanzi, and in any case they can be represented accurately using special fonts with the same encodings as modern ideographs.

(4) Rules of character selection

An Old Hanzi may have different glyphs due to their differences in the time period of use, place of origin and aesthetic purposes. All such different glyphs of an Old Hanzi would be selected.

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N

Date: 2005-05-17

Doc. Type:	Member body contribution
Title:	Problems regarding old Hanzi
Source:	TCA
Status:	Input to IRG
Action:	For Discussion in old Hanzi expert group
Distribution:	IRG Members and old Hanzi Experts
Reference:	IRG N1102
No. of pages:	3
Medium:	Electronic

We found the following problems which need to raise discussion during the meeting:

1. A detail definition of the standard rules of Glyph Determ should be clarified.

   TCA propose:

   Those words which could be sorted in Glyph Determ can be defined by Kai-Hua (the formal style), based on the original word shapes. Kai-Hua is different from Kai-Ti (Kai script). Kai-Hua is not written as in successive order as Kai-Ti.

2. The rules should clearly distinguish the differences between Yi-Gou and Yi-Xie

   TCA propose:

   (1) Yi-Gou: In the construction of words, some parts of a word is written differently, or the original part A (甲)is replaced by part B (乙); then, the word which is constructed by part B is Yi-Gou.

   (2) Yi-Xie: Referring to some written words, there are slightly different in successive orders of writing but it will not change the basic shape of the word. It is Yi-Xie.

   Note 1: However, because the old Han-Zi(script) has been formed long time ago, the judgment will affect the word shapes of the basic successive order of written words. For example, when Dian(dot)turns to Heng(Horizontal line), it is Yi-Gou. But the size of dots (big or small) and the size of the horizontal lines (long or short) belong to Yi-Xie.

   Note 2: If based on the successive order of a written word as a measurement, all those words that their bodies have been changed or replaced are termed Yi-Gou. However, the

number of strokes increase or decrease in formation then it belongs to Yi-Xie. If it affects the word structure then it still belongs to Yi-Gou.

3. Can Yi-Gou and Yi-xie be coded separately and independently.

   TCA propose:

   (1) Yi-Gou can be collected as one file and coded

   (2) Yi-Xie can not be collected as a file and not be coded

4. Serial No. the ways of arranging code

   TCA propose:

   0001-03-07: Code 0001-- for radicals (according to the listing order in Shuowen Jiezi), which is 4 digits; Code 03-- for section 22, which is two digits, such as (01) for Shan Dynasty Jia-Gu(Ancient. script on tortoise shells or bones), (02) for West-Zhou Dynasty Jia-Gu, (03) for Shan Dynasty Jin-Wen(inscriptions on ancient bronze vessels) and so on; (07) is serial number.

5. What radical should be taken?

   TCA propose:

   Take the radical order based on Duan-Yu-Cai-Zhu of Shuowen Jiezi Zhu(footnotes) by Xue-Shen as a rule.

Notes: 22 block --

Jiaguwen (甲骨文)	1 Shang Jiaguwen (商甲骨文)	2 Western Zhou Jiaguwen (西周甲骨文)
Jinwen (金文)	3 Shang Jinwen (商金文)	4 Western Zhou Jinwen (西周金文)
	5 Spring and Autumn Period Jinwen (春秋金文)	6 Warring Kingdoms Chu Jinwen (戰國楚系金文)
	7 Warring Kingdoms Jin Jinwen (戰國晉系金文)	8 Warring Kingdoms Qi Jinwen (戰國齊系金文)
	9 Warring Kingdoms Yan Jinwen (戰國燕系金文)	10 Warring Kingdoms Qin Jinwen (including Zhouwen) (戰國秦系金文(含籀文))
Jade and stone writings (玉石文字)	11 Shang Jade and stone writings (商玉石文字)	12 Jade and stone writings (西周玉石文字)
	13 Spring and Autumn Period Jade and stone writings (春秋玉石文字)	14 Warring Kingdoms Chu Jade and stone writings (戰國楚系玉石文字)
	15 Warring Kingdoms Jin Jade and stone writings (戰國晉系玉石文字)	16 Warring Kingdoms Qi Jade and stone writings (戰國齊系玉石文字)
	17 Warring Kingdoms Yan Jade and stone writings (戰國燕系玉石文字)	18 Warring Kingdoms Qin Jade and stone writings (戰國秦系玉石文字)
Bamboo and silk writings (簡帛文字)	19 Warring Kingdoms Chu bamboo and silk writings (戰國楚系簡帛文字)	20 Warring Kingdoms Qin bamboo and silk writings (戰國秦系簡帛文字)
Small Juahn (小篆)	21 Small Juahn (小篆)	
Miscellanea	22 ancient pottery writings, seal writings, currency writings, lacquer ware writings, etc. (古陶文、璽印文字、貨幣文字、漆器文字)	

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1135

Date: 2005-05-25

Title:	**Report from the Old Hanzi Expert Group**
Source:	**Old Hanzi Expert Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**6**
Medium:	**Electronic**

The Old Hanzi Expert Group has discussed the issues of the repertoire development of Old Hanzi. Out of four types of Old Hanzi (Oracle bone, Bronze, Small Seal, others), the group decided to focus on the development of Oracle bone inscription repertoire first. The group came up with the following agreements:

1. Format of submission (Update)

ID	Rep. Script/ Glyph	Original Shape/ Glyph	Source	Period/ Epoch	Area/ Terrain	Material	SW Radical	SW Radical Number	Glyph Determ.	Corresp. Modern Char	Notes
1											
2											

The finalized format for the member submission is as follows:

† ID is the unique id that consists of one or two letters member id

(G, T, K, KP, J, V, S, H, M) followed by four digit sequential numbers assigned by submitters.

Example: T0001 is one IRG global unique ID assigned to an Old Hanzi submitted by TCA.

† SW Radical: ShuoWen Radical

† SW Radical number : 1 - 540

† Note. YiGou features will be noted for clarification.

† The format of bitmap images for rep. glyphs, original shapes and 540 SW Radicals are specified as follows:

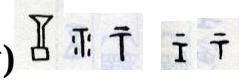
128x128，Black and white bitmap

2. The principles of Old Hanzi selection

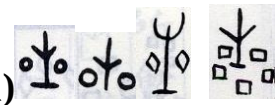
2.1 Distinction principles

Two or more instances of Old Hanzi are considered different from each other when either one of the following conditions apply:

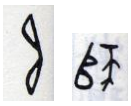
2.1.1 One or more types of components are different.

(莫)  **，(示)** 

2.1.2 The number of components or lines is different.

(介)  **，(星)** 

2.1.3 The direction (e.g. mirror image) of a component is different.

(人)  **，(師)**  **，(至)** 

2.1.4 The position of one or more components is different.

(好) 

2.1.5 Whether the same set of components are connected each other or not.

(伐) 

2.1.6 One or more line types (straight line, curve, circle, rectangle, closed line or curve filled inside) are different.

(王) 

2.2 Unification principles

Two or more instances of Old Hanzi with the following differences will be unified unless there's any difference in the meaning:

2.2.1 The length of corresponding line is different.

2.2.2 The thickness of corresponding line is different.

2.2.3 The size of each component of the same set is different.

3. The work plan

3.1 Working plan of the Oracle Bone Inscription

Member parties are requested to submit their Old Hanzi collection of the Oracle Bone Inscription based on the six references listed below to the group Chairperson (Prof. Li Guoying), in the format specified above, according to the following schedule.

(The IRG#25 will be held 2005-11-28)

Target Date	Work Items	Assignment
One month before IRG#25 (Oct. 28)	Submit collection of the Oracle bone inscription	Each member body
10 days after the work item above (Nov. 7)	Consolidation of member contributions and Distribution of consolidated documents to all member bodies.	Chairperson
10 days before IRG#25 (Nov. 18)	Submit feedback and comments to the Chairperson for discussions in IRG#25.	Each member body

References of the Oracle Bone Inscription

The collection of the Oracle Bone Inscription will be developed to cover the following six books:

1. 《甲骨文合集》
2. 《甲骨文合集补编》
3. 《花园庄东地甲骨》
4. 《山东济南大辛庄甲骨》
5. 《周原甲骨》
6. 《小屯南地甲骨》

3.2 The purpose and use of the Old Hanzi

Each submission should include detail information on the purpose and use of the Old Hanzi in order for the group to develop appropriate encoding model to the proposed Old Hanzi collection.

3.3 Contributions on the other type of Old Hanzi

The group accepts IRG N1134 as a new contribution on the other types of Old Hanzi references. Member bodies are encouraged to contribute further on the other types of Old Hanzi, namely Bronze inscription (金文 from 殷 to 周 or 先秦) and Small Seal (小篆) of Old Hanzi.

Members attended the meeting were Li Guoying, Yin Jianghong, Wang Xiaoming, Zhang Zaixing, He Zheng-an, Julie S. C. Chuang, Hsu Hsueh Jen, Selena Wei, Fang Chuan Huang, ATSUJI Tetsuji, KUDO Motoo, MORI Masashi, Xiaolin Longxiong (invisible contributor), ANAN Kanghong, Lu Qin.

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1165

Date: 2005-11-29

Doc. Type:	Member body contribution
Title:	Summary of Old Hanzi Forum Held in Taipei 2005.10.26-28
Source:	TCA
Status:	Input to IRG
Action:	To be discuss
Distribution:	IRG Members and Ideographic Experts
Reference:	
No. of pages:	2
Medium:	Electronic

On October 26-28, TCA hosted an old Hanzi forum. Experts from Taiwan and mainland China made some conclusions in the meeting. All experts agreed that a common consensus should be obtained between Taiwan and Mainland China before processing the realignment work separately.

Based on the conclusion of the form, TCA proposes the arrangement of Old Hanzi attributes plan as follow:

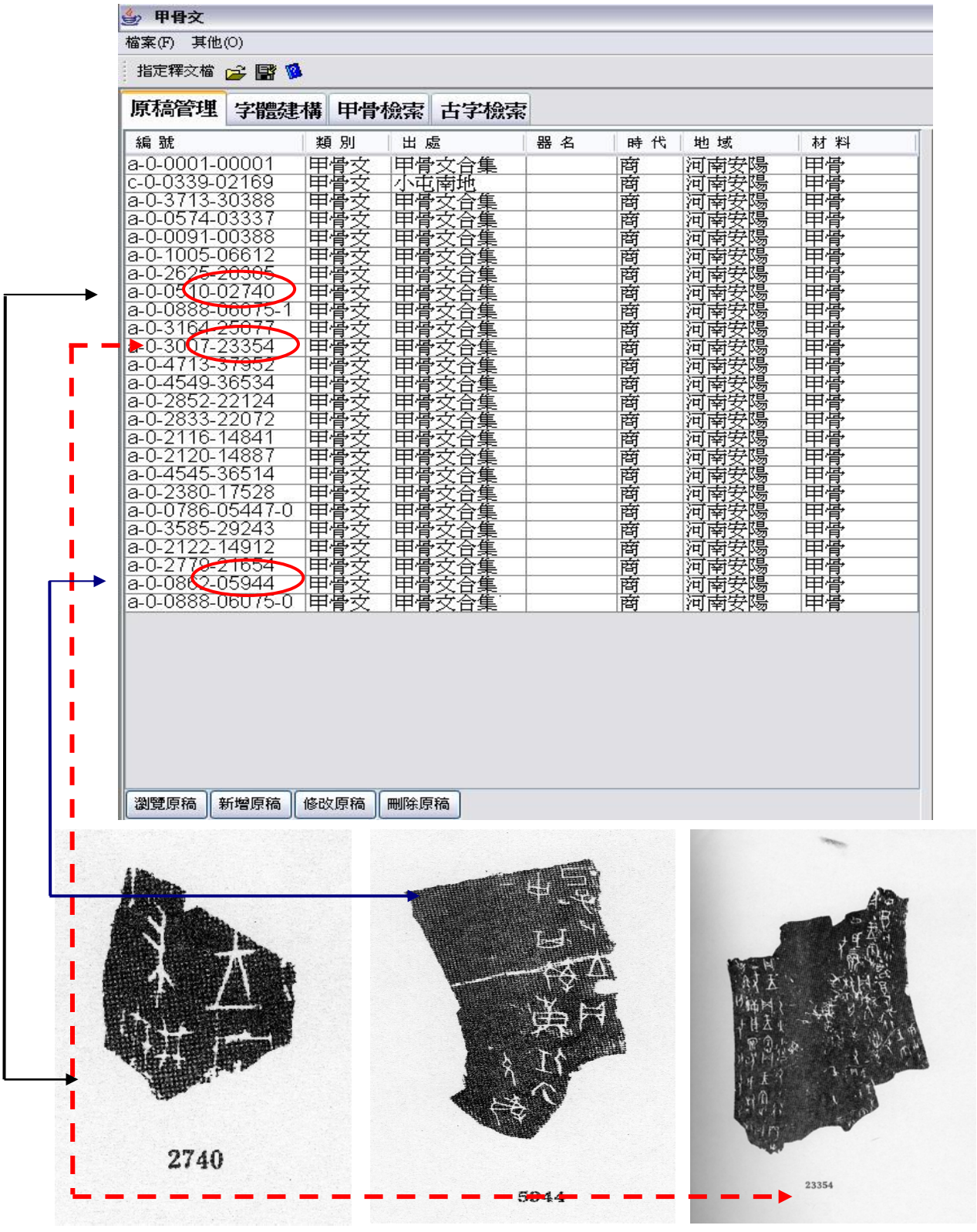
1. Start with building up (construction) an old Hanzi inscription database, including:
   (1) Original image (figure) database,

   (2) The database of glyph determination (隸定), and

   (3) Attribute database of Old Hanzi inscription.

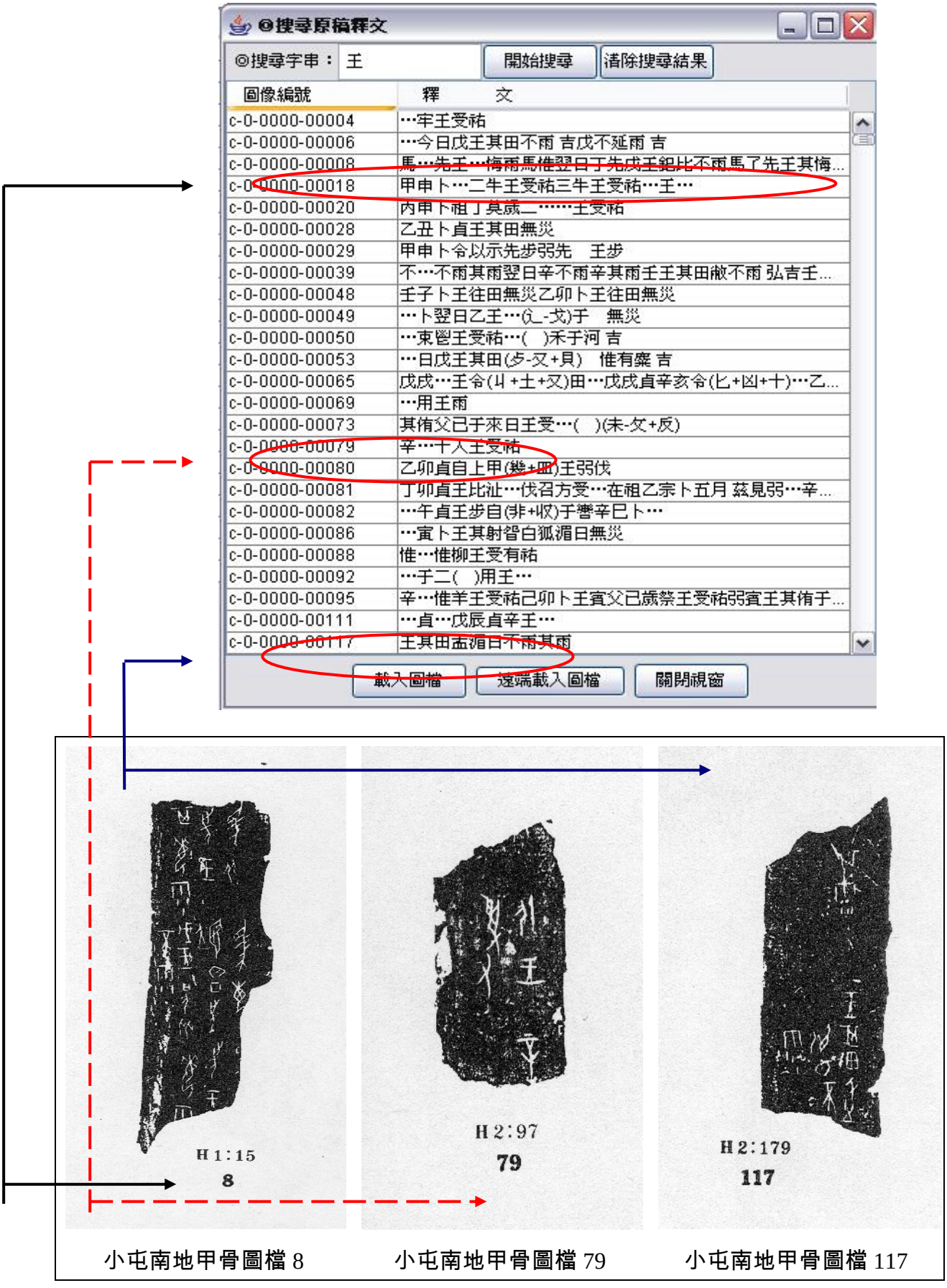
2. This Old Hanzi platform includes
   (1) The image (figure) database: Store original images of the Old Hanzi and the explanation archives . Users can refer to the original image when needed. (refer to Fig. 1 & Fig. 2)
   (2) The database of glyph determination: Store glyph determination and Corresponding Modern Character. (refer to Fig. 3)
   (3) Attributes Database: Stores the attribute data of collected Old Hanzi inscriptions. (refer to Fig. 4)

Fig. 1　the image database frame



甲骨文合集圖檔 2740　　甲骨文合集圖檔 5944　　甲骨文合集圖檔 23354

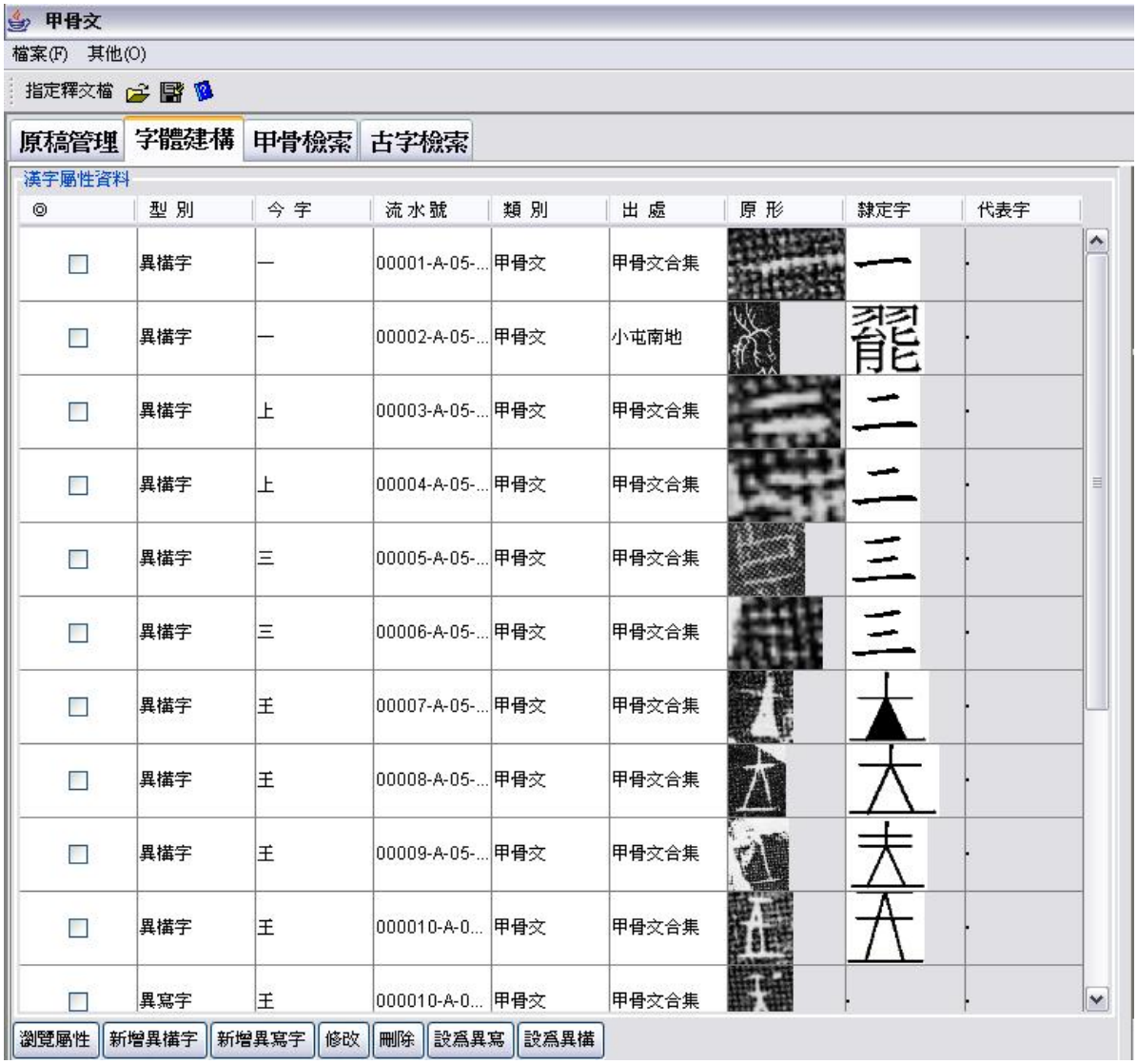
Fig. 2　the frame of using explanation archives search



小屯南地甲骨圖檔 8　　小屯南地甲骨圖檔 79　　小屯南地甲骨圖檔 117

Fig. 3　the database of glyph determination

Fig. 4　attributes database



3. The arrangement work procedure are:

(1) Build up image database: scan all oracle inscription images, and input all the explanation archives (translations) as text files, thus forming an image database.

