

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




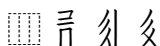



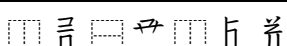

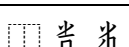
**Doc Type:** Working Group Document  
**Title:** Proposal to encode five additional Tangut ideographs  
**Source:** Andrew West, Viacheslav Zaytsev, Sun Bojun  
**Status:** Individual Contribution  
**Action:** For consideration by JTC1/SC2/WG2 and UTC  
**Date:** 2016-04-21

### 1. Introduction

Since the finalization of the encoding of the set of 6,125 Tangut ideographs in ISO/IEC 10646:2014 Amendment 2 and Unicode version 9.0, five additional Tangut ideographs that were not included in any of the sources used for the original Tangut proposal have been identified. None of these characters can be considered to be glyph variants of existing encoded characters, and so are not appropriate for representation as variation sequences. We therefore propose adding these five characters to the end of the existing Tangut block, in the range 187ED..187F1 (see Table 1).

We also discuss three characters that occur as phonetic glosses in *Pearl in the Palm*, which may be candidates for encoding (see Section 3); and four unidentified Tangut characters used by Prof. Kychanov, which we have rejected for encoding (see Section 4).

**Table 1: Proposed Characters**

Code Point	Glyph	IDS	Radical/ Strokes	Source
U+187ED			195.14	Li Fanwen 2012 #6075
U+187EE			308.14	Li Fanwen 2012 #6076
U+187EF			415.13	Li Fanwen 2012 #6077
U+187F0			308.17	This document
U+187F1			415.11	This document

## 2. Tangut Characters Proposed for Encoding

### 2.1 U+187ED..187EF 𪚱 𪚲 𪚳

Prof. Li Fanwen has identified three previously unknown Tangut characters in the Tangut translation of the *Classic of Filial Piety* (Chinese *Xiào Jīng* 孝經). These characters are found in the cursive manuscript version in the collection of the Institute of Oriental Manuscripts in Saint Petersburg [Tang. 1, Inv. No. 1, old Inv. No. 2627], as shown in Fig. 28–29 or Fig. 30–31 (see Appendix).

The regular forms of these three characters were published by Li Fanwen in 2012, in a paper dedicated to Prof. Kychanov under the heading “Three new Tangut characters” (see Fig. 1).

Fig. 1: Li Fanwen 2012 (Тангуты в Центральной Азии) p. 212

1192 80 𪚱 6075	(牙音 ṇa 音雅) ода (кит. я 雅)	𪚱 𪚲 𪚳 𪚴 𪚵 𪚶 𪚷 𪚸 𪚹 𪚺 𪚻 𪚼 𪚽 𪚾 𪚿 Кит. 《大雅》云：『無念爾祖，聿脩厥德。』 (《孝經》開宗明義章 第一)
1224 82 𪚲 6076	(正齒音 zja 音社) в соч. алтарь (кит. шэцзи 社稷)	𪚲 𪚳 𪚴 𪚵 𪚶 𪚷 𪚸 𪚹 𪚺 𪚻 𪚼 𪚽 𪚾 𪚿 Кит. 富貴不離其身，然後能保其社稷，而和其民人 (《孝經》諸侯章 第三)
9824 22 𪚳 6077	(齒頭音 tsjwak 音稷) в соч. алтарь (кит. шэцзи 社稷)	𪚲 𪚳 𪚴 𪚵 𪚶 𪚷 𪚸 𪚹 𪚺 𪚻 𪚼 𪚽 𪚾 𪚿 Кит. 富貴不離其身，然後能保其社稷，而和其民人 (《孝經》諸侯章 第三)

Li Fanwen subsequently included these three characters as nos. 6075, 6076 and 6077 in the 2012 abridged edition of his 2008 *Tangut-Chinese Dictionary* (see Fig. 2). Nos. 6076 and 6077 are written slightly differently in these two publications, but based on the cursive forms in the original manuscript, and their presumed character construction, it would seem that the character form for no. 6077 in the 2012 dictionary is correct, but the character form for no. 6076 in the 2012 dictionary is slightly wrong, and should be as given in the 2012 paper, with 𪚲 rather than 𪚲 | in the middle.

No. 6075 𪚱 is used to transcribe Chinese *yǎ* 雅 in the title of the *Dà Yǎ* 大雅 section of the *Book of Songs*.