(2) Choose glyphs: classify and choose words based on the constructed image database.

(3) Glyph Determination: arrange attributes of the chosen glyphs and proceed glyph determination, data will be stored in database. The procedure of glyph determination are as follows:

a) Proceed glyph determination to the inscriptions of the Oracle Bones refer to the comments of experts

b) Compare and analysis the structure of glyphs with the same glyph determination result, judge between construct-variance and writing-variance according to the

conclusion made in the Old Hanzi Encoding and Modern Application Conference 2005.

c) choose the most representing (complete construct, clear strokes) construct-variance and writing variance for the glyph, and store it into the database.

(4) Counting representatives: there are many words for oracle bone inscription, every word has its own construct-variance and writing-variance, the representative is the word that represents the whole group of glyphs. Hence after the attributes arrangement, counting the number of representative is necessary.

(5) Edit a wordlist: edit an oracle bone inscription wordlist, for WG2 as reference for encoding

Universal Multiple-Octet Coded Character Set

UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1168

Date: 2005-11-29

Doc. Type:	Member body contribution
Title:	Principles of Old Hanzi Selection
Source:	China
Status:	Input to IRG
Action:	
Distribution:	IRG Members and Ideographic Experts
Reference:	
No. of pages:	
Medium:	Electronic

China experts discussed the selection and the procedure of Old Hanzi with experts of Taiwan in the meeting held on Oct. 26-28, 2005 in Taiwan. Both parties agreed a consensus should be made before the processing of Old Hanzi, and a further cooperation is expected in the future work.

Chinese delegation has the following proposition:

1. The principle of Old Hanzi Selection:

The Old Hanzi should have a complete coverage and an exact map of the attributes of sources. Complete coverage is to guarantee the multiple use of Old Hanzi (research, publishing, virtual library).

To be complete, the glyphs to be selected should include all existing individual glyphs of all public materials.

Characters to be selected should include:

A. different characters

B. character with different construction

C. character with different aesthetical style

D. all components, their variants, varying writing forms

E. signs of function (repetition symbol, sentence symbol etc.)

Attributes of sources should include:

A. sources of characters to be selected.

B. sources of glyphs covered by selected characters.

2. Procedures:

To facilitate the processing of Old Hanzi, a database should be built. It should have the following platforms:

A. Contextual information

The text in which the Old Hanzi appears, the attribute of region, time, scholarly research work, and other related information. (see fig. 1)

B. Glyph information

The original image, decomposing component, representative glyph, corresponding modern character (if any), coverage index of the selected glyph. (see fig. 2)

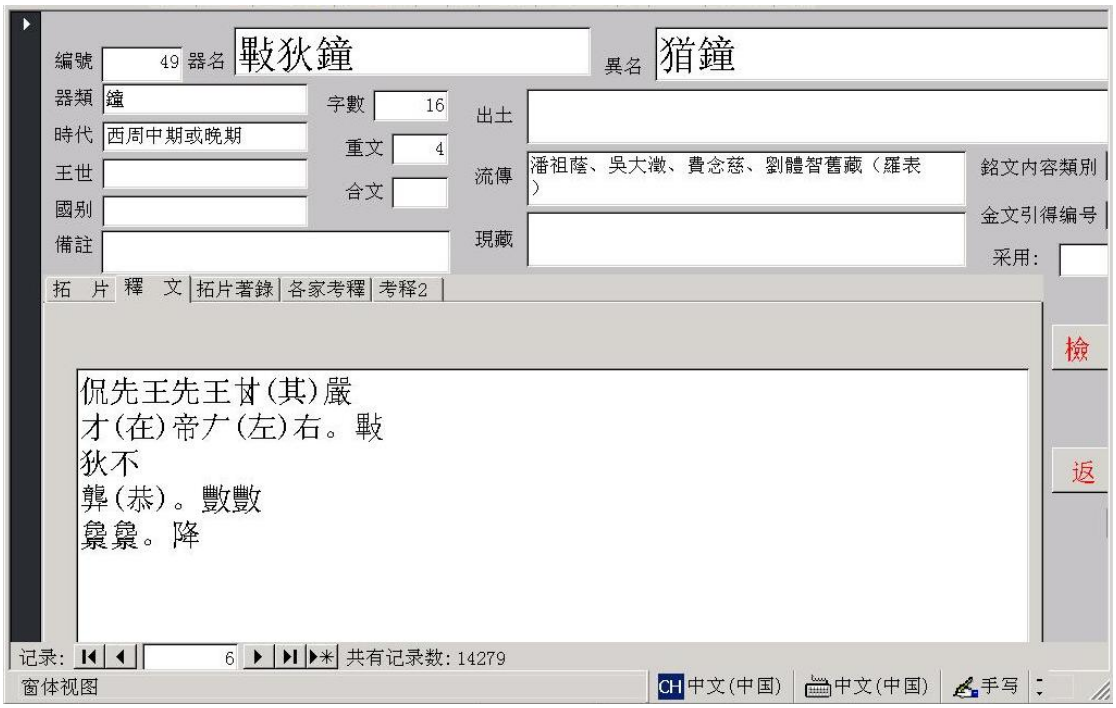


Figure 1

Figure 2

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N1182
Date: 2005-12-1

Title:	**Report from the Old Hanzi Interest Group**
Source:	**Old Hanzi Interest Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**6**
Medium:	**Electronic**

The Old Hanzi Interest Group has discussed the issues of the repertoire development of Old Hanzi. The group decided to obtain a common consensus between all the members before processing the realignment work separately, and to focus on the development of Oracle bone inscription repertoire first. The submission document will follow the N 1102 and N 1135, with one addition, the field "Unifiable Shapes" described below under "3. Format Update".

The Group discussed N1165 submitted by TCA regarding an Old Hanzi Interest Group forum on old Hanzi information format and working procedure.

The group came up with the following agreements:

1. The Old Hanzi experts group plans to arrange the Old Hanzi attributes as follows:

(1) Start with constructing an Old Hanzi script database, including these fields:

(1.1) Original image (figure)

(1.2) Glyph determination

(1.3) 9 Attributes of Old Hanzi inscription

(2) Edit a wordlist: edit an Oracle Bone inscription wordlist, for IRG consideration for submit to WG2 for information.

(3) Integrate all data into one consolidated database for easy use and make the database open to the IRG member bodies.

(4) Glyph determination: Establish a set of common rules for glyph determination in the next IRG meeting by consulting Old Hanzi experts after the IRG#25 meeting.

(5) The scope of mapping between Old Hanzi and modern ideographs：

(5.1) The mapping to modern ideographs is provided using encoded UCS characters.

(5.2) When there are mappings to both Simplified and Traditional characters, both will be listed.

2. The consolidated Old Hanzi database will include:

(1).The original images of the Old Hanzi and the explanation archives.

(2) The glyph determination and corresponding modern character.

(3) The attribute data of collected Old Hanzi inscriptions.

(Please refer to database illustrations in IRG N 1165 and N 1168.)

3. Format Update

ID	Rep. Script/ Glyph	Original Shape/ Glyph	Source	Period/ Epoch	Area/ site	Material	SW Radical	SW Radical Number	Glyph Determ.	Corresp. Modern Char	Unifiable Shapes	Notes
1												
2												

There are two changes of the format from what was defined by IRG N1135 as follows:

1. There is a new field, “Unifiable shapes”.

2. The field “Area/Terrain” is changed to “Area/site"

4. The work plan:

(1) Members are requested to submit the set of characters from Oracle Bone inscription categorized under the first 180 radicals of Shuowen Jiezi.

(2) Members should submit the images and attributes of the Oracle Bone inscriptions by 8th May, 2006.

(3) Data Format For Old Hanzi Data Exchange

For the data exchange and review work, members are going to use the data format specified as follows:

Images format:

3.1 use PNG storage format.

3.2 The original glyph should be scanned at 300 dpi (dots per inch).

3.3 The transcribed glyph images are named [ID]+[_R] (for example, if the ID is T00001, the transcribed glyph images should be named T00001_R).

3.4 The original glyphs are named [ID]+[_O] (for example, if the ID is T00001, the original glyph should be named T00001_O).

3.5 Glyph determination images are named [ID]+[_D] (for example, if the ID is T00001, the glyph determination image should be named T00001_D).

3.6 The images of unifiable shapes are named [ID]+[Unifiable shapes ID] (for example, if the Old Hanzi ID is T00001 and the unifiable shape ID is 000, the image of unifiable shape should be named T00001_000).

XML Schema:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified">
  <xs:element name="OldHanZi">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="Character" minOccurs="0"
maxOccurs="unbounded"/>
      </xs:sequence>
      <xs:attribute name="version" type="xs:string" use="required"
fixed="1.0"/>
    </xs:complexType>
  </xs:element>
  <xs:element name="Character">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="Source"/>
        <xs:element ref="Period"/>
        <xs:element ref="Area"/>
        <xs:element ref="Material"/>
        <xs:element ref="Radical"/>
        <xs:element ref="ModernChar" minOccurs="0"/>
        <xs:element ref="Unified" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element ref="Note" minOccurs="0"/>
      </xs:sequence>
      <xs:attribute name="id" use="required">
        <xs:simpleType>
          <xs:restriction base="xs:string">
            <xs:pattern value="(G|T|K|KP|J|V|S|H|M)[0-9]+"/>
          </xs:restriction>
        </xs:simpleType>
      </xs:attribute>
    </xs:complexType>
  </xs:element>
  <xs:element name="Source" type="xs:string"/>
  <xs:element name="Period" type="xs:string"/>
  <xs:element name="Area" type="xs:string"/>
  <xs:element name="Material" type="xs:string"/>
```

```
    <xs:element name="Radical">
      <xs:simpleType>
        <xs:restriction base="xs:unsignedShort">
          <xs:minInclusive value="1"/>
          <xs:maxInclusive value="540"/>
        </xs:restriction>
      </xs:simpleType>
    </xs:element>
    <xs:element name="ModernChar" type="xs:string"/>
    <xs:element name="Unified">
      <xs:complexType>
        <xs:attribute name="id" type="xs:string" use="required"/>
      </xs:complexType>
    </xs:element>
    <xs:element name="Note" type="xs:string"/>
  </xs:schema>
```

XML example:

```
  <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
  <OldHanZi version="1.0">
    <Character id="T00001">
      <Source>甲骨文合集</Source>
      <Period>商</Period>
      <Area>河南安陽</Area>
      <Material>甲骨</Material>
      <Radical>001</Radical>
      <ModernChar>一</ModernChar>
      <Unified id="0000"/>
      <Unified id="0001"/>
      <Note/>
    </Character>
    <Character id="T00002">
      <Source>甲骨文合集</Source>
      <Period>商</Period>
      <Area>河南安陽</Area>
      <Material>甲骨</Material>
      <Radical> 005</Radical>
      <ModernChar>王</ModernChar>
      <Unified id="0000"/>
```

<Note/>
</Character>
</OldHanZi>

Members attended the meeting were Dai Hong,Li Guoying, Yin Jianghong, Zhang Deshao, Shi Jianqiao, Julie S. C. Chuang, Selena Wei, Hsu Hsueh Jen ,Chen Wen-Nan, ANAN Kanghong, Lu Qin, Tom Bishop, Cook Richard.

End of document

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N1215
Date: 2006-06-08

Title:	**Report from the Old Hanzi Expert Group**
Source:	**Old Hanzi Expert Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**7**
Medium:	**Electronic**

The Old Hanzi Expert Group recapped the last report at IRG#25 meeting and discussed open issues in detail.

This report is organized as follows:

- **Revisit of the previous report**
- **Summary of discussions at this meeting**
- **Revised work plan for the development of Old Hanzi glyph repertoire.**

1. Review of the document N1182 (Old Hanzi Expert Group Report at the IRG #25 Meeting):

The Old Hanzi Expert Group has reviewed the document N1182 and reached the following agreements.

Revising N1182

(1) Item 2 of paragraph 1 in IRG N1182 should read: '(2) Compile a glyph repertoire: develop the glyph repertoire of Oracle Bone inscription to be submitted to WG2.

(2) Paragraph 3 'Format Update' has been amended as follows:

ID	Rep. Script/ Glyph	Original Shape/ Glyph	Source	Period/ Epoch	Area/ site	Material	SW Radical	SW Radical Number	Glyph Determ. (Liding)	* Corresp. Modern Char. (UCS Code)	* Unifiable Shapes	* Notes
1												
2												

There are four changes to the format:

(a) The group decided that the last three columns are 'optional' field and they are indicated with an asterisk "*", the other fields are mandatory.

(b). The field "Corresp. Modern Char." is changed to "Corresp. Modern Char. (UCS Code)".

(c). If the field "Corresp. Modern Char.(UCS Code)" is blank, then the "Notes field" must be filled with justifications to indicate the glyph is well-understood, for example, the meaning of the 'Rep. Script / Glyph'.

(d). For the field 'Corresp. Modern Char.(UCS Code)' there may be no corresponding Simplified character only corresponding Traditional characters.

2. <u>Summary of discussions:</u>

The Old Hanzi Expert Group reviewed all related comments from TCA and China, and came up with the following recommendations.

(1) The Old Hanzi Expert Group adopts TCA's 'Old Hanzi Online Editing System' as a tool for submissions, reviews and consolidation. To facilitate our work, the following proposals are made.

(a) The “Source” field consists of two data elements with one optional element. They will be concatenated with hyphen character ‘-’.

- (mandatory) The 1st letter indicates the book reference number. The possible values are:
  (A) stands for《甲骨文合集》
  (B) stands for《甲骨文合集補編》
  (C) stands for《花園莊東地甲骨》
  (D) stands for《山東濟南大辛莊甲骨》
  (E) stands for《周原甲骨》
  (F) stands for《小屯南地甲骨》

- **(**mandatory**)** Oracle Bone number (甲骨拓片的編號) which consists of 5 digits assigned uniquely to each Oracle Bone**.**

- **(optional)** Identifier to determine the side of the Oracle Bone which consists of 1 digit. The possible values are ‘0’ for front side, ‘1’ for back side. If an inscription is carved only on one side, this element will be omitted.

  Three examples of the “Source” field are listed below.

  A-00001 (does not have front and back side)

  A-00001-0 (front side)

  A-00001-1 (back side)

(b) . The “Source” field is the primary key of the data system for collation.

(c) Add a new search criterion by IRG meeting numbers to easily find the change logs.

(d) Build a website containing the above information accessible by other experts to get their feedback. The availability of the website will be subject to the agreement of the IRG Old Hanzi Expert Group..

(2). Definitions

(a) ‘Rep. Script/Glyph’: the truthful trace from ‘Original Shape/ Glyph’

[代表字的定義：忠實摹寫原形的字形。]

(b) ‘Original glyph’: the glyph selected according to the principles in the item 2) in document N1135R.

[原形的定義：根據選字原則(N1135文件)選定的字形。]

(3). The ‘Rep. Script/Glyph’ should be of standardized size and dimension.

(4). The ‘Rep. Script/Glyph’ shall be in EPS format (resolution 1024×1024).

(5). The “Glyph determination (Liding)” is a transcription of ancient script forms originally into ‘clerical-style forms’ (Li style), and today into Kai style with preservation of the shapes of the former. For this “Glyph determination (Liding)”, the extended Kai style forms will be used without limitation of known Kai strokes.

[隸定字的定義：將古漢字的原形用楷書或隸書的筆畫加以轉寫成的字形。]

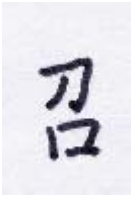
(6) “Glyph determination (Liding)” rules.

(a) Retain the number of the components of the original scripts. Examples are shown below. [保持原形構字部件的數量。]

Original Script/Glyph	Rep. Script/Glyph	Glyph determination (Liding)
		 3 components
		 8 components
		 7 components

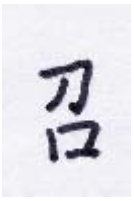
(b) Retain the relative positions and faces of the components of the original scripts. Examples are shown below.

[保持原形構字部件的相對位置及方向。]

Original Script/Glyph	Rep. Script/Glyph	Glyph determination (Liding)
An example of two inscriptions with different face of the same component. the top component in the first inscription is a mirrored image of '刀'.		
		
		
An example of two inscriptions which consists of components with different relative position.		
		
		

(c) In the transcription of Oracle Bone forms into Kai style, known Kai strokes must be used when there are equivalents in Kai style. Examples are shown below.

[甲骨文的構字部件如在楷書中已有對應，即用楷書部件形體轉寫。]

Original Script/Glyph	Rep. Script/Glyph	Glyph determination (Liding)
		

(d) The glyph should be in PNG file format (2cm × 2cm, resolutions in 300dpi)。

(7) Work plan

(a) Estimated size of the repertoire will be around 6,000 glyphs.

(b) Target completion date is IRG #32 Meeting in 2009 (three years from now)

(c) By IRG #27 Meeting, the following should be completed: radical 1 to 80 in Shuowen (Please refer to N1102 for character collation sequence principles)

- Submit glyphs selected by experts under radical 1 to 80 in Shuowen by 20 October 2006;
- Prepare PDF file for checking by all member bodies by 27 October 2006;
- Discussion by the IRG Old Hanzi Expert Group in the Meeting on 28 ~ 30 November 2006.

(d) By IRG #28 Meeting, the following should be completed: radical 81 to 160 in Shuowen

(e) By IRG #29 Meeting, the following should be completed: radical 161 to 240 in Shuowen

(f) By IRG #30 Meeting, the following should be completed: radical 241 to 320 in Shuowen

(g) By IRG #31 Meeting, the following should be completed: radical 321 to 420 in Shuowen

(h) By IRG #32 Meeting, the following should be completed: radical 421 to 540 in Shuowen

3.Appreciation

The Old Hanzi Expert Group expresses its sincere appreciation to TCA for their generous support for providing the 'Old Hanzi Online Data System' for efficient review work of the Old Hanzi Expert Group.

Members attended the meeting were Dai Hong, Li Guoying, Yin Jianghong, Shi Jianqiao, Hsu Hsueh Jen , Wei Lin Mei(Selena), Chen Wen-Nan, ANAN Kanghong, Park Jong Woo, Jung Gwang Hwa, Cheng Wai Hong(Peter), Francis Chuen.

End of document

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N1261
Date: 2006-11-26

Title:	**Old Hanzi Collection from China (Radical 1 to 80 in Shuowen Jiezi)**
Source:	**China**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**2**
Medium:	**Electronic**

This submission document followed the IRG N1215.

Documents submitted to the IRG #27 by China consist of two parts:

1. The list of the Oracle Bone Inscription. The list includes 1011 glyph.

(1) A PDF file, for Oracle Bone Inscription for reviewing (see Appendix A).

The PDF file is making up to approximately 26MB.

(2) A zip file (a compressed data file), contains individual images of Old Hanzi (Radical 1 to 80 in Shuowen Jiezi) in xml format for data exchange (see Appendix B).

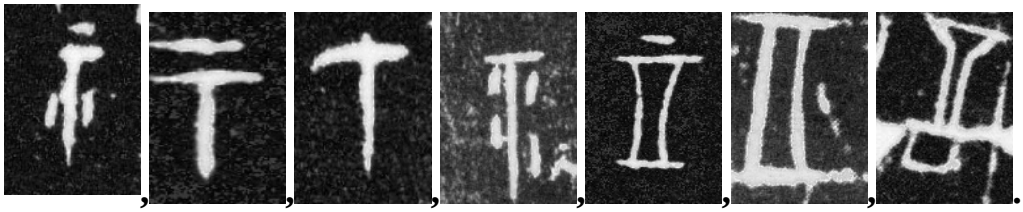
The zip file is making up to approximately 30MB.

2. Some issues need to be discussed.

(1) One glyph occupies one position in the list only. The form, not usage, form basis for our arrangement the position of the glyph. For example,

used as “茵” in the oracle inscription, this glyph is arranged in “口” of the radical of Shuowen Jiezi only, not in “艸”。

(2) If one character has more than one glyph , the sequence of these glyph in the list is that the glyph of same with or similar to the small seal of Shuowen Jiezi is arranged in the first, others are arranged according to the principle of similarity. For example, the glyph of the character “示” are arranged as follow:



(3) References of the Oracle Bone Inscription add up to eight books from six books. Additional References of the Oracle Bone Inscription are:

①《英国所藏甲骨集》;

②《怀特氏等所藏甲骨文集》。

End of document

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1267

Date: 2006-11-30

Title:	**Report from the Old Hanzi Expert Group**
Source:	**Old Hanzi Expert Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	
Medium:	**Electronic**

The Old Hanzi Expert Group recapped the last report at IRG#26 meeting and discussed open issues in details.

This report is organized as follows:

- **Revisit of the previous report**
- **Summary of discussion at this meeting**
- **Revised work plan for the development of Old Hanzi inscriptions repertoire.**

1. Review of the document N1215 (Old Hanzi Expert Group Report at the IRG #26 Meeting):

The Old Hanzi Expert Group has reviewed the document N1215 and reached the following agreements.

Revising the document N1215.

Increase the number of reference materials on Oracle Bone Inscriptions.

The collection of the Oracle Bone Inscriptions will be increased to 11 books. The additional five books are listed below:

補充整理甲骨文字集之選用材料增加五本，分別如下：

(G) 英國所藏甲骨集

(H) 懷特氏等所藏甲骨文集

(I) 天理大學附屬天理參考館藏甲骨文字

(J) 德瑞荷比所藏一些甲骨錄

(K) 瑞典斯德哥爾摩遠東古物博物館藏甲骨文字

2. Summary of discussion in this meeting

2.1 The principles of radical classification are listed below:

(1) The glyph should be classified into Shuowen radicals according to the Original Oracle Bone Inscriptions, e.g. “斷” should be classified under radical 斤.

歸部問題：以甲骨文字形為主，如“斷”入斤部。

Example 1(範例一)

Rep. Script/Glyph	Original Shape/Glyph	S.W. Radical
		斤

(2) the shapes of the Original Oracle Bone inscriptions are different, but they share the same meaning and usage. Eventually, they have evolved into a pair of variants , with two different radicals in Shuowen. According to the radical classification in Shuowen, the characters are put under different radicals. e.g. 兀 and 元.

甲骨文異形同用，後世分為兩字，說文分見兩部，則依《說文解字》收入不同部首。如“元”、“兀”。

Example 2 (範例二)

Rep. Script/Glyph	Original Shape/Glyph	S.W. Radical
		一
		儿

(3) the shape, meaning and usage of the Original Oracle Bone inscriptions are identical but new components have been added over time. If a radical can be found in Shuowen, the characters will be put under the corresponding radical according to the Original Oracle Bone inscriptions. e.g. 枭 put under the 枭 radical and similarly 畐 in 畐 radical.

甲骨文同形同用，後世增添偏旁，《說文解字》另有部首者，則依甲骨文原形歸入相應部首。如“畐”入《說文解字》畐部，“㚔”入《說文解字》㚔部。

Example 3 (範例三)

Rep. Script/Glyph	Original Shape/Glyph	S.W. Radical
		畐
		㚔

(4) the shapes of the Original Oracle Bone inscriptions are the same but they have many meanings and usages. Eventually, they have evolved into different characters. The glyph of these characters will be determined according to the shapes of the Original Oracle Bone inscriptions, and put under the corresponding radical in Shuowen.

e.g. 史吏事.

甲骨文同形多用，後世分為多字，依甲骨文字形分別隸定，歸入《說文解字》相應同一個部首。如：“史”、“吏”、“事”。

Example 4(範例四)

Rep. Script/Glyph	Original Shape/Glyph	S.W. Radical
		一

		一

2.2 The principles of sorting the order of the glyphs of the same Oracle Bone Inscription
同字之字形排序原則

(1) For Oracle Bone Inscriptions that are similar with Small Seal, Li style and Kai style font, they will be placed first .
甲骨文字形與小篆、隸書和楷書相合者(接近，相似)，放在首位。

(2) If one or more types of components or radical are different, then those with smaller difference will be placed first, and those with greater difference behind.
異構字依字形差異大小排序，差異較小者置於前，差異較大者置於後。

(3) Glyph variants will be placed after the main glyph.
異寫字置於主形之後。

3. Work plan

(1)An ad hoc meeting before March 2007 is needed in view of the very heavy workload at this stage of character table submission and examination.

(2) By 28 March 2007, complete the revisions of and amendments to the Oracle Bone inscriptions character table discussed at IRG#27

(3) By 28 March 2007, complete the Oracle Bone inscriptions of the first 100 radicals in Shuowen by all interested parties and submit the results to Prof. Li Guoying.

(4) By 10 April 2007, Selina Wei finishes the consolidation of all the

submissions provided by all interested parties and email to all members of the Old Hanzi Expert Group.

(5) From 11 April to 20 May 2007, examination of the Oracle Bone inscriptions character table.

(6) Before IRG#28, submit the commented Oracle Bone inscriptions character table to IRG Rapporteur.

Appendix: A combined character table of Oracle Bone Inscriptions

Members attended the meeting were Dai Hong, Li Guoying, Wang Ning, Zhao Cheng, Wang Yunzhi, Zhou Xiaowen, Zhang Zaixing, Yin Jianghong, Shi Jianqiao,Wei Lin Mei(Selena), Hsu Hsueh Jen , Yuen Kwok-wa, Chen Wen-Nan, Cheng Wai Hong, AU Yung-yee.

End of document

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1271

Date: 2006-12-13

Title:	**Old Hanzi Principles and References (Version 2)**
Source:	**Old Hanzi Expert Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**15**
Medium:	**Electronic**

The Old Hanzi principles and references (version 2) extracted and compiled from the following documents:

IRG N1135R, IRG N1182, IRGN1215, IRG N1267.

1. Format of submission

ID	Rep. Script/ Glyph	Original Shape/ Glyph	Source	Period/ Epoch	Area/ site	Material	SW Radical	SW Radical Number	Glyph Determ.	* Corresp. Modern Char (UCS Code)	* Unifiable Shapes	* Notes
1												
2												

2. Definitions

2.1 ID: It is the unique id that consists of one or two letters member id (G, T, K, KP, J, V, S, H, M) followed by four digit sequential numbers assigned by submitters.

Example: T0001 is one IRG global unique ID assigned to an Old Hanzi submitted by TCA.

2.2 Rep. Script/Glyph: The truthful trace from 'Original Shape/ Glyph'

[代表字的定義：忠實摹寫原形的字形。]

2.3 Original glyph: The glyph selected according to the principles of Old Hanzi selection (in the item 4) in this document.

[原形的定義：根據選字原則(本文件第 3 章)選定的字形。]

2.4 SW Radical: ShuoWen Radical

2.5 SW Radical number : 1 – 540

2.6 The "Glyph determination (Liding)": It is a transcription of ancient script forms originally into 'clerical-style forms' (Li style), and today into Kai style with preservation of the shapes of the former. For this "Glyph determination (Liding)", the extended Kai style forms will be used without limitation of known Kai strokes.

[隸定字的定義：將古漢字的原形用楷書或隸書的筆畫加以轉寫成的字形。]

3. Rules

3.1 The 'Rep. Script/Glyph' should be of standardized size and dimension.

3.2 The 'Rep. Script/Glyph' shall be in EPS format (resolution 1024×1024).

3.3 The format of bitmap images for rep. glyphs, original shapes and 540 SW Radicals are specified as follows:

128x128，Black and white bitmap

3.4 The last three columns are ‘optional’ field and they are indicated with an asterisk “*”, the other fields are mandatory.

3.5 If the field “Corresp. Modern Char.(UCS Code)” is blank, then the “Notes field” must be filled with justifications to indicate the glyph is well-understood, for example, the meaning of the ‘Rep. Script / Glyph’.

3.6 For the field ‘Corresp. Modern Char.(UCS Code)’ there may be no corresponding Simplified character only corresponding Traditional characters.

3.7 Glyph determination (Liding)

(a) Retain the number of the components of the original scripts. Examples are shown below.

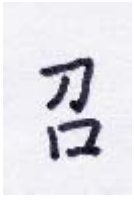
[保持原形構字部件的數量。]

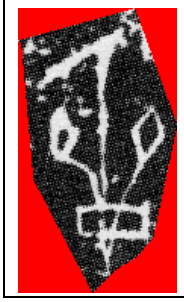
Original Script/Glyph	Rep. Script/Glyph	Glyph determination (Liding)
		 3 components
		 8 components

		 7 components

(b) Retain the relative positions and faces of the components of the original scripts. Examples are shown below.

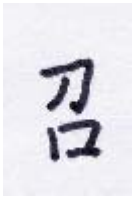
[保持原形構字部件的相對位置及方向。]

Original Script/Glyph	Rep. Script/Glyph	Glyph determination (Liding)
An example of two inscriptions with different face of the same component. the top component in the first inscription is a mirrored image of '刀'.		
		
		
An example of two inscriptions which consists of components with different relative position.		

(c) In the transcription of Oracle Bone forms into Kai style, known Kai strokes must be used when there are equivalents in Kai style. Examples are shown below.

[甲骨文的構字部件如在楷書中已有對應，即用楷書部件形體轉寫。]

Original Script/Glyph	Rep. Script/Glyph	Glyph determination (Liding)
		

(d) The glyph should be in PNG file format (2cm × 2cm, resolutions in 300dpi)。

3.8 Source: The “Source” field is the primary key of the data system for collation.

The “Source” field consists of two data elements with one optional

element. They will be concatenated with hyphen character '-'.

- (mandatory) The 1st letter indicates the book reference number. The possible values are:

  (A) stands for《甲骨文合集》

  (B) stands for《甲骨文合集補編》

  (C) stands for《花園莊東地甲骨》

  (D) stands for《山東濟南大辛莊甲骨》

  (E) stands for《周原甲骨》

  (F) stands for《小屯南地甲骨》

  (G) stands for《英國所藏甲骨集》

  (H) stands for《 懷特氏等所藏甲骨文集》

  (I) stands for《天理大學附屬天理參考館藏甲骨文字》

  (J) stands for《德瑞荷比所藏一些甲骨錄》

  (K) stands for《瑞典斯德哥爾摩遠東古物博物館藏甲骨文字》

- (mandatory) Oracle Bone number (甲骨拓片的編號) which consists of 5 digits assigned uniquely to each Oracle Bone.

- (optional) Identifier to determine the side of the Oracle Bone which consists of 1 digit. The possible values are '0' for front side, '1' for back side. If an inscription is carved only on one side, this element will be omitted.

Three examples of the "Source" field are listed below.

- A-00001 (does not have front and back side)
- A-00001-0 (front side)
- A-00001-1 (back side)

4. The principles of Old Hanzi selection

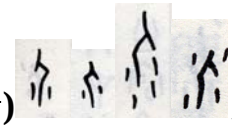
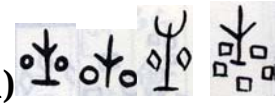
4.1 Distinction principles

Two or more instances of Old Hanzi are considered different from each other when either one of the following conditions apply:

4.1.1 One or more types of components are different.

(莫)  ，**(示)** 

4.1.2 The number of components or lines is different.

(介)  ，**(星)** 

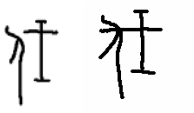
4.1.3 The direction (e.g. mirror image) of a component is different.

(人)  ，**(師)**  ，**(至)** 

4.1.4 The position of one or more components is different.

(好) 

4.1.5 Whether the same set of components are connected each other or not.

(伐) 

4.1.6 One or more line types (straight line, curve, circle, rectangle, closed line or curve filled inside) are different.

(王) 

4.2 Unification principles

Two or more instances of Old Hanzi with the following differences will be unified unless there's any difference in the meaning:

4.2.1 The length of corresponding line is different.

4.2.2 The thickness of corresponding line is different.

4.2.3 The size of each component of the same set is different.

5. The principles of radical classification

5.1 The glyph should be classified into Shuowen radicals according to the Original Oracle Bone Inscriptions, e.g. "斷" should be classified under radical 斤.
歸部問題：以甲骨文字形爲主，如"斷"入斤部。

Example 1(範例一)

Rep. Script/Glyph	Original Shape/Glyph	S.W. Radical
		斤

5.2 The shapes of the Original Oracle Bone inscriptions are different, but they share the same meaning and usage. Eventually, they have evolved into a pair of variants , with two different radicals in Shuowen. According to the radical classification in Shuowen, the characters are put under different radicals. e.g. 兀 and 元.

甲骨文異形同用，後世分爲兩字，說文分見兩部，則依《說文解字》收入不同部首。如“元”、“兀”。

Example 2 (範例二)

Rep. Script/Glyph	Original Shape/Glyph	S.W. Radical
		一
		儿

5.3 The shape, meaning and usage of the Original Oracle Bone inscriptions are identical but new components have been added over time. If a radical can be found in Shuowen, the characters

will be put under the corresponding radical according to the Original Oracle Bone inscriptions. e.g. 彔 put under the 彔 radical and similarly 畐 in 畐 radical.

甲骨文同形同用，後世增添偏旁，《說文解字》另有部首者，則依甲骨文原形歸入相應部首。如“畐”入《說文解字》畐部，“彔”入《說文解字》彔部。

Example 3 (範例三)

Rep. Script/Glyph	Original Shape/Glyph	S.W. Radical
		畐
		彔

5.4 The shapes of the Original Oracle Bone inscriptions are the same but they have many meanings and usages. Eventually, they have evolved into different characters. The glyph of these characters will be determined according to the shapes of the Original Oracle Bone inscriptions, and put under the corresponding radical in Shuowen. e.g. 史吏事.

甲骨文同形多用，後世分爲多字，依甲骨文字形分別隸定，歸入《說文解字》相應同一個部首。如：“史”、“吏”、“事”。

Example 4(範例四)

Rep. Script/Glyph	Original Shape/Glyph	S.W. Radical
		一
		一

6. The principles of sorting the order of the glyphs of the same Oracle Bone Inscription

同字之字形排序原則

6.1 For Oracle Bone Inscriptions that are similar with Small Seal, Li style and Kai style font, they will be placed first .

甲骨文字形與小篆、隸書和楷書相合者(接近，相似)，放在首位。

6.2 If one or more types of components or radical are different, then those with smaller difference will be placed first, and those with greater difference behind.

異構字依字形差異大小排序，差異較小者置於前，差異較大者置於後。

6.3 Glyph variants will be placed after the main glyph.

異寫字置於主形之後。

7. References of the Oracle Bone Inscription

The collection of the Oracle Bone Inscription will be developed to cover the following 11 books:

1. 《甲骨文合集》
2. 《甲骨文合集补编》
3. 《花园庄东地甲骨》
4. 《山东济南大辛庄甲骨》
5. 《周原甲骨》
6. 《小屯南地甲骨》
7. 《英國所藏甲骨集》
8. 《懷特氏等所藏甲骨文集》
9. 《天理大學附屬天理參考館藏甲骨文字》
10. 《德瑞荷比所藏一些甲骨錄》
11. 《 瑞典斯德哥爾摩遠東古物博物館藏甲骨文字》

8. Data Format For Old Hanzi Data Exchange

For the data exchange and review work, members are going to use the data format specified as follows:

- **Images format:**

  8.1 use PNG storage format.

  8.2 The original glyph should be scanned at 300 dpi (dots per inch).

8.3 The transcribed glyph images are named [ID]+[_R] (for example, if the ID is T00001, the transcribed glyph images should be named T00001_R).

8.4 The original glyphs are named [ID]+[_O] (for example, if the ID is T00001, the original glyph should be named T00001_O).

8.5 Glyph determination images are named [ID]+[_D] (for example, if the ID is T00001, the glyph determination image should be named T00001_D).

8.6 The images of unifiable shapes are named [ID]+[Unifiable shapes ID] (for example, if the Old Hanzi ID is T00001 and the unifiable shape ID is 000, the image of unifiable shape should be named T00001_000).

■ XML Schema:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified">
  <xs:element name="OldHanZi">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="Character" minOccurs="0"
maxOccurs="unbounded"/>
      </xs:sequence>
      <xs:attribute name="version" type="xs:string" use="required"
fixed="1.0"/>
    </xs:complexType>
  </xs:element>
  <xs:element name="Character">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="Source"/>
```

```
            <xs:element ref="Period"/>
            <xs:element ref="Area"/>
            <xs:element ref="Material"/>
            <xs:element ref="Radical"/>
            <xs:element ref="ModernChar" minOccurs="0"/>
            <xs:element ref="Unified" minOccurs="0" maxOccurs="unbounded"/>
            <xs:element ref="Note" minOccurs="0"/>
          </xs:sequence>
          <xs:attribute name="id" use="required">
            <xs:simpleType>
              <xs:restriction base="xs:string">
                <xs:pattern value="(G|T|K|KP|J|V|S|H|M)[0-9]+"/>
              </xs:restriction>
            </xs:simpleType>
          </xs:attribute>
        </xs:complexType>
      </xs:element>
      <xs:element name="Source" type="xs:string"/>
      <xs:element name="Period" type="xs:string"/>
      <xs:element name="Area" type="xs:string"/>
      <xs:element name="Material" type="xs:string"/>
      <xs:element name="Radical">
        <xs:simpleType>
          <xs:restriction base="xs:unsignedShort">
            <xs:minInclusive value="1"/>
            <xs:maxInclusive value="540"/>
          </xs:restriction>
        </xs:simpleType>
      </xs:element>
      <xs:element name="ModernChar" type="xs:string"/>
      <xs:element name="Unified">
        <xs:complexType>
          <xs:attribute name="id" type="xs:string" use="required"/>
        </xs:complexType>
      </xs:element>
      <xs:element name="Note" type="xs:string"/>
    </xs:schema>
```