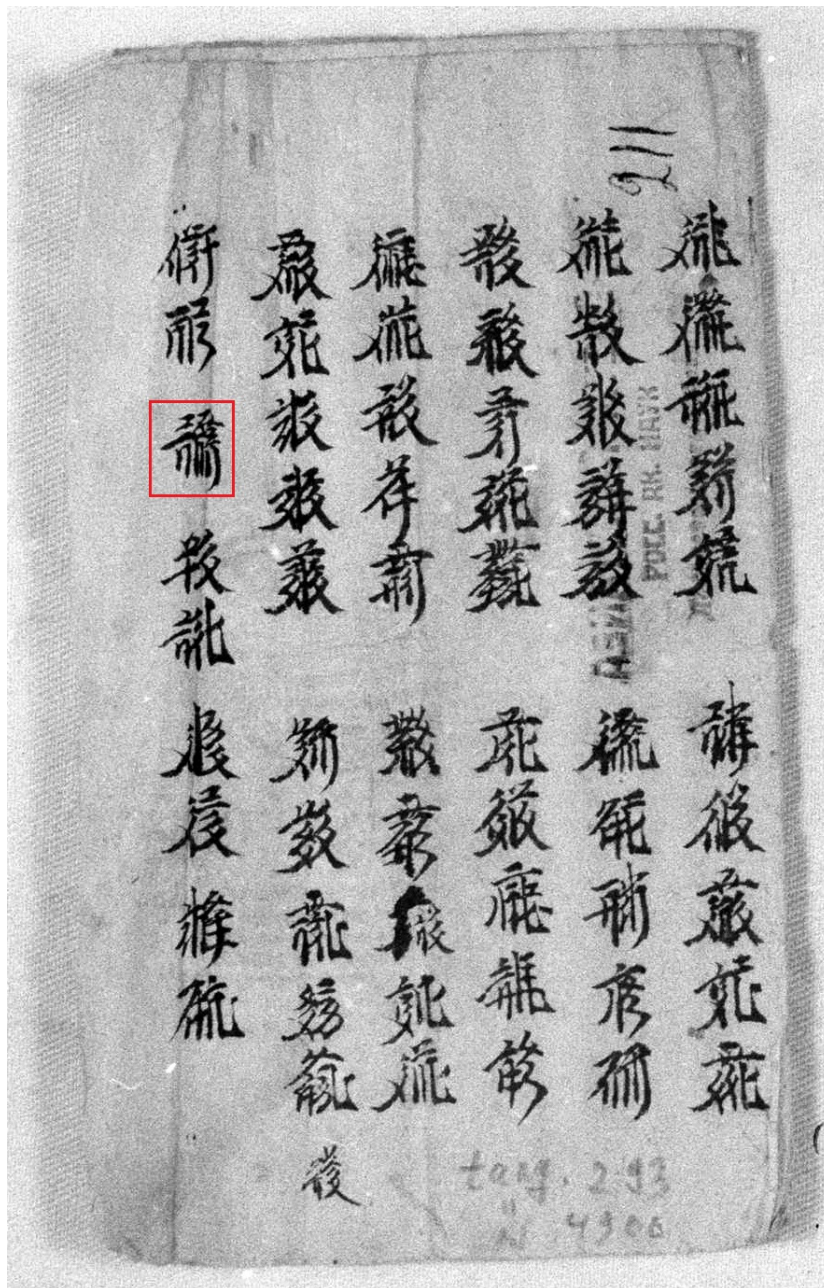
Nos. 6076 and 6077 𪚲 𪚳 are used to transcribe the Chinese word *shèjì* 社稷 ‘altars to the land and the grain’.



2.2 U+187F0 𪛗

Prof. Sun Bojun has drawn attention to a recently-identified Tangut character in a Tangut Buddhist manuscript in the collection of the Institute of Oriental Manuscripts in Saint Petersburg [Tang. 293, Inv. No. 1149, old Inv. No. 4900] (Tangut title 𪛗𪛗𪛗𪛗𪛗𪛗; Chinese title *Fānyán Jīngāngwáng Chénggēn* 番言金剛王乘根 on p. 594 of Prof. Kychanov's 1999 catalogue, No. 697).