■ XML example:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<OldHanZi version="1.0">
   <Character id="T00001">
      <Source>甲骨文合集</Source>
      <Period>商</Period>
      <Area>河南安陽</Area>
      <Material>甲骨</Material>
      <Radical>001</Radical>
      <ModernChar>一</ModernChar>
      <Unified id="0000"/>
      <Unified id="0001"/>
      <Note/>
   </Character>
   <Character id="T00002">
      <Source>甲骨文合集</Source>
      <Period>商</Period>
      <Area>河南安陽</Area>
      <Material>甲骨</Material>
      <Radical> 005</Radical>
      <ModernChar>王</ModernChar>
      <Unified id="0000"/>
      <Note/>
   </Character>
</OldHanZi>
```

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1304
Date: 2007-05-30

Title:	**Consolidated Old Hanzi Submission (Radical 65 to 100 in Shouwen)**
Source:	**IRG N1303(TCA), IRGN 1307(China)**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**1**
Medium:	**Electronic**

This submission document is according as the consolidation of two documents. Both of the documents are TCA IRG N1303 and China IRG N1307.

The method of integrating the data from these two documents is using some ordinal key fields to sort all the data in one format. First of all, using the filed of SW Radical Number to group all the relate records. If there are too many records relate to one SW Radical Number, and then will use the second field which is Corresp. Modern Character, to sort again. Finally, if after using the above two keys still couldn't organize all the records into a fine structure, then will use "ID" field to be the last way to order the rest of the records.

Documents submitted to the IRG #28 Meeting are listed below:

The electronic files are a PDF file, for Oracle Bone Inscription for reviewing (IRGN1304-Appendix A1-Consolidated Old Hanzi Submission).

End of document

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1325

Date: 2007-06-07

Title:	**Report from the Old Hanzi Expert Group**
Source:	**Old Hanzi Expert Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**4**
Medium:	**Electronic**

The Old Hanzi Expert Group recapped the last report at IRG#27 meeting and discussed open issues in details.

This report is organized as follows:

- **Revisit of the previous report for new members.**
- **Summary of discussion at this meeting.**
- **Revised work plan for the development of Old Hanzi inscriptions repertoire.**

1. **Review of the document N1267 (Old Hanzi Expert Group Report at the IRG #27 Meeting)(no change)**
2. **Review of the document N1271 (Old Hanzi Principles and References Version 2)**

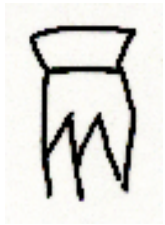
The Old Hanzi Expert Group has reviewed the document N1271 and reached the following agreements.

Revising the document N1271.

To consider adding one principle regrding glyph determination (Liding) for Oracle Bone Inscriptions to clarify how the glyph determination (Liding) should be made when the shape of Liding cannot be decided based on the Original Script/Glyph of the Oracle Bone Inscriptions, but rather based on the context**”**. Details of this principle will be finished at the next Old Hanzi Group meeting.

甲骨文原形無法據其形隸定者，其隸定原則於下次會議討論。

Example 1: Glyph determination according to the original script is not possible 無法據其形隸定

Original Shape/Glyph	Rep. Script/Glyph	Glyph Determ. (Liding)
		
		

Example 2: Glyph determination according to the original script is possible 可據其形隸定

Original Shape/Glyph	Rep. Script/Glyph	Glyph Determ. (Liding)

		启

3. Summary of discussion in this meeting

3.1 According to the principles in document N1271, China and TCA both have submitted SW Radical 1-100 of Oracle Bone Inscriptions.

依 N1271 原則，中國與台灣已提交甲骨文字形之說文部首 1- 100。

3.2 Old Hanzi Expert Group has finished the discussion of SW Radical 1 to 3 and 65-100 of Oracle Bone Inscriptions.

古漢字專家工作組已討論完成說文部首 1-3 和 65-100。

3.3 SW Radical 4 to 64 of Oracle Bone Inscriptions will be discussed at the next meeting.

說文部首 4-64 之甲骨文字形，將於下次古漢字專家工作組中討論。

3.4 The Old Hanzi Expert Group has decided to advance its meeting in mid September in Harbin (tentative), and submit its discussion results before IRG#29. The Old Hanzi Expert Group is sorry that no meeting can be held in San Jose, USA due to budgetary consideration and other work commitments.

All other parties interested in our work are welcome to the meeting to be held in Harbin.

由於經費和時間的問題，古漢字專家工作組會議將無法於美國 SanJose 舉行。古漢字專家工作組一致決議，九月中旬(暫定)於哈爾濱提前召開古漢字專家工作組會議，並在 IRG #29 次會議前提交工作組成果報告。

4. Work plan

(1) By the middle of July 2007, complete the revisions of and amendments to the Oracle Bone inscriptions character table corresponding to the 1-100 radicals in Shuowen discussed at IRG#28.

(2) By 31 July 2007, complete the Oracle Bone inscriptions of the 101-140 radicals in Shuowen by all interested parties and submit the results to Prof. Li Guoying.

7 月 31 日 提交 SW 部首 101-140 之甲骨文。

(3) By 8 August 2007, Wei Lin-Mei (Selina) finishes the consolidation of all the submissions of the Oracle Bone inscriptions corresponding to the 101-140 radicals in Shuowen provided by all interested parties and email to all members of the Old Hanzi Expert Group.

8 月 8 日整合完成並返回，開始審查之工作。

(4) From 9 to 27 August 2007, examination of the Oracle Bone inscriptions character table corresponding to the 101-140 radicals in Shuowen and submit the review report for discussion.

8 月 9-27 日審查甲骨文字表，並於 27 日相互提交審查意見。

(5) The Old Hanzi Expert Group will hold its meeting from 10 to 13 September 2007 in Harbin or other place.

(6) Before IRG#29, submit the commented Oracle Bone inscriptions character table from SW Radicals 1 to 140 to the IRG Rapporteur.

Appendix: A combined character table of Oracle Bone Inscriptions

Members attended the meeting were Dai Hong, Li Guoying, Zhao Cheng, Zhou Xiaowen, Shi Jianqiao, Wang Danhui, Wei Lin Mei(Selena), Hsu Hsueh Jen , Yuen Kwok-wa, Cheng Wai-hong (Peter), Matsuoka Eiji and Yasuhiro Anan.

End of document

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1337

Date: 2007-11-6

Title:	**Report from the Old Hanzi Expert Group**
Source:	**Old Hanzi Expert Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**7**
Medium:	**Electronic**

The Old Hanzi Expert Group held a group meeting in Harbin, China on 13th -15th of September, 2007 in prior to IRG#29. And recapped the last report at IRG#28 meeting and discussed open issues in details.

This report is organized as follows:

- **Summary of discussion at this meeting**
- **Revised work plan for the development of Old Hanzi inscriptions repertoire.**

1. **Review of the document N1325 (Old Hanzi Expert Group Report at the IRG #28 Meeting) (no change)**
2. **Review of the document N1336 (Old Hanzi Principles and References Version 3).**
3. **Summary of discussion in this meeting**

Agenda was set on the proposal (N1346) brought up by Japanese delegacy, and Shuowen radical 101-180 of Oracle **B**one inscriptions that were submitted by China and Taiwan respectively, according to the work project decided at the IRG28 Meeting.

Japanese delegacy's proposal N1346 was discussed in the meeting. In the proposal it stated that the foundation of ordering should be established on Yin-hsu chia-ku ko-tzu lei-chuan《殷墟甲骨刻辭類纂》, edited by Yao Hsiao-sui(姚孝遂), which was developed from Yin-hsu pu-tzu tsung-lei《殷墟卜辭綜類》edited by Shima Kunio(島邦男) in Japan. In fact, the ordering method applied by the above two books is called "Natural Classification". It is invented by Tang Lan(唐蘭), a researcher of Chinese ancient characters (Shima Kunio mentioned this point in his book). Tang Lan did not push "Natural Classification" to a full length because he is aware that there are some drawbacks to use natural classification for radical classification of Oracle Bone inscriptions. Examples are listed below.

Example 1

"have"「有」：  甲骨文 1289、  甲骨文《殷契粹編》一三、
 Jin-Wen 金文《盂鼎》、
 Warring Kingdoms inscriptions 戰國文字《石經》、
 Warring Kingdoms inscriptions 戰國文字-簡帛《信陽楚簡》、
 Small Seal 小篆《說文》

Example 2

"morning"「晨」：  甲骨文《殷契粹編》二五一、
 Warring Kingdoms inscriptions 戰國文字-簡帛《楚帛書》、

 Warring Kingdoms inscriptions 戰國文字《古璽》、
 Small Seal 小篆《說文》

Hanzi has had a history of over three thousand years since Oracle Bone inscriptions. No matter Oracle Bone inscriptions, Jin-Wen, Warring Kingdoms inscriptions, Big Seal (including Zhou-Wen), Small Seal, Li style or Kai style, there exists a link of inheritance and they can not be divided without consideration. If one has enough understanding on the development of Hanzi it would not be hard to find out from the above examples that if using “Natural Classification” characters with identical sound or meaning but different shape will be scattered and are rather difficult to be unified. Furthermore, applying “Natural Classification” is insufficient. For example, in Yin-hsu chia-ku ko-tzu lei-chuan《殷墟甲骨刻辭類纂》there are more than one hundred characters that can not be classified into any radicals.

After thorough discussion at the meeting and we think in the field of Oracle Bone inscriptions there are always two different methods of ordering, Natural Classification or Shuowen. Both methods hold advantages and disadvantages. Natural Classification is more convenient for glyph search of Oracle Bone inscriptions but it has the disadvantage of unable to group variant Chinese characters, and it also can not be applied by the same classification method like some of the ancient characters Jin-Wen, Warring Kingdoms inscriptions and Small Seal. If ordering is done according to Shuowen, it is easy to group variant Chinese characters and can be applied by the same classification method like some of the ancient characters Jin-Wen, Warring Kingdoms inscriptions and Small Seal. However there is a drawback, ordering for characters not included in Shuowen has to be handled in a special way. IRG’s job on the classification of Old Hanzi covers a complete

classification on Oracle Bone inscriptions, Jin-Wen, Warring Kingdoms inscription and Small Seal. It is better to choose identical ordering method, and such objective can be reached only by following Shuowen ordering. Thus, in the meeting agreement was reached unanimously to continue following IRG's confirmed ordering principle, which means to use Shuowen ordering method to process Old Hanzi ordering including Oracle Bone inscriptions.

In the meeting, discussion was completed on Shuowen radical 101-180 of Oracle Bone inscriptions that were submitted by China(IRGN 1344) and Taiwan(IRGN 1343) respectively. Thus, it was closed according to plan.

4. Work plan

(1) Next work project was also discussed in the meeting, and decided to complete Shuowen radical 181-260 of Oracle Bone inscriptions in next IRG meeting.

(2) By 28 March 2008, complete the Oracle Bone inscriptions of the 180 -260 radicals in Shuowen by all interested parties and submit the results to Prof. Li Guoying.

(3) By 10 April 2008, Wei Lin-Mei(Selena) finishes the consolidation of all the submissions provided by all interested parties and email to all members of the Old Hanzi Expert Group.

(4) From 11 April to IRG#30 2008, examination of the Oracle Bone inscriptions character table.

(5) Before IRG#30, submit the commented Oracle Bone inscriptions

character table to IRG Rapporteur.

Appendix: A combined character table of Oracle Bone Inscriptions

Members attended the meeting were Li Guoying, Zhao Cheng, Dong Lianchi, Zhou Xiaowen, Chen Shuangxin, Zhao Fu, Wei Lin-Mei(Selena), Jung Bor-Sheng(鍾柏生), Yuen Kwok-wa(袁國華), Wang Xiao, Cao Ying .

End of document

29 次會議古漢字決議（中文稿）

2007 年 9 月 13 日至 9 月 15 日古漢字專家組在中國哈爾濱市預先召開了 IRG#29 會議的小組會議。

會議收到日本代表團提出的提案（N1346）、中國和臺灣按照 IRG28 會議的決議確定的工作計畫提交的《說文解字》部首 101-180 的甲骨文字形。

會議討論了日本代表團提出的提案（N1346）。提案提出，應該把在日本島邦男《殷墟卜辭綜類》的基礎上發展起來由姚孝遂主編的《殷墟甲骨刻辭類纂》作爲排序的基礎。其實以上兩書採用的排序方法，叫做「自然分類法」，是由中國古文字學家唐蘭先生所發明（島邦男書中已言之），唐蘭先生鑑於以「自然分類法」將甲骨文分部，並非全無缺點，故此沒有大力提倡。舉例說明如下：

字例 1「有」：  甲骨文 1289、  甲骨文《殷契粹編》一三、  金文《盂鼎》、  戰國文字《石經》、  戰國文字-簡帛《信陽楚簡》、  小篆《說文》

字例 2「晨」：  甲骨文《殷契粹編》二五一、戰國文字-簡帛《楚帛書》、  戰國文字《古璽》、  、小篆《說文》

漢字自甲骨文以來，已有超過三千的歷史，無論甲骨文、金文、戰國文字、大篆（包括籀文）、小篆、隸書，以及楷書，都有一脈相承、承先啓後的關係，不宜任意切割。對於漢字發展有足夠了解的話，便不難從上舉字例發現，如果採用「自然分類法」，音、義相同而字形不同的字，便會散置多處，難以統合。更何況即使使用「自然分類法」亦存在不足之處，以《殷墟甲骨刻辭類纂》爲例，至少有一百個以上的字還是無法排入任何部首的。

會議經過充分的討論認爲，甲骨學界歷來存在自然分類和按《說文解字》排序兩種不同的排序方法，兩種方法各有優缺點。自然分類法便於甲骨文自身的據形檢索，缺點是異體字不能類聚在一起，同時不能和金文、戰國文字、小篆等古文字採用相同的分類方法。按《說文解字》排序便於異體字的類聚，且可以和金文、戰國文字、小篆等古文字採用相同的分類方法，缺點是有個別《說文解字》未收的字排序需要特殊處理。IRG 的古漢字整理工作是要對包括甲骨文、金文、戰國文字、小篆在內的古漢字全部處理，最好的選擇是採用相同的排序方法，只有按《說文解字》排序才能實現這個目標。會議一致同意，仍舊按 IRG 已經確定的排序原則，即按《說文解字》排序處理包括甲骨文在內的古漢字排序。

會議討論完成了由中國(IRGN 1344)和臺灣(IRGN 1343)分別提交的《說文解字》部首 101-180 的甲骨文字形，按計劃完成了收字工作。

會議討論確定了下一步的工作計畫，決定到下次 IRG 會議完成《說文解字》部首 181-260 的甲骨文字形。

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1345
Date: 2007-9-1

Title:	**Consolidated Old Hanzi Submission (Radical 101 to 180 in Shouwen)**
Source:	**IRG N1303(TCA), IRGN 1307(China)**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**1**
Medium:	**Electronic**

This submission document is according as the consolidation of two documents. Both of the documents are TCA IRG N1343 and China IRG N1344.

The method of integrating the data from these two documents is using some ordinal key fields to sort all the data in one format. First of all, using the filed of SW Radical Number to group all the relate records. If there are too many records relate to one SW Radical Number, and then will use the second field which is Corresp. Modern Character, to sort again. Finally, if after using the above two keys still couldn't organize all the records into a fine structure, then will use "ID" field to be the last way to order the rest of the records.

Documents submitted to the IRG #29 Meeting are listed below:

The electronic files are a PDF file, for Oracle Bone inscriptions for reviewing (IRGN1345-Appendix A1-Consolidated Old Hanzi Submission).

End of document

JTC 1/SC2/WG2/IRG N 1346

Date: September 09, 2007

ISO/IEC JTC 1/SC2/WG2/IRG

Ideographic Rapporteur Group

Source: Dr. SUZUKI, Atsushi

Meeting: Old Hanzi Special Group Meeting held in Harbin, China, from 2007-9-12 to 2007-9-14.

Title: Input to Old Hanzi Expert Group

Keywords: (none)

Status: Expert Contribution

Short Description:

This is a cover page to the attachment written by a Japanese Jaguwen expert who asks the Old Hanzi Expert Group to take full advantage of the achievement made by Jaguwen researchers.

Proposed Conclusion / Requested Action: the Old Hanzi Expert Group to discuss.

Input to Old Hanzi Expert Group

Japanese SC2 recently invited Dr. SUZUKI, Atsushi, one of the most active Jaguwen experts in Japan, for consultation on the IRG Old Hanzi projects. After briefing of the project status, Dr. SUZUKI raised two major concerns on the following points:

- ✓ Some of the distinction/unification principles described in IRG N1271 do not seem reflect the achievement of Jaguwen research already done in the past.
- ✓ Adding Shuowen radicals to Jaguwen would work for convenience's sake, while Shuowen radicals might be of little help in attempting to classify them.

Based on his observation, he asks the Old Hanzi Expert Group if it is useful for the group to take full advantage of the achievement that Jaguwen research field has done in the past twenty years. Especially he emphasizes the following points from experience in the field.

- ✓ 《甲骨文編》（中國社會科學院考古研究所編 科學出版社 1965 年）,attempting to classify Jaguwen according to Shuowen radicals, resulted in inconsistent unification and the Shuowen radicals for more than a half of all Jaguwen were left unresolved. The Old Hanzi Expert Group might want to avoid repeating this.
- ✓ 《殷墟甲骨刻辭類纂》(姚孝遂編 中華書局 1989年) should contain the most recent and successful taxonomy of Jaguwen by the research field, which covers most from 11 sources that the Old Hanzi Expert Group is referring to. The group could reduce the time and efforts in developing Jaguwen repertoire by leveraging the work already established by Jaguwen experts themselves.

The attached are (A) the abstract of the original document (Chinese) and (B) the original contribution (Japanese).

Attachment-A : Abstract of "Input to Old Hanzi Expert Group" in Chinese.

Attachment-B : The original document "Input to Old Hanzi Expert Group" in Japanese.

----end----

关于Old Hanzi 审议情况的意见

2007 年9 月4 日

日本・茨城大学 铃木敦

摘要[1]

现在，根据Old Hanzi（以及WG2）系统而对甲骨文字进行的符号化作业（从零开始）正在进行之中。但在甲骨学界，类似的工作过去已经实施过2种方法。

其一是《甲骨文编》（中国社会科学院考古研究所编 科学出版社 1965 年）之编纂。它过分依赖《説文》而加以整理，很明显其包摄范围設定自身即不合理，而且可能出现LABELING之处达到全文的 2 / 5 左右，参照现在的水准而言以失败告终也是不得以的。

另一为《殷墟卜辞综类》（岛邦男编 大安 1967 年）以及在此基础上发展的《殷墟甲骨刻辞類纂》（姚孝遂编 中华书局 1989年）的编纂。这种方法以公开发表的所有甲骨文字为对象，避开《説文》的束缚，以甲骨文字自身部首为基础进行了全面彻底的整理，虽然一些细节部分还有问题，但现已成为甲骨学界的规范标准。

现行根据Old Hanzi（以及WG2）所进行的工作，没有沿袭以上的成果，基本采用与曾以失败而告终的《甲骨文编》编纂方法同样的方法，其结局很明朗化。认识甲骨学界的经验和成果，活用《殷墟甲骨刻辞類纂》的成果，尽早朝这个方向转换，合理进行“符号化作业”，我认为可以节省庞大的劳动力。

注 1．作为本论文的依据，笔者的诸篇论文全部刊载于以下网页：
http://www.itscj.ipsj.or.jp/domestic/sc02/Old_Hanzi/lmd_docs/
另外其中大部分内容亦转载于宋镇豪・段志洪主编的《甲骨文献集成》四川大学出版社 2001 年第 14・19・38・40 卷，同电子版见 http://www.cn-oracle.com/book.php。论文中所使用的“文字域”概念，相当于“包摄范围”。

關於Old Hanzi 審議情況的意見

2007 年9 月4 日

日本・茨城大学 铃木敦

摘要[1]

現在，根據Old Hanzi（以及WG2）系統而對甲骨文字進行的符號化作業＜從零開始＞正在進行之中。但在甲骨學界，類似的工作過去已經實施過2種方法。

其一是《甲骨文編》（中國社會科學院考古研究所編 科學出版社 1965 年）之編纂。它过分依赖《説文》而加以整理，很明显其包摄范围設定自身即不合理，而且可能出现LABELING之处达到全文的 2 / 5 左右，参照现在的水准而言以失败告终也是不得以的。

另一爲《殷墟卜辭綜類》（島邦男編 大安 1967 年）以及在此基礎上發展的《殷墟甲骨刻辭類纂》（姚孝遂編 中華書局 1989年）的編纂。這種方法以公開發表的所有甲骨文字爲対象，避開《説文》的束縛，以甲骨文字自身部首爲基礎進行了全面徹底的整理，雖然一些細節部分還有問題，但現已成爲甲骨學界的規範標準。

現行根據Old Hanzi（以及WG2）所進行的工作，沒有沿襲以上的成果，基本採用與曾以失敗而告終的《甲骨文編》編纂方法同樣的方法，其結局很明朗化。認識甲骨學界的經驗和成果，活用《殷墟甲骨刻辭類纂》的成果，儘早朝這個方向轉換，合理進行“符號化作業”，我認爲可以節省龐大的勞動力。

注 1・作爲本論文的依據，筆者的諸篇論文全部刊載於以下網頁：
http://www.itscj.ipsj.or.jp/domestic/sc02/Old_Hanzi/lmd_docs/
另外其中大部分内容亦转载于宋镇豪・段志洪主编的《甲骨文献集成》四川大学出版社 2001 年第 14・19・38・40 卷，同电子版见 http://www.cn-oracle.com/book.php。論文中所使用的“文字域”概念，相當於“包攝範圍”。

中文要旨（GB & BIG5）

关于Old Hanzi 审议情况的意见

2007 年9 月4 日

日本・茨城大学 铃木敦

摘要[1]

现在，根据Old Hanzi（以及WG2）系统而对甲骨文字进行的符号化作业（从零开始）正在进行之中。但在甲骨学界，类似的工作过去已经实施过2种方法。

其一是《甲骨文编》（中国社会科学院考古研究所编 科学出版社 1965 年）之编纂。它过分依赖《説文》而加以整理，很明显其包摂范围設定自身即不合理，而且可能出现LABELING之处达到全文的2 / 5 左右，参照现在的水准而言以失敗告终也是不得以的。

另一为《殷墟卜辞综类》（岛邦男编 大安 1967 年）以及在此基础上发展的《殷墟甲骨刻辞類纂》（姚孝遂编 中华书局 1989年）的编纂。这种方法以公开发表的所有甲骨文字为对象，避开《説文》的束缚，以甲骨文字自身部首为基础进行了全面彻底的整理，虽然一些细节部分还有问题，但现已成为甲骨学界的规范标准。

现行根据Old Hanzi（以及WG2）所进行的工作，没有沿袭以上的成果，基本采用与曾以失敗而告终的《甲骨文编》编纂方法同样的方法，其结局很明朗化。认识甲骨学界的经验和成果，活用《殷墟甲骨刻辞類纂》的成果，尽早朝这个方向转换，合理进行“符号化作业”，我认为可以节省庞大的劳动力。

注 1. 作为本论文的依据，笔者的诸篇论文全部刊载于以下网页：
http://www.itscj.ipsj.or.jp/domestic/sc02/Old Hanzi/lmd docs/
另外其中大部分内容亦转载于宋镇豪 段志洪主编的《甲骨文献集成》四川大学出版社 2001 年第 14・ 19・ 38・ 40 卷，同电子版见 http://www.cn-oracle.com/book.php。论文中所使用的“文字域”概念，相当于“包摂范围”。

關於Old Hanzi 審議情況的意見

2007 年9 月4 日

日本・茨城大学 铃木敦

摘要[1]

現在，根據Old Hanzi（以及WG2）系統而對甲骨文字進行的符號化作業<從零開始>正在進行之中，但在甲骨學界，類似的工作過去已經實施過2種方法。

其一是《甲骨文編》（中國社會科學院考古研究所編 科學出版社 1965 年）之編纂，它过分依赖《說文》而加以整理，很明显其包摄范围設定自身即不合理，而且可能出现LABELING之处达到全文的2 / 5 左右，參照現在的水准而言以失敗告終也是不得以的。

另一爲《殷墟卜辭綜類》（島邦男編 大安 1967 年）以及在此基礎上發展的《殷墟甲骨刻辭類纂》（姚孝遂編 中華書局 1989年）的編纂，這種方法以公開發表的所有甲骨文字爲対象，避開《說文》的束縛，以甲骨文字自身部首爲基礎進行了全面徹底的整理，雖然一些細節部分還有問題，但現已成爲甲骨學界的規範標準。

現行根據Old Hanzi（以及WG2）所進行的工作，沒有沿襲以上的成果，基本採用與曾以失敗而告終的《甲骨文編》編纂方法同樣的方法，其結局很明朗化，認識甲骨學界的經驗和成果，活用《殷墟甲骨刻辭類纂》的成果，儘早朝這個方向轉換，合理進行“符號化作業”，我認爲可以節省龐大的勞動力。

注1，作爲本論文的依據，筆者的諸篇論文全部刊載於以下網頁：
http://www.itscj.ipsj.or.jp/domestic/sc02/Old_Hanzi/lmd_docs/
另外其中大部分内容亦转载于宋镇豪 段志洪主编的《甲骨文献集成》四川大学出版社 2001 年第 14、19、38、40 卷，同电子版见 http://www.cn-oracle.com/book.php。論文中所使用的“文字域”概念，相當於“包攝範圍”。

Old Hanzi　審議状況への意見

2007 年 9 月 4 日
日本・茨城大学　鈴木敦

要旨［注１］

現在、Old Hanzi（ならびに WG2）による甲骨文字の符号化作業が＜ゼロから＞進行中である。しかし、甲骨学界においては、既に過去２通りの方法で同様の作業が行われている。

一つは『甲骨文編』(中国社会科学院考古研究所編　科学出版社 1965 年)の編纂である。『説文』に引きずられた整理のため、包摂範囲の設定自体に不合理が目立ち、かつラベリング可能なものが全体の２／５程度に留まる等、現在の水準に照らせば失敗に終わったと言わざるを得ない。

もう一つは『殷墟卜辞綜類』（島邦男編　大安 1967 年）並びにこれを発展させた『殷墟甲骨刻辞類纂』（姚孝遂編　中華書局　1989 年）の編纂である。既刊の甲骨文字全体を対象に、『説文』の呪縛を逃れて甲骨文字自体の部首立てに基づく徹底的な整理が行われており、細部に問題を残しながらも、現在では甲骨学界のスタンダードとして機能している。

現行の Old Hanzi（ならびに WG2）による作業は、以上の成果を踏まえることなく、かつ失敗に終わった『甲骨文編』の編纂方法と基本的に同一の方法で進められている。その帰結は明らかである。**甲骨学界の経験と成果を認識し、『殷墟甲骨刻辞類纂』の成果を活用する方向へと早急に転換することが、符号化作業を合理的に行い、かつ膨大な労力の節減になると考える。**

Ⅰ：現状認識

A：Old Hanzi 並びに WG2 における甲骨文字の符号化作業は、以下の手順で進行中である。

１：既刊の甲骨文字資料全体を視野に入れ、一定の基準に基づき甲骨文字のリスト化・DB 化を行う。[担当：Old Hanzi]

→個々の字形の差異をどこまで「有意の差異」と見なすかが問題になるが、この段階では、かなり細かなレベルまで「有意」として区別している**［注２］**。

２：１の成果を、一定の基準に基づいて検討し、同一字として統合できるものは統合して、包摂範囲を確定し、レパートリーを作成する。[担当：Old Hanzi]

→結果的に、上記１ほど細かな区別は行わない。

３：２で定めた個々の文字の包摂範囲に対してコードを振る。[担当：WG2]

→コードの順番はどんなものでもよいが、一つの包摂範囲に対して一つのコード番号が一意に対応しなければならない。換言すれば、「１：多」「多：１」「多：多」の対応は、いずれも不可である。

４：２で定めた個々の文字の包摂範囲に対して、ラベリングを行う。[担当：WG2]

→現状では、『説文』部首によるラベリングが企図されている。

B：審議には古文字の研究者も関わってきたが、古文字の中でも特に甲骨文字を専門とする人間は関わってこなかった。その結果、甲骨文字の実態にそぐわない部分も見受けら

れる。

II：斯界の経験の有効活用を

甲骨文字研究の世界では、既に上記の Old Hanzi の手順と基本的に同一の方法で甲骨文字の包摂範囲整理が行われている（中国社会科学院考古研究所『甲骨文編』科学出版社 1965 年）**【注３】**。結果は以下に記す通りであり、結論としてこの方法は失敗に終わったと言わざるを得ない。

ａ：甲骨文字の字形の偏差は、篆書以降の文字に比べて甚だしく大きい。『甲骨文編』編者は、上記１段階相当の作業では、Old Hanzi N1271-4 同様、字形弁別について一定の基準を立てようとしたと想像されるが、実際の編纂作業に当たって全ての字形に統一的に当てはめることは不可能であったと見えて、結果は甚だ混乱したものとなっている。

ｂ：『甲骨文編』編者は、上記２段階相当の作業では、「字義の同一性」に基づいて「字形の（小さな）差異」を無視して統合する、という作業を行ったと想像されるが、甲骨文字の字義は「現代の文字との字形的な相似」では知り得ない。「ある甲骨文字が、甲骨文の中でどのように使われているか」を、膨大な用例から帰納する以外に方法がない**【注４】**。『甲骨文編』は、そのような検討を行うことなく統合作業を行った結果、種々問題のある包摂範囲を設定することになった。

ｃ：『甲骨文編』編者は、上記３・４段階相当の作業では、各包摂範囲を「『説文』部首に基づいてラベリングし、『説文』部首順に配列し、冒頭から４桁の通し番号を振る」方法を採った。その結果、包摂範囲と番号との関係が「１：多」（例えば「翌」に対して「0004（吏）」「0388（史）」「0389（事）」「1003（使）」が対応）となったり「多：１」（例えば「処」と「翰」が共に「1704（巳）」に対応）となったりという事態が多発した**【注５】**。

ｄ：甲骨文字の中には、『説文』編纂以前に消滅したものが多数あるため、上記４段階で、『甲骨文編』編者が立てた包摂範囲約 5,000 字の内、ラベリングができたものは約 2,000 字に留まり、残りの約 3,000 字はラベリング不能・検索不能に陥った。

「全ての甲骨文字を対象として、統一的かつ論理的な基準に基づいて包摂範囲を定め、一意に対応する番号を割り当て、ラベリングを行う」という作業は、甲骨学界においては戦前に始まり、『甲骨文編』の失敗を経て、開始から半世紀以上たった 80 年代末に漸く『殷墟甲骨刻辞類纂』を生むに至った。その背後には多数の甲骨学者の膨大な努力がある**【注６】**。**これと同様の作業を「専門家以外の人間集団で・ゼロから・ほんの数年で」実行し、かつ『殷墟甲骨刻辞類纂』を凌駕する合理性と使い勝手の良さを備えた成果を纏めようとすることは無謀と言わざるを得ない。**

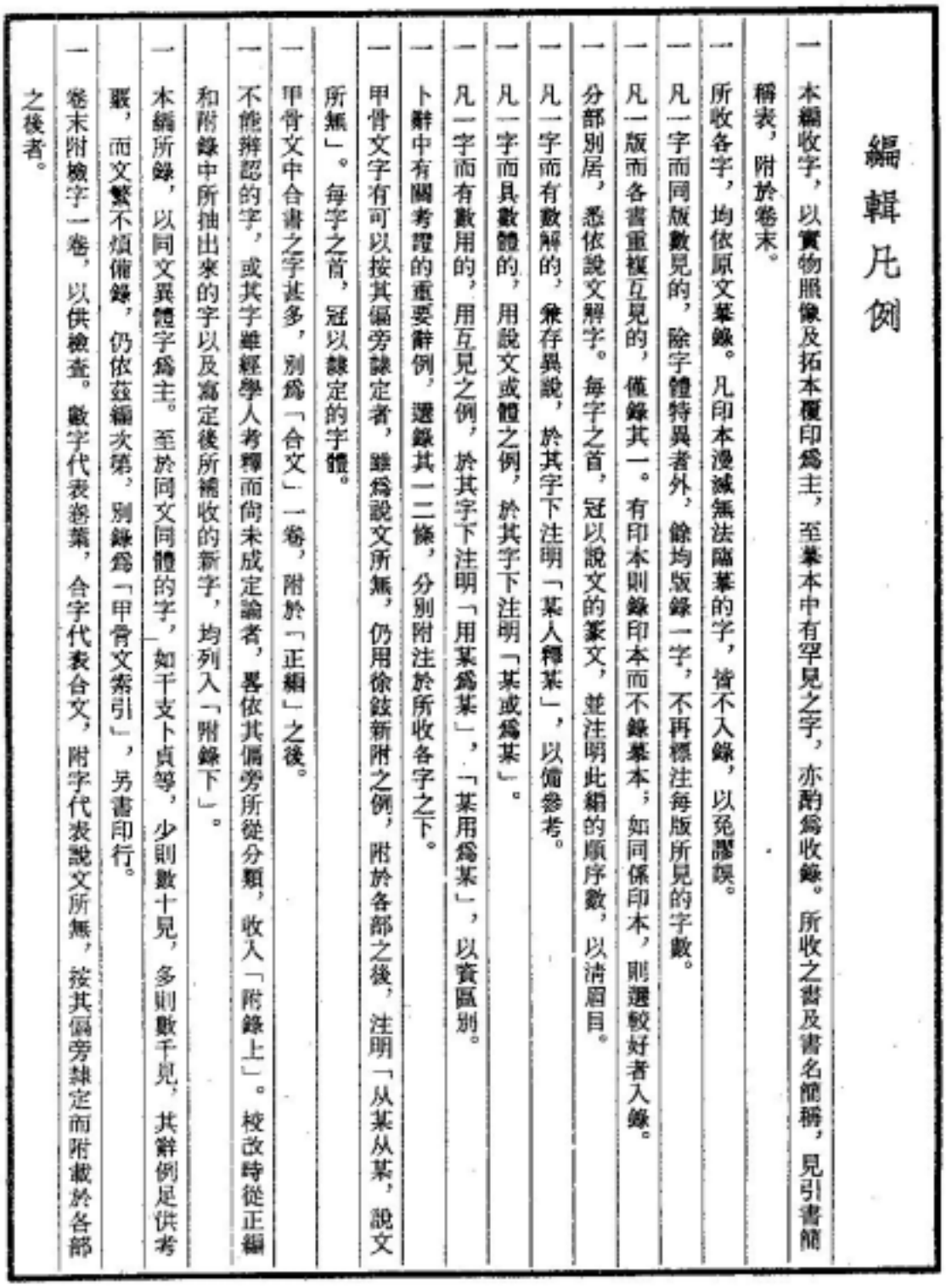
斯界には、既に上記の作業結果としての『殷墟甲骨刻辞類纂』が存在する。そのベースとなったのは、同書の編纂時＝1980 年代までに公刊されていた全ての甲骨文字資料であり、Old Hanzi の議論が底本とする 11 種の甲骨資料の殆どをカバーしている。つまり、少なくとも上記１・２段階の作業は、既に専門家の手で基本的に完了しているのである。上記１・２段階の作業で、現段階で行うべきは、以下の２点と考える。

ａ：『殷墟甲骨刻辞類纂』に漏れている資料について精査し、既存の包摂範囲に含められない文字があれば（＝少数と予想される）包摂範囲を追加新設して収める **[注７]**。
ｂ：『殷墟甲骨刻辞類纂』の包摂範囲の内容を再検討し、もし不合理があれば然るべく修正する **[注８]**。

注

１：本件の論拠となる筆者の諸論文は、
http://www.itscj.ipsj.or.jp/domestic/sc02/Old_Hanzi/lmd_docs/
に、全て掲載している。また、その大部分は
宋鎮豪・段志洪主編『甲骨文献集成』四川大学出版社　2001 年　第 14・19・38･40 巻
同・電子版　http://www.cn-oracle.com/book.php
にも転載されている。なお論文中で用いた「文字域」という概念は、「包摂範囲」に相当する。
２：IRGN1271 の 3 および 4 に基づく認識である。「点の数」等への拘りもさることながら、甲骨文字の世界ではごく普通に存在するミラーリバースまでも「有意の差異」の如く扱っていることには（仮に後々統合されるにしても）違和感を覚える。
３：同書冒頭の「編集凡例」を図１に示す。現行の Old Hanzi 並びに WG2 の整理方法と比較されたい。
４：この方法の創始者・島邦男による『殷墟卜辞綜類』あとがきを図３に示す。包摂範囲の設定法、字義の確定法等、図１に示した『甲骨文編』のそれと対比して戴きたい。
５：該当部分を図２に示す。この問題は IRGN1271 の 5-4 でも取り上げられ、「一つの部首に収められる」との方針が示されている。『説文』の部首立てに拠る以上、妥当な方針であるが、特定の例については採りうる方針が甲骨文字全てに対して・遺漏無く適応できるとは限らない。図２に示すように『甲骨文編』も同様の方針を立てているが、実際の整理においてはこれを貫徹できず、様々な混乱を生じている。
６：『殷墟甲骨刻辭類纂』の部首立てを図４に、包摂範囲を列挙した字形総表（の一部）を図５に、その背景にある用例研究（の一部）を図６に示す。
７：『殷墟甲骨刻辭類纂』は、IRGN1271 の 3 に示された 11 編の内、(A)（F）(G)（H）を対象としている。資料数は約５万片に上り、(B)（C）(D)（E）(I)（J）(K）の合計１万数千片を大きく凌駕している。加えて、後者の大部分（＝約 14,000 片）を占める（B）は、書名が示す通り（A）の編纂に漏れた資料の補遺という性格である。符号化すべき甲骨文字の字形で、前者に含まれないものは少数に留まると判断される。
８：作業の膨大さに鑑み、某かの変更が生じる可能性は否定できないが、包摂範囲の体系を揺るがすような大変更はないと予想している。

図１：『甲骨文編』編輯凡例



編輯凡例

一 本編收字，以實物照像及拓本覆印爲主，至摹本中有罕見之字，亦酌爲收錄。所收之書及書名簡稱，見引書簡稱表，附於卷末。

一 所收各字，均依原文摹錄。凡印本漫滅無法臨摹的字，皆不入錄，以免謬誤。

一 凡一字而同版數見的，除字體特異者外，餘均版錄一字，不再標注每版所見的字數。

一 凡一版而各書重複互見的，僅錄其一。有印本則錄印本而不錄摹本；如同係印本，則選較好者入錄。

一 分部別居，悉依說文解字。每字之首，冠以說文的篆文，並注明此編的順序數，以清眉目。

一 凡一字而有數解的，兼存異說，於其字下注明「某人釋某」，以備參考。

一 凡一字而具數體的，用說文或體之例，於其字下注明「某或爲某」。

一 凡一字而有數用的，用互見之例，於其字下注明「用某爲某」，「某用爲某」，以資區別。

一 卜辭中有關考證的重要辭例，選錄其一二條，分別附注於所收各字之下。

一 甲骨文字有可以按其偏旁隸定者，雖爲說文所無，仍用徐鉉新附之例，附於各部之後，注明「从某从某，說文所無」。每字之首，冠以隸定的字體。

一 甲骨文中合書之字甚多，別爲「合文」一卷，附於「正編」之後。

一 不能辨認的字，或其字雖經學人考釋而尚未成定論者，畧依其偏旁所從分類，收入「附錄上」。校改時從正編和附錄中所抽出來的字以及嘉定後所補收的新字，均列入「附錄下」。

一 本編所錄，以同文異體字爲主。至於同文同體的字，如干支卜貞等，少則數十見，多則數千見，其辭例足供考覈，而文繁不煩備錄，仍依茲編次第，別錄爲「甲骨文索引」，另書印行。

一 卷末附檢字一卷，以供檢查。數字代表卷葉，合字代表合文，附字代表說文所無，按其偏旁隸定而附載於各部之後者。

図２：『甲骨文編』における「史」「吏」「使」「事」の整理

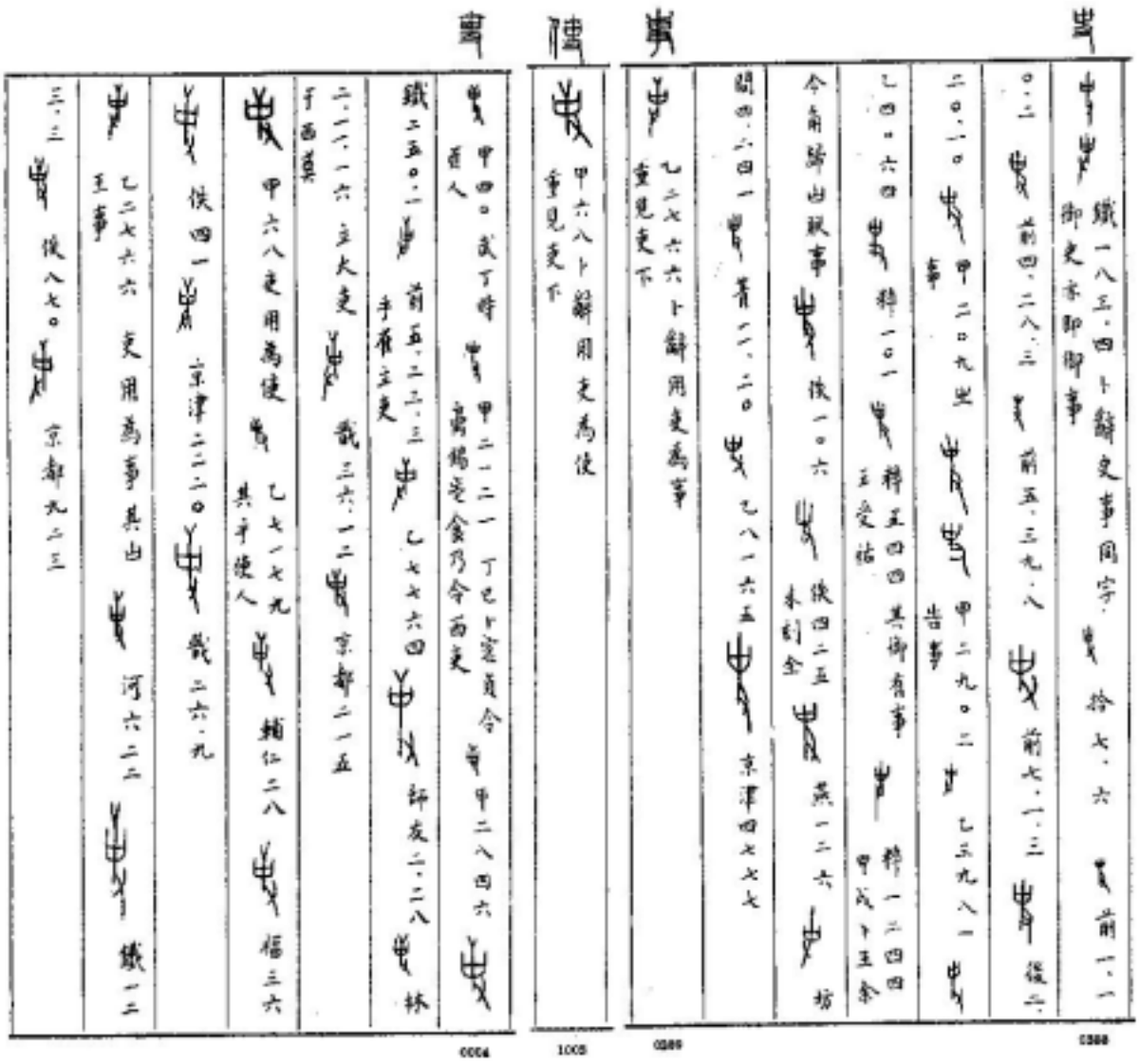


図３：『殷墟卜辞綜類』あとがき

あとがき

島邦男

今から三十數年前、私がまだ學生の頃、宇野哲人先生が講義の餘談に、金文から甲骨文を話されて、甲骨文字は僞作であると申された。私はひどく興味をそそられ、いつかはこれを研究してやろうと考えた。これから興味は次第に文字の方面に傾き、乏しい財布をはたいては文字學の書物や、甲骨文字の資料を購入したが、唐蘭氏の古文字學導論を讀んで驚いたり、孫海波氏の甲骨文編に望洋の嘆を發する未熟者で、甲骨學に入る戸口は解らず、ただ積んで置くだけであった。その後偶〻「上帝」の神格を調べて、朱子から溯って詩・書に及んだがどうも解らず、甲骨文字の世界ではどうなっているだろうかと考えた。そこで勇敢にも、上帝に關する甲骨文を全部調べて解讀しようと決心し、文求堂で何かと教えを受けて、これに取りかかったのが、私が卜辭研究に入った動機である。

入って見て驚いたことは、調べれば調べるほど、甲骨文字の解釋は管見と總說に滿ちていて、正に辭宵が象を撫ずる觀を呈していることであった。私は字形から文字を解釋するやり方に、失望すると同時にその限界を感じ、又貞讀からの解釋するやり方にひどく不安を覺え、よりよい方法がないものかと日夜考えるようになった。ある時ふと漱石の小說の中に「理學は對象を固定することによって成立している。研究對象が不安定である限り、學問は成り立たない。」という意味のことがあったのを思い出し、卜辭の資料を確定することができないものかと考えた。然し甲骨文編に記されている資料が、果して日本で見られるかどうかをも知らない私にとっては、これは夢に過ぎなかった。上帝に關して平岡武夫教授の著書をみると、甲骨文字が利用されているので、ふとこの先生に會ったならば、日本に於ける甲骨文字の研究の狀態がわかりはしまいかと思いつき、混雜する汽車に乘って、のこ〳〵でかけて行ったのは終戰後間もない頃であった。

平岡教授との出會いは、私を甲骨學に深く結びつける契機となった。これから幾度か東北の片田舍から、京都の人文科學研究所に足を運び、時には幾日か滯在して甲骨文字の資料の吸收に努めた。その頃の或日、教授は索引のことを話され、そのついでに話はアメリカ人の日本語研究の方法に及び、彼等は日本語の語の結びつき方を、數冊の書物から精密に分析歸納する仕方をとっていると、なに氣なく話された。これを聞いていた私は「これだ、この方法を卜辭の解讀に攝り入れることだ」と突嗟に思いついた。これからは研究所で得た資料によって「祭祀の研究」をやる傍、乾乾と資料の分析と整理に努め、どうにかお粗末なものを作り上げることができた。

次いで京都に於ける日本中國學會の際、貝塚茂樹教授から、臺灣から贈られて來たばかりの小屯を見せてもらった。初めて見るこの膨大な資料を前にして、私は驚きと喜びと同時に、今までの整理の努力が水泡に歸したことを感じた。再び小屯を含めての資料の整理をやり直さねばならないからである。それから殆ど二年を費して、略〻その整理を終ったが、その頃から新しい資料が續々と刊行されたので、三度整理をやり直さねばならなくなった。

そこで私は新資料を含む全資料を徹底的に整理することを決意して、文部省に補助を申請した。それは今から八年前の昭和三十四年である。先づ資料の一片、一片の字數に從って複寫を作る計畫を立て、殆ど二年を費やし原資料の略〻十三倍の資料を作り上げた。次いで二千餘の本のケースを研究室に所狹しと並べて一字々々に分類したが、これだけで殆ど一年を費した。複寫機を操ることも、一片一片に切りとることも、その一片一片をその字の小箱を探して、これに投入することも、すべて自分一人でやる外はなく、この準備段階の作業は誠に容易ではなかった。

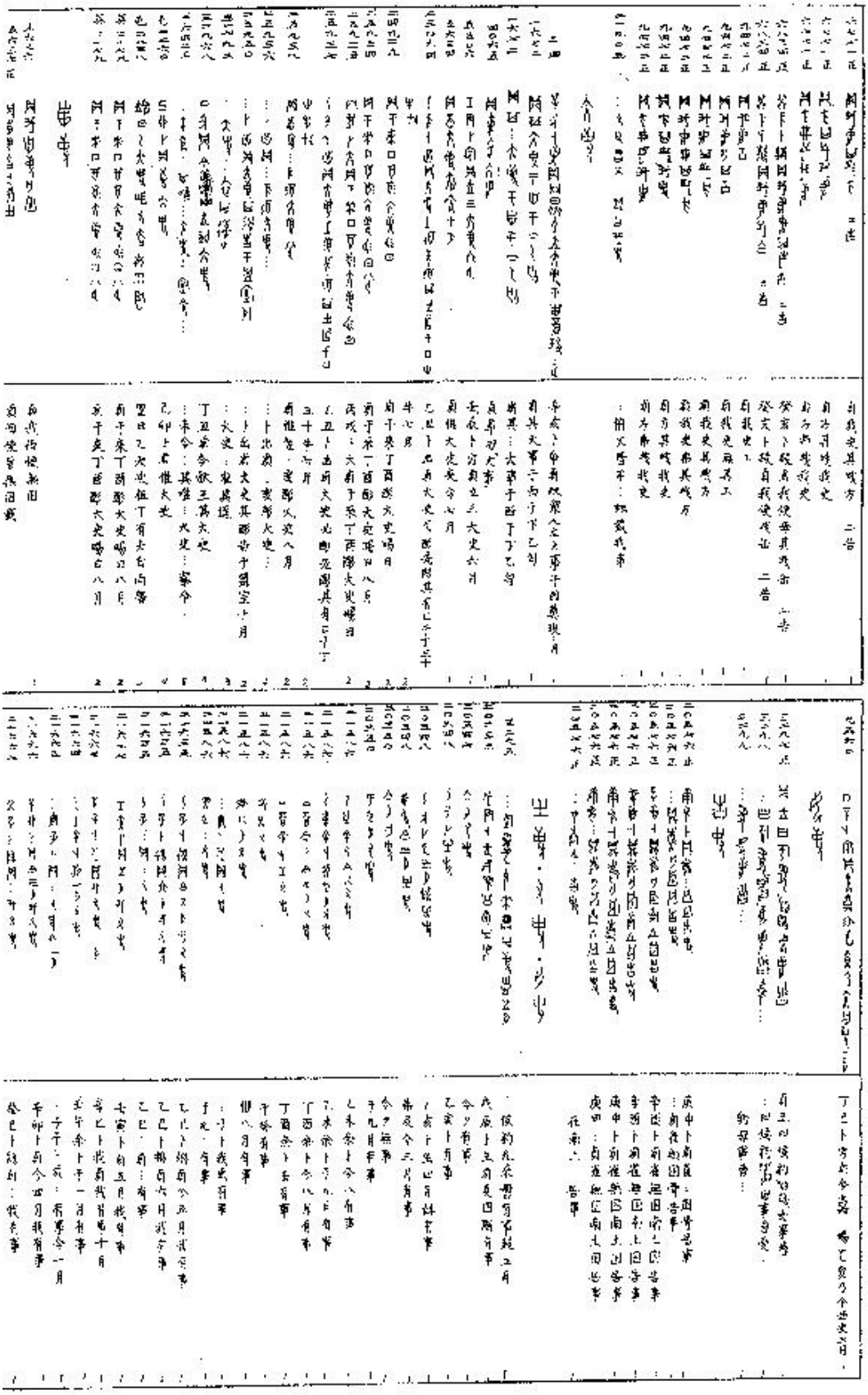
その頃赤塚忠教授が研究室に訪ねて來られたが、この態をみて、あきれた顏をされるだけで、一言の批評の言葉もなかったことを覺えている。準備だけで既に三年、一字一字を更に用例に從って分類し、時代順に整理して筆記してゆくとなると、この先幾年ででき上るのだろうかと、その頃はしきりに氣があせった。それからは初心忘るべからず、愚公が山を移したではないかと念じて、整理に沒頭する毎日を續けたが、淨書の段階が近づくにつれて、字釋をつけるか否かが、心を惱ますものとなってきた。字釋をつけるのは容易だが、どの說を採るかとなるとこれは大問題である。幸に一昨年季孝定氏が甲骨文字集釋を刊布したので、字釋はこれに讓ることに心を決めることができた。こうして準備段階が終ってから五年後の本年九月六日に、漸く本文を完成することができたのである。これが殷墟卜辭綜類である。顧みれば私の生涯の事業であった。

昭和四十二年九月二十四日　記

図４：『殷墟甲骨刻辭類纂』部首表

図５：『殷墟甲骨刻辞類纂』字形総表（部分・全体は 26 頁）

図６：『殷墟甲骨刻辭類纂』本文＝文例整理（部分。全体は 1,473 頁）



Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1460

Date: 2007-06-12

Title:	**Report of the Old Hanzi Expert Group**
Source:	**Old Hanzi Expert Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**6**
Medium:	**Electronic**

The Old Hanzi Expert Group recapped Old Hanzi Principles and References (Version 2) and discussed open issues in details at this meeting.

This report is organized as follows:

The Old Hanzi Expert Group held its meeting from 10 June to 14 June 2008 in Busan, the Republic of Korea. Mr Kobayashi, the SC2 Chair, reiterated the principles and procedures concerning the encoding of ideographs and made it clear that different glyphs of a representative character could not be encoded separately. After discussing IRG N1424 and reviewing the character tables submitted by China and TCA (Shuowen radical 181 to 280), the meeting finalized its future work plan. The following resolutions were made:

2008 年 6 月 10 日至 14 日古漢字專家組在韓國釜山召開會議。SC2 主席小林先生於會上詳細說明文字編碼的原則和程序，並明確指示同一個代表字的不同字形，不能分別編碼。本次會議並討論 IRG N1424 文件，審查中國代表團和 TCA 提交的字表(說文部首 181~280)，決定了日後之工作計畫。

經過小組的討論作出了如下決議：

The following resolutions were made:

- **Revisit of the Old Hanzi Principles and References (Version 2)**
- **Discussion of IRG N1424**
- **Summary of discussion at this meeting**
- **Revised work plan for the development of Old Hanzi inscriptions repertoire**

1. Revisit of the Old Hanzi Principles and References (Version 2)

1.1 Based on the conclusions made at this meeting, the principles would be revised (Version 3) (IRG N1471)

依據本次會議結論，將從新整理古漢字原則，為第 3 版 (IRG N1471)

1.2 By 15 July 2008, Selena will send IRG N1471 to all the experts, whose comments should be sent to Selena by 10 October 2008 for consolidation and subsequent submission to the IRG。

7 月 15 日前 selena 將寄出 IRG N1471 給所有的專家，10 月 10 日前將意見返回 selena，整理後提交給 IRG。

2. Discussion of IRG N1424

A better understanding was obtained after exchange of views following Suzuki Atsushi's presentation on IRG N1424. Dr. Suzuki Atsushi, when in Japan, did not have sufficient information on the work of the Old Hanzi Expert Group. He gained much

information at this meeting and a broad consensus was reached after thorough discussion. He will re-visit the matter and submit his views again at IRG#31.

Suzuki Atsushi 說明 IRG N1424 文件，會議中進行討論後，增進了大家的了解。**Suzuki 先生於日本看了之前的文件，因資訊不足，無法詳細了解，討論後取得了大致的共識。參加本次會議以後，Suzuki 先生得到了很多資訊，回日本以後將進行整理，並於 IRG #31 會上提出意見。**

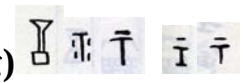
3. Summary of discussion at this meeting

3.1 Revise the principles of Old Hanzi selection 修訂選字原則

Characters with the same meaning and pronunciation but with different structure or component will be put into the same category. A representative glyph will then be chosen.

字形之分類概念，將音義相同，字形構字或部件不同的字歸為一類，並選取一個字作為代表字。

Example:

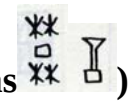
(莫)  ，**(示)** 

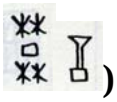
3.2 The principles of selecting a representative glyph:

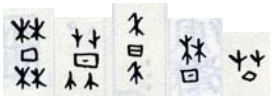
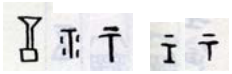
代表字的選取原則

Based on the origin shape of the Oracle Bone Inscriptions, the most representative glyph will be selected.

以甲骨文原形來做判斷，選取最有代表性之字形。

Example: use the first glyph as the representative (such as )

範例：以第一個字為代表字(如 )

(莫)  ，**(示)** 

3.3 Principle of radical classification

歸部原則

For a character that is in Shuowen, the same radical as in Shuowen will be used. For a character that is not in Shuowen, the most suitable component in it will be chosen as the radical, on the condition that the radical chosen is also in Shuowen.

若有對應說文的字，則依照說文歸部;如不包含說文部首之形，則依照該字形適當之偏旁附註部 2。

For example 1:

 說文《  部》有對應  字的重文『明』，則  歸  部。

For example 2:

 說文《邑部》有對應的『邦』字的重文  ，則 

歸邑部，並在附註中註明部首 2 田部。

3.4 To benefit from other experts in Oracle Bone Incriptions, outside experts will be invited to join the Old Hanzi Expert Group.

更加廣泛聽取全世界甲骨文專家的意見，日後將邀請有關專家參加 IRG 古漢字專家組的工作。

3.5 To launch an IRG Oracle Bone Inscriptions website. Consolidated information and findings will be released so that other experts may provide their comments.

建立 IRG 甲骨文網站，將整理的過程和結果於網站上公佈，聽取專家的意見。

4. Work plan

4.1 Progress(選字進度)

- **Range of selection: from Shuowen Radical 281-400. Radical 281-340 in IRG N1473(TCA) and Radical 341-400 in IRG N1472(China)**

  選字範圍，從說文部首 281-400，281-340 IRG N1473(TCA)，341-400 IRG N1472(China).

- **Submit the character table by 15 October 2008 to Selena.**

  10 月 15 日以前提交字形表

- **Selena will generate a consolidated table (IRG N1474) between 16 October and 31 October 2008.**

  10 月 16-31 進行整合字表 IRG N1474

- **Send out the consolidated table (IRG N1474 radical 281-400) by 22 October for discussion at IRG#31.**

  10 月 22 寄出整合後之字表 IRG N1474 (281-400)，並於 IRG#31 會議中進行討論。

4.2 Revision of Old Hanzi Principles and References (Version3)

Production of Old Hanzi Principles and References (Version3)

- **By 15 July 2008, Selena will send IRG N1471 to all the experts.**

  7 月 15 日前 selena 將寄出 IRG N1471 給所有的專家。

- **Comments sent to Selena by 10 October 2008 for consolidation and subsequent submission to the IRG.**

  10 月 10 日前將意見返回 selena，整理後提交給 IRG。

Appendix: A combined character table of Oracle Bone Inscriptions(181-280).

Members attended the meeting were Dai Hong, Li Guoying, Zhao Cheng, Zhou Xiaowen, Chen Shuangxin, Wei Lin Mei(Selena), Jung Bor Sheng, Chi Hsiu Sheng, Kim Jinhe, Tang Yingmin, Suzuki Atsushi, Kobayashi Tatsuo（小木木）.

End of document

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1522
Date: 2008-11-09

Source:	Japan
Meeting:	IRG#31, Kunming, China
Title:	Questions on the policy of old hanzi expert group works
Status:	Member's submission
Actions required:	FYI
Distribution:	IRG Old Hanzi Group
Medium:	Electronic
Pages:	7

1. Introduction

Study on JiaGuWen for standardization has been under the way for nearly 3 years. Recently, Japan National Body (JNB) has been attending to this standardization process, and has several concerns on the working process. In this document, Japan describes questions derives from those concerns and requests members for the clarification.

2. Questions on Old hanzi group works

2.1. Supposed user requirement on this repertoire

Users of JiaGuWen (甲骨文) characters may vary from Historians, Linguists, Calligraphic experts to the others. Depend on their needs for displaying and searching characters, encoding policy may differ. What kind of user community is supposed on this work?

Example: Attached paper distinguish the differences of JiaGuWen character for ages (from first age to fifth age 第一期～第五期) and their variants such as official versus non-official (王朝卜辭／非王朝卜辭). Furthermore, some researchers may also need to distinguish other variants. On the other hand, such precise classifications may not be useful for calligraphers.

2.2. Unification Policy

In the current work, unification policy for each character is not clearly stated nor written down as a document. The basic unification rules, common to all characters, are stated in IRG N1460 or IRGN 1271, however, individual issues agreed by discussions are not clearly stated or recorded. Many of these individual issues are not yet known among the

JiaGuWen experts or researchers in the world. So it is requested to establish and form the policy that this group works based on. Does the group have any plan on this request?

2.3. Use of method for the shape variants

For the ideographic characters it is useful to use compatibility characters or IVS to treat the shape variants. Otherwise it causes complexity on searching or comparing text if codes are assigned to all glyphs without references to variant relations. What is the encoding policy of this group works? Is it expected to use such methodology?

2.4. Font Design Policy

Japan also wonders about how to produce font for publishing future standards.

Basically, there are two choices on font design policy. One is to faithfully copy the representative oracle bone inscription chosen from various sources. Another is to design our own representative glyph, which is an average shape of various sources.

FYI: Currently, for example, 『殷墟甲骨刻辭類纂』 adopts the `faithful copy' design policy, while 『殷墟卜辭綜類』 adopts the `average design from various sources' design policy.

In any case, old hanzi group will be requested to provide font for publishing standard when IRG submits to WG2. Who is expected to provide font and how are the glyphs designed?

2.5. Ordering and Attributes

Currently, standardization on JiaGuWen characters are based on ShuoWen JieZi (『説文解字』). This means that all JiaGuWen characters can be mapped to equivalent modern Hanzi Character appeared in ShuoWen JieZi, however, a majority of JiaGuWen characters may be defunct. (One theory states that among 6,000 JiaGuWen characters, only 900 are survived ShuoWen). When we are to arrange oracle bone characters based on ShuoWen, we must determine how to classify those already defunct JiaGuWen characters.

Also, several mappings from JiaGuWen character to modern character are not stable. For the JiaGuWen characters which have multiple candidates of corresponding modern character, we must determine which modern character should be adopted, and the other modern character candidates should be attached as informational attributes.

Is there any solutions to treat such unmapped JiaGuWen characters?

2.6. Procedural Recording Policy

Japan is also afraid if the discussions and decisions made by the group are not sufficiently recorded.

Studies on JiaGuWen are still immature, and new findings or theories are reported every year. In the future, we may need to modify, or add new characters to, once established standard. At that time, referring what current Old Hanzi Group had discussed or decided would certainly be needed. So, it is expected to establish policy recording and sharing decisions clearly.

2.7. Document on the principles

Old Hanzi group was resolved to produce its principle and references document revision (IRG M30.8), however it is not registered in IRG web site. Japan will be happy if answers to the questions above will be resolved in the document. Furthermore, it will be useful for users that the document is united with IRG's PnP document.

3. Conclusion.

The above questions and requests are all related to the standardization 'policies'. Japan would like to get suggestions to the question above as an IRG document for sharing idea of the work among our experts.

Japan feels that such document will also be helpful not only for JiaGuWen, but also for future JinWen(金文) and ZhangGuoWenZi (戦国文字) standardizations, too.

(Attachment)

甲骨文中黄与寅的关系及其相关问题

秦晓华

(广东 广州 510275)

摘要：有学者认为甲骨文中的“矢”、“寅”、“黄”三字同源，皆源于“矢”。笔者对《甲骨文合集》中的“矢”、“寅”、“黄”三字作了穷尽式的调查，从分期的角度辨析了三字的关系。最后，得出结论：“黄”字从“大”而非从“矢”，与其它二者不存在同源的关系。另外，有学者所言的“寅尹”应该都是“黄尹”的误释。

关键词：黄，寅，矢，甲骨文，黄尹，寅尹

中图分类号：K871.3　　文献识别码：A　　文章编号：1001-0327(2008)01-0117-04

一、“矢”、“寅”、“黄”辨

“矢”、“寅”、“黄”三字的甲骨文形体相近，释者每易混淆。姚孝遂先生以为是一字所分化：“契文‘矢’、‘寅’、‘黄’本同源，以用各有当，渐致分化。”[1]而王国维先生在《戬寿堂殷墟文字》中却指出“黄”、“寅”二字实有从“矢”从“大”之别：“郭罗参事释寅父，然卜辞寅字皆从矢，而人名之  尹皆从大，疑非寅字也。”[2]王氏所谓“尹”之“”实即“黄”字。孰是孰非，学界仍未有统一看法。

为了对这三个字的甲骨文形体及其相互关系有一个清楚的认识，笔者不揣浅陋，拟从分期的角度，再参以金文，对“矢”、“寅”、“黄”三字的形体关系作一历史考查。在考查中，笔者对《甲骨文合集》作了穷尽式的调查[3]，现将调查结果表示如下(表一)(各个形体右下脚数字为其在本期出现的总次数，形体残缺者和漫漶不清者暂不计其内)：

根据上表，我们先纵向地分析一下每个字的演变情况：

甲骨文中“矢”字的形体确如罗振玉所言，

表一：

字 / 分期	矢	寅	黄
一期	3	 656 4	 94 27 5 2 1
一期(附)	——	153 30 4	3 1 1
二期	2	269	4 1 1 1
三期	1	68 1	4 1
四期	1	132 2 1	2
五期	1	104 48 21 10 10 9 3 2 1 1 1 1 1	8 5
西周金文			

注：笔者倾向于将《甲骨文合集》第七册，即“一期附”看作是武丁后期的卜辞，因此笔者在统计时将其放在一期之后。

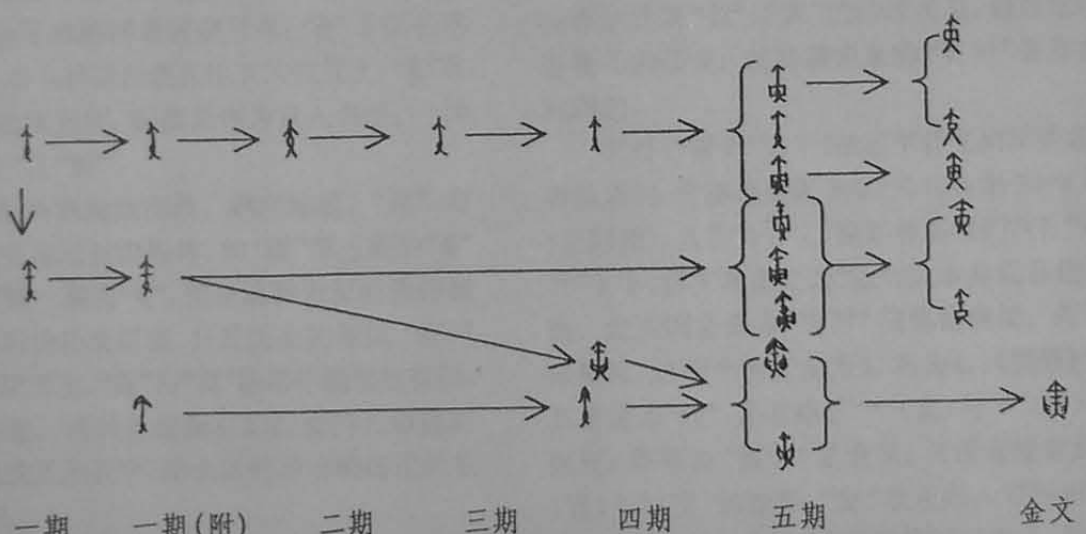
收稿日期：2004-06-07

作者简介：秦晓华(1978—)，男，河南安阳人，中山大学中文系 2005 级古文字博士研究生。

"象镝干括之形"[4]。整个殷商时期，"矢"的形体都比较稳定，第一期和第五期相比，没有显著的变化，而且与后代金文的继承关系也显而易见：甲骨文"矢"的箭尾交叉状，变为金文箭干上一点，这一点继而变为一横。

"寅"字的形体，郭沫若谓："字于甲骨文作若，均象矢若弓矢形。"[5]对此种说法，姚孝遂在《甲骨文字诂林》"寅"字条下按语中指出："其作，不得谓为从弓，乃区别之文。"姚说是也。至于为何借"矢"为"寅"，郭氏之"与引射同意"[6]可备一说。

相对而言，"寅"字由于使用频繁，其形体在殷商时期变化较为剧烈，从早期的与"矢"无别演变为晚期的与"矢"迥异，其具体演化情况见下图所示：



从上图我们可以看出，和是"寅"早期的形体，则更为象形，但并不常见，从第一期至第四期一直占统治地位还是，它是当时人们通行的写法，导源出了后期各式各样的形体。到了第五期，"寅"字的写法大变，虽然还有这样传统的形体，但已是罕见，常见的却是于其之上增加饰笔而成的各式各样的变体。另外，值得我们注意的是，在这些变体中有一个细微的变化，那就是"寅"字所从之""(矢)出于简便的原因，箭尾交叉之形消失，变为""，与""(大)颇为近似，这样就使得"寅"字的某些形体和"黄"字相近。到了西周金文，"寅"字变化多端的形体又趋于统一，常写作。

关于"黄"字的甲骨文形体，郭沫若谓其像佩玉之形[7]，唐兰认为是系市的带子[8]。不管是佩玉抑或佩带，此形体与人有关是勿庸置疑的。从"黄"字的早期甲骨文形体也可以看出，"黄"字确实从正面人之形"大"，而非从"矢"，王国维的认识是正确的。整个殷商时期，"黄"的字形变化比较细微，其变化主要表现在所从之""(大，象人正立之形)的两腿由弯变直，作""，这可能是书写快捷所致。这种形体早在第一期就出现了，但只是偶尔书之，并不普遍，通行的写法还是两腿弯曲的形体。直到第三期，""这样两腿作直撇的形体才通行起来，弯曲者已不见于甲骨文。

另外，王襄谈到："疑是黄之省文。"[9]在甲骨文当中，我们发现了""这样的形体，与西周金文的""(黄)无别，应为"黄"在甲骨文中的标准写法，但先人出于刻写便利的原因，往往简而书之作""，所以""这一写法在甲骨文中用的较少，总共仅发现两例(另一例出现在屯南2182之上)。而金文多用于庄重场合，所以取其标准写法。其情形与"贞"既作""(用鼎为贞)又作""(贞的简体)相类。

弄清楚每个字的演变脉络之后，再来分析它们之间到底是否同源就比较容易了。前人之所以认为"矢"、"寅"、"黄"三字同源，就在于没有弄清楚"黄"和"寅"的关系，我们只要把两

者的关系理顺了，这个问题就迎刃而解了。

我们先横向地比较。从上表中可以看出，第一期到第四期，“寅”和“黄”形体区别明显，决不相混。但是到了第五期，由于“寅”字所从“矢”的箭尾交叉状消失，再加上又增加了饰笔，这样就与同时代的“黄”字出现了相近似的形体。两者虽然很容易混淆，但是只要我们仔细观察，还是会发现此二者有细微的差别，这种差别主要表现在字中间的“日”字形的部件：“  ”字所从作方框状，而“  ”字所从为椭圆状。虽然文字形体相近，但是我们知道它们的形构是不一样的：“”（寅）字所从之“”是“”（矢）的简化，中间的“日”字形部件是区别符号；而“”（黄）字所从之“”是“”（大）字的简化，“日”字形部件为人所佩戴之物。另外，从文例上我们也不难将两者区别开来：“”于第五期仅两见，并且都是出现在地支的位置上；“”在第五期出现五次，四次是作为贞人名字，一次是“黄牛”之“黄”。

我们再纵向地比较。纵向地看，“黄”与“寅”也有相近似的形体，如“寅”第五期的“”与“黄”第一期的“”，但是这种近似的形体却处在不同的历史层面，只是历史的相似。在同一历史层面上，“黄”与“寅”的形体则迥然有别，不易相混。若只是根据它们历史的相似就将两字说成是形近字，那么这样得出的结论就是草率的。

这样，由于“黄”和“寅”纵向和横向上的错综关系，就形成了“矢”、“寅”、“黄”三字“同源”的误解。

总之，通过以上的分析，我们可以得出这样的结论：“黄”字从“大”而非从“矢”，“黄”与“矢”、“寅”不同源。

二、“黄尹”与“寅尹”

正因为“黄”与“寅”具有这样错综复杂的关系，而导致一些学者把“黄尹”改释为“寅尹”。

最早改释“黄尹”为“寅尹”的是杨树达先生。上文提到，王国维怀疑罗振玉所释“尹”之“”不是“寅”字，而杨氏在《卜辞琐记》中谈到：“《殷虚书契前编》三卷七之四‘戊寅’寅字作，《龟甲兽骨文字》一卷十六之二‘△寅己卯’寅字作，并与此字同，罗释寅，是。王疑非寅，郭释黄，并非也。”[110]这里，杨氏没有注意到“黄”和“寅”在形体上的历史关系。罗振玉所举之“”出现在第二期，而杨树达所举干支“寅”字出现在第五期。通过前面的分析我们知道，第二期的“黄”、“寅”区别是明显的，而杨树达先生拿第五期的“寅”字的形体与第二期的“”之“”相比附，得出的结论自然就会有问题。再者，据笔者考查，关于“黄尹”的卜辞直出现在前两期，前两期的“黄”和“寅”区别明显，并不相混。因此，王国维、郭沫若释“黄尹”是，而杨氏改释“寅尹”非。

然而，后来的学者和杨树达先生一样，都没有注意到“黄”、“寅”的错综关系，继续延续着杨氏的错误。这样释出来的“寅尹”是存在问题的。

另外，“寅尹”见于《殷墟甲骨文刻辞类纂》者仅两例，分别为合集3097片与合集3098片（见附图），且皆为合文，编订者将它们收在“寅尹”条下，释文却隶定为“效”，这本身就自相矛盾。此两例是否为“寅尹”很值得商榷。细审原拓片，这两个合文左为，右为。《类纂》将其隶定为“效”，是忽略了“”（攴）与“”（尹）的区别；而释为“寅尹”之合文，又没有注意到（寅）与（交）的差异：“交”象正面人双腿相交之形，两腿呈弯曲状，“寅”则借“矢”为之，矢之括与两腿的分别是明显的。笔者认为，这两处都应该是“黄寅”的合文。首先，我们从形体上进行分析。这两个形体右半为“尹”，确定无疑。左半确实象“交”，但是应该是“黄”字的误刻。在甲骨文中“交”、“黄”相混的情况是存在的。如合集9177正之烄作“”，字所从之“交”与本期的“黄”字相近似；又如，京津636之“黄”作“”，与“交”字相近似。我们再看文例，此两片的文例分别为：

丙戌卜，争贞：取丁人娘？（合集3097）（附图，1）

入丁人？（合集3098）（附图，2）

与之相对应，“黄尹”也有相同的文例：

癸卯卜，贞……甲令取黄丁人？七日。（合集 22）

贞：于乙亥入黄尹丁人（合集 3099）

根据以上的分析，我们认为：甲骨文中“寅尹”是不存在的，过去释“寅尹”的卜辞应该都是“黄尹”。

附图：

1
合集 3097

2
合集 3098

注释：

[1]见《甲骨文字诂林》2537 页姚孝遂先生按语，中华书局 1996 年版。

[2]其10]见《甲骨文字诂林》2534 页，“黄”字条下杨树达先生语，中华书局 1996 年版。

[3]第十三册为摹本，调查的时候没有将其包括在内。

[4]见《甲骨文字诂林》2527 页，“矢”字条下罗振玉语，中华书局 1996 年版。。

[5][6]郭沫若：《甲骨文字研究》，科学出版社，1962 年版，195~196 页。

[7]郭沫若：《金文丛考·释黄》，科学出版社，1956 版，180~192 页。

[8]唐兰：《用青铜器铭文来研究西周史》，载于《文物》，1976 年 6 期，35 页。

[9]见《甲骨文字诂林》2532 页，“黄”字条下王襄语，中华书局 1996 年版。

The Relation of 黄 and 寅 in Jia-Gu Wen and The Other Questions Related to Them

Qin Xiao-hua
(Guangzhou,Guangdong 510275)

Abstract: Some scholar think that 矢,寅,黄 have the same origin, which all come from the character 矢. The author investigates thoroughly these characters in Jia Gu Wen Collection of Material and differentiates and analyzes the relation of them from the angle of historical period. At last, the writer reaches a conclusion that the parts formed the character 黄 is not 矢 but 大 and the character 黄 hasn't the same origin with two others. Besides, 寅尹 shoulde be the mistke of 黄尹.

Key words: 黄；寅；矢；jia Gu Wen；黄尹；寅尹

（责任编辑、校对：陈丽新）

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1524

Date: 2008-11-13

Title:	**Report of the Old Hanzi Expert Group**
Source:	**Old Hanzi Expert Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**6**
Medium:	**Electronic**

The Old Hanzi Expert Group held its meeting from 10 Nov. to 13 Nov. 2008 in Kunming, China. The Old Hanzi Expert Group completed its work according to the schedule set in its Editorial Report (N1460) passed at IRG#30. The Expert Group finished discussing and reviewing the character tables submitted by China and TCA (Shuowen radicals 281 to 400), and discussed N1522 submitted by Japan national body. The Group also finalized its future work plan.

This report is organized as follows:

- **Discussion and reviewing of IRG N1474 consolidated Old Hanzi documents.**
- **Revisit of the Old Hanzi Principles and References (Version 2)**
- **Discussion of N1522 submitted by Japan(JNB)**
- **Proceed to the next stage of the work plan**

1. Work plan

1.1 Progress(選字進度)

The Old Hanzi Expert Group will proceed to Radicals 401 to 540 (the last radical in Shuowen) after IRG#31 and will finalize its work on this part at IRG#32.

The division of work is as follows:

TCA	Radicals 401 to 470 (N1557)
China	Radicals 471 to 540 (N1556)

1.2 With the completion of work under Work Plan 1.1, the Old Hanzi Expert Group will generate a complete list from Radicals 1 to 540 for a comprehensive review.

- Submit the character table by 2 May 2009 to the Project Secretary (Selena Wei).

  5 月 2 日以前提交字形表

- The Project Secretary will generate a consolidated table (IRG N1558) between 9 May and 13 May 2009.

  5 月 9-13 日進行整合字表 IRG N1558

- The Project Secretary will send out the consolidated table (IRG N1558 radicals 401-540) by 13 May 2009 for discussion at IRG#32.

  5 月 13 日寄出整合後之字表 IRG N1558 (401-540)，並於 IRG#32 會議中進行討論。

- The Old Hanzi Expert Group will generate a complete list from Radicals 1 to 540 for a comprehensive review by IRG #33.

2. Discussion of N1522 submitted by Japan(JNB)

The Old Hanzi Expert Group's response to IRG N1522 is as follow:

In response to the questions (2.1 to 2.7) in N1522 submitted by Japan(JNB), the Old Hanzi Expert Group would like to provide the following feedback:

2.1 User community: To all in the world who are interested in learning and studying more about Old Hanzi.

2.2 IRG N1271 is still valid. Since no revision has been made, it is not necessary to produce N1471 as promised in IRG M30.8. Results of discussion have been recorded in the system developed by the Old Hanzi Expert Group. (Refer to 2.6 and 2.7 below.)

2.3 The issue of variants is handled according to N1271. When we have finished our project, we will submit our work to WG2 through the IRG. We would be happy to consider the views of the IRG and WG2 on our work.

2.4 We will use the original glyph form. The production of font will be done according to the established practice of the IRG.

2.5

A. There are 540 radicals in Shuowen. For Oracle Bone inscriptions not found in Shuowen, they will be placed after the 540 Shuowen radicals. (Refer to N1271)
B. If an Oracle Bone inscription may be mapped to several modern characters, it will be mapped to all these modern characters, not mapped to just one. (Refer to N1271)

2.6 Results of discussion have been recorded in the system developed by the Old Hanzi Expert Group. (Refer to 2.2 above.)

2.7 Principles and references are found in N1271 (which is a valid, standing document). Since no revision has been made, it is not

necessary to produce N1471 as promised in IRG M30.8. (Refer to 2.2 above.)

3. Conclusion

The Old Hanzi Expert Group has concluded that there is no need to put forward new principles, so the Group will continue to rely on N1271 (Version 2). The Expert Group welcomes specific amendments to its principles from all member bodies and other experts. It will consider suggestions received and explore the need of producing Version 3 of N1271.

Appendix: A character table of Oracle Bone Inscriptions (281-400) (agreed after discussion).

Members attended the meeting were Dai Hong, Li Guoying, Zhao Cheng, Zhou Xiaowen, Chen Shuangxin, Wei Lin Mei(Selena), Chuang Suh-chyin, Huang Fang-chuan, Jung Bor Sheng, Chi Hsiu Sheng, Kim Jinhe, Tang Yingmin, Suzuki Atsushi, Caoying, Liu Zhiji and Peter Cheng Wai-hong.

End of document

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1558
Date: 2009-06-15

Title:	**Consolidated Old Hanzi Submission (Radical 401 to 570 in Shouwen)**
Source:	**IRGN 1556(China) ,IRG N557 (TCA-CMEX))**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**1**
Medium:	**Electronic**

This submission document is according as the consolidation of two documents. Both of the documents are China IRG N1556 and TCA IRG N1557.

The method of integrating the data from these two documents is using some ordinal key fields to sort all the data in one format. First of all, using the filed of SW Radical Number to group all the relate records. If there are too many records relate to one SW Radical Number, and then will use the second field which is Corresp. Modern Character, to sort again. Finally, if after using the above two keys still couldn't organize all the records into a fine structure, then will use "ID" field to be the last way to order the rest of the records.

Documents submitted to the IRG #32 Meeting are listed below:

The electronic files are a PDF file, for Oracle Bone inscriptions for reviewing (IRGN1558-Appendix A1-Consolidated Old Hanzi Submission).

End of document

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N1588
Date: 2009-06-18

Title:	**Report from the Old Hanzi Expert Group**
Source:	**Old Hanzi Expert Group**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**2**
Medium:	**Electronic**

The Old Hanzi Expert Group held a group meeting in Hong Kong SAR, China from 15 to 18 June 2009. The meeting recapped the last editorial report at IRG#31 and discussed a number of issues in detail.

1. Discussion on IRGN 1558

(a) Members completed the discussion on the Shuowen radicals 401-540 with reference to IRGN 1271 (Old Hanzi Principles and References Version 2). After four years of review on the Oracle Bone inscriptions, the work on Shuowen radicals 1-540 has been completed.

(b) The second round of review on the Oracle Bone inscriptions will commence next. The review results are tentatively planned to be submitted to the IRG by the end of 2010.

2. Revised work plan for the development of Old Hanzi inscription repertoire

(a) Wei Lin-Mei (Selena) will prepare a list of the Oracle Bone inscriptions based on the Shuowen radicals 1 to 540. A sample is attached below:

ID	Rep. Script/Glyph	Original Shape/Glyph	Source	Period/ Epoch	Area/ Site	Material	SW Radical	SW Radical Number	Glyph Determ.	Corresp. Modern Char	Unifiable Shapes	Notes
G04700			A-35706	商代	河南安阳	甲骨	亥	540		亥		
G04703			A-34295	商代	河南安阳	甲骨	亥	540		亥		
G04702			A-34294	商代	河南安阳	甲骨	亥	540		亥		

The consolidated list will be emailed to all Old Hanzi Expert Group members and uploaded to the IRG website by 20 July 2009.

(b) Members and all interested parties are requested to send the comments on the missing attributes of about 2,000 Oracle Bone inscriptions to Selena for consolidation by 30 September 2009.

(c) Selena will prepare a consolidation list to the IRG Rapporteur and members by 20 October 2009 for discussion at IRG#33.

Work Schedule

20 July 2009	Selena to prepare a consolidated list of the Oracle Bone inscriptions (IRGN 1603)
30 September 2009	Comments from members to reach Selena
20 October 2009	Selena to prepare a consolidated list with comments incorporated for discussion at IRG#33 (IRGN 1604)
At IRG#34	Submit the completed consolidated list to the IRG

Appendix: A consolidated discussion on Shuowen radicals 401-540 of Oracle Bone Inscriptions

Members attended the meeting were Li Guoying, Wei Lin-Mei (Selena), Zhao Cheng, Jung Bor-Sheng (鍾柏生), Zhou Xiaowen, Caspar Lee (李欽堯), Lai Sio Leng and Au Yung-yee, Bryan.

End of document

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1615

Date: 2009-11-13

Source:	Report from the Old Hanzi Expert Group
Title:	Old Hanzi Expert Group
Status :	Input to IRG
Actions required	
Distribution:	IRG Members and Ideographic Experts
Reference:	
No. of pages:	3
Medium :	Electronic

The Old Hanzi Expert Group held a group meeting in Beijing, China, from 11 to 13 November 2009. The meeting recapped the last editorial report at IRG#32 and discussed a number of issues in detail.

1. Discussion on IRGN 1603

(a) Members completed the primary selection of 1-540 radicals of Shuowen Jiezi (說文解字).
完成了1至 540首之初選之工作

(b) The second round of review on the Oracle Bone inscriptions was finished, characters left before the group meeting were further dissussed.
進行甲骨文第二輪之審查，對之前遺留之字，做了進一步之討論.

2. Future Work Plans:

- Work items:
  (1) The third round of review on radicals 1-540.
  對 1 至 540 部首之全面之審核
  (2) The contents of review include adding missing characters, merging duplicate characters, and deleting characters that are not necessarily included.
  審核之內容包含缺字之增補，重複字之合併，不必要收錄之形體予以刪除。
  (3) Full-scale ordering or all characters.
  統一排序

(4) Completing determining glyphs and corresponding modern characters.
完成隸定及對應今字之工作

(5) Representative scripts (glyphs) revisions.
摹寫修定

(6) Writing descriptions of proposed documents.
編寫提案文件說明

- Work Projects

(1) Add the relevant attributes of rep. scripts (glyphs), glyphs determinations, and corresponding modern characters.
補摹寫、隸定、對應今字之相關屬性

(2) Output the radical font summary V3.0 listing radicals 1-540.
產出 1 至 540 部首之字形總表 V3.0

(3) Submit to members and all interested parties for review.
提交專家審查

(4) Submit the comments of review.
提交審查意見

(5) Output the radical font summary V4.0 listing radicals 1-540, which will be discussed by the Old Hanzi specialist group.
產出 1 至 540 部首之字形總表 V4.0，於古漢字專家組討論

(6) Output the radical font summary V5.0 listing radicals 1-540.
產出 1 至 540 部首之字形總表 V5.0

(7) Write descriptions of submitted documents.
撰寫提交文件說明

3. Work Schedule

(1) 10th February, 2010, completing the supplementation of missing attributes such as the relevant attributes of rep. scripts (glyphs), glyph determinitions and corresponding modern characters. (The completed files will be sent to Ms. Wei Lin-mei.)
2010 年 2 月 10 日，完成遺漏屬性之補充工作，如摹寫、隸定、對應今字之相關屬性，並將修定後之檔案提交 selena。

(2) 29th March, 2010, Wei Lin-mei (Selena) will prepare a consolidation list of radicals 1-540 (V3.0) to the IRG Rapporteur and members for discussion.
2010 年 3 月 29 日，selena 須產出 1 至 540 部首之字形總表 V3.0 版，並提交給 IRG 會員。

(3) 30th May, 2010, members and all interested parties are requested to send

the comments of review to Selena.
2010 年 5 月 30 日，古漢字專家須於將審查意見提交給 Selena

(4) June, 2010 (one week before the IGR#34 meeting), Selena shall output the summary of the experts' comments for the Old Hanzi experts group discussions.
2010 年 6 月 IRG # 34 次會議前一星期，Selena 產出專家意見彙集表，提供古漢字專家組討論

(5) 30th September, 2010, Old Hanzi experts shall complete the task of reviewing the table of all Oracle Bone inscriptions.
2010 年 9 月 30 日，古漢字專家完成甲骨文部全部字表之複審工作

(6) 20th October, 2010, Selena shall output a consolidation list of radicals 1-540 (V4.0) to the IRG Rapporteur and members for review.
2010 年 10 月 20 日，Selena 產出 1 至 540 部首之字形總表 V4.0，提供古漢字專家進行審查。

(7) Discussion of V4.0
- Arrange additional Old Hanzi experts group meeting to review Oracle Bone inscriptions V4.0 one more time prior to the IRG#35 meeting.
  在 IRG#35 次會議之前，多安排一次古漢字專家組會議，針對甲骨文 V4.0 版再次進行審查。
- Tentative venue for the meeting is Taipei.
  地點暫定台北

(8) V4.0 review shall be completed in the IRG#35 meeting, as well as output of V5.0, and prepare to submit documents to WG2 for Oracle Bone inscriptions coding.
IRG # 35 會議，完成 V4.0 之審查工作，並產出 V5.0，準備提交 WG2 做為甲骨文之編碼工作。

Drafted by Ms. Wei Lin-mei, 2009-11-13
Edited by Mr. Chen Zhuang, 2009-11-24

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N1649
Date: 2010-06-14

Title:	**The third round of review on the Oracle Bone inscriptions from China**
Source:	**China**
Status:	**Input to IRG**
Action:	
Distribution:	**IRG Members and Ideographic Experts**
Reference:	
No. of pages:	**1**
Medium:	**Electronic**

Followed the IRG N1615 the third round of review on the Oracle Bone inscriptions 3520 characters was finished by China.

The main work is as follows:

(1) The third round of review on the Oracle Bone inscriptions 3520 characters.
对其中的 3520 个甲骨文字形进行第三次审核

(2) The contents of review include adding missing characters, merging duplicate characters, and deleting characters that are not necessarily included.
审核的内容包含缺字之增补，重复字之合并，不必要收录之形体予以删除。

(3) Full-scale ordering or all characters（with TCA）
（与 TCA 部分）统一排序

(4) Completing determining glyphs and corresponding modern characters.
完成隶定及对应今字之工作

(5) Representative scripts (glyphs) revisions.
摹写修定

Documents submitted to the IRG #34 Meeting by China consist of seven appendixes.

End of document

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRGN1703 (2010-06-24)

Title:	Report of the Old Hanzi Expert Group
Source:	Old Hanzi Expert Group
Status:	Input to IRG
Action:	
Distribution:	IRG Members and Ideographic Experts
Reference:	
No. of pages:	3
Medium:	Electronic

At the plenary session of IRG#34 on 2010-06-22, the IRG asked the Old Hanzi Expert Group to provide the selected glyph collection of Oracle Bone scripts with tracking information (why a glyph was dropped, why a glyph was unified with another one, and so on). However, providing all tracking information was considered not practical. Therefore, the Old Hanzi Expert Group agreed to provide only the following tracking information:

- Why a glyph was dropped for "Deleted" glyphs.
- Which glyph was unified with another glyph.

Also, the dated snapshot of the database was also required to record when the decision was made.

Japan requested that the data format of the selected glyph collection should be some spreadsheet format editable by popular applications (like Excel) or some XML, because it would be difficult to examine PDF data with electronic tools. After discussion, the Old Hanzi Expert Group agreed to provide all database materials to JTC 1/SC 2/WG 2.

The issue of copyright should be cleared before submission to the IRG. The source glyph image and the partial scan of the rubbing of Oracle Bone might cause copyright problems. Both China and TCA were concerned that the representation glyph could be some calligraphic opus and might cause copyright issues. Japan was concerned that the database without original and representation glyph image was difficult to discuss, and difficult to use for font production to publish JTC1 standard.

Because the publishing of source and representation glyph images as individual pictures could cause copyright issues, the Old Hanzi Expert Group agreed to submit only 2 documents to the IRG:

1. PDF including all texts and images in the database.
2. Excel including all texts but excluding images in the database.

Also Japan proposed to add new property for each glyph in the database--unique IDs for the collected glyphs. It should not be changed even if the glyph was dropped in the working process. The Old Hanzi Expert Group agreed to add unique IDs for every glyph in the database and keep them. The Project Secretary would add unique IDs after this meeting.

Japan was concerned that the current working process might be different from the process described in IRG N1271 (revision 2), IRG N1460, IRG N1473 and IRG N1524. Japan proposed to produce Version 3 to reflect the current working process. The Old Hanzi Expert Group appointed Professor Atsushi Suzuki to consolidate the materials for preparation of Version 3.

WORK SCHEDULE

1. Project Secretary (Selena Wei) to add unique IDs to the glyphs to generate new Version 2 of the database to include all data by 2010-August-31, and distribute to all members of the Old Hanzi Expert Group. Version 2(IRGN 1705) would be submitted in PDF format.
2. Project Secretary to submit Excel file to all members of the Expert Group by 2010-08-31.
3. Professor Atsushi Suzuki to submit a consolidated document(IRGN 1706) about the Old Hanzi Expert Group’s principles and references for discussion by 2010-09-30.
4. Members of the Expert Group to submit feedback on item 3 to the Project Secretary by 2010-10-22.
5. Project Secretary to consolidate the feedback on item 4 to all members of the Old Hanzi Expert Group by 2010-11-05.
6. Members to discuss their feedback on item 1 at IRG#35.

Members attended the meeting were Li Guoying, Wei Lin-Mei (Selena), Zhao Cheng, Jung Bor-Sheng (鍾柏生), Dai Hong, Zhou Xiaowen, Atsushi Suzuki, Toshiya Suzuki, Tatsuo Kobayashi, Chen shuangxin and Peter Cheng (鄭偉康).

End of Document

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC 1/SC 2/WG 2 IRG Nxxxx

Date: 2010-10-22

Source:	SUZUKI Toshiya, Faculty of Integrated Arts and Science, Hiroshima University
Title:	Comment on IRG N1706 (Consolidated P&P of Old Hanzi) from Japan NB
Actions required:	For consideration at IRG #35
Distribution:	IRG Members and Old Hanzi Experts
Medium :	Electronic

Japan reviewed IRG N1706 and found that the current P&P (IRG N1271) is specific to Oracle Bone script only; it is not applicable to Bronze script and Small Seal script. Japan proposes to improve P&P for Bronze and Small Seal scripts, before starting the glyph collecting works for these scripts.

When Old Hanzi expert group was working for Oracle Bone script, many important rules (especially for glyph determination and ordering) were added or changed during the glyph selection. It made the reliability of the database unstable; Oracle Bone experts had to review the glyph collection twice, and the full disclosure of the database, IRG N1705, had delayed 1 month after the promised deadline.

Considering this experience on Oracle Bone script, the stabilization of P&P is clearly essential before starting glyph selection work. In addition, it is important for IRG experts to review the results of Old Hanzi experts' discussion, the template format of the discussion record for each glyph must be approved by IRG. All proposed glyphs, cancellation, unification, postpone (to next extension) are trackable in IRG documents. The output from Old Hanzi should be described on the same level.

(end of document)