Fig. 3: Tang. 293, Inv. No. 1149, old Inv. No. 4900



𪛗

This character occurs at the end of the last page as part of the Tangut transcription of the Sanskrit title of this text, the *Vajrayānamūlāpatti* ('Root downfalls of Vajrayana') by Aśvaghōṣa (see Fig. 3):

y a n a      m u l a      p a t i  
 𑖑𑖑𑖑 𑖑𑖑 𑖑𑖑𑖑 𑖑𑖑𑖑 𑖑𑖑𑖑

Sanskrit language: *Vajrayānamūlāpatti*

The character 𑖑𑖑 stands for the Sanskrit word *vajra*, and is constructed from the left side of 𑖑 (ba), the top and left of 𑖑 (dzji), and the left side of 𑖑 (rjar), the three parts together transcribing *vajra*. In the mid 14th-century Tangut inscription on the east wall of the Cloud Platform at Juyong Pass in Beijing, the Sanskrit word *vajra* is transcribed once using these three characters 𑖑𑖑𑖑 (see Fig. 4), and so 𑖑𑖑 in the manuscript can be regarded as a portmanteau abbreviation for these three characters.

**Fig. 4: Juyong Pass East Wall col. 10: rubbing and Nishida's transcription**  
 (Murata Jirō (ed.) *Kyoyōkan* 居庸關 (1957) rub. IV & p. 182)



𑖑𑖑 𑖑𑖑 𑖑𑖑 𑖑𑖑 𑖑𑖑 𑖑𑖑  
 \*le bha      \*ra \*ga — \*ya  
 vajra-kāya-

𑖑𑖑𑖑

## 2.3 U+187F1 𪛗

In the past some modern scholars considered the component 𪛗 to be a cursive variant of 𪛗, and characters with the 𪛗 component were sometimes normalized to use the 𪛗 component. For example, in Li Fanwen's 1997 *Tangut-Chinese Dictionary* the characters U+178EB 𪛗 and U+178C5 𪛗 are both written identically, using the 𪛗 component (see Fig. 5).

Fig. 5: Li Fanwen *Xià-Hàn zìdiǎn* 夏漢字典 (1997 ed.) p. 705

2942 [牙音 ŋwu 2.1 音兀]  
<sup>74</sup>  
 𪛗 surname  
 3797 [兀]族姓也。(音)  
 𪛗 𪛗 ŋu 2.1 lhi 2.10 [兀叻](族姓)(同  
 20B2)。  
 [齒頭音 tsəj 1.40 𪛗 𪛗 𪛗 祖來切 音  
 栽]  
 𪛗 small; little; yong  
 3798 小、少、幼也。(形)  
 𪛗 𪛗 zǐə 1.92 tse 1.40 [邪栽]稀少(同  
 33B5)。

However, in the revised 2008 edition of Li Fanwen's *Tangut-Chinese Dictionary* the two characters U+178EB 𪛗 and U+178C5 𪛗 are clearly distinguished with different right hand components (see Fig. 6).

- 2942 〔牙音  $\eta wu$  2.1 音兀〕  
74  
𪛗 a surname  
3797 〔兀〕族姓也。(音)  
𪛗𪛗  $\eta u$  2.1 lhi 2.10〔兀叻〕(族姓)(同 20B2)。  
𪛗𪛗𪛗〔兀〕:族姓(同丁21A72背注)。
- 𪛗 〔齒頭音  $tsəj$  1.40 𪛗𪛗𪛗𪛗祖來切 音栽〕  
3798 small; little; young  
小、少、幼也。(形)  
𪛗𪛗  $zĩə$  1.92  $tse$  1.40〔邪栽〕微少(同 33B5)。

It seems that the character U+18307 𪛗 unifies two different characters with the 𪛗 and 𪛗 components.

- 𪛗  $niə$  (Initial class III, rising tone, rime 63) : ‘an ear of grain.’
- 𪛗  $ku$  (Initial class V, rising tone, rime 4) : either a transcription of the Chinese surname Gu 顧, or ‘millet sprouts’.

These are both rare characters with few attestations, and in some modern dictionaries the two characters have been treated as a single character. This is the case in both the 1997 and 2008 editions of Li Fanwen’s *Tangut-Chinese Dictionary*, where the two characters are both included under entry No. 5990 (see Fig. 7 and Fig. 8). As a result of this mistaken unification, the two characters have been encoded as a single character (U+187F1).



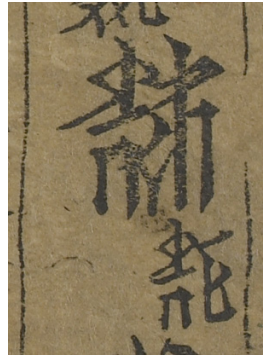


In original Tangut texts the two meanings and readings of these characters are distinguished by their glyph forms. In the *Homophones* (*Tóngyīn* 同音), where the character is read as *niō* ‘ear of grain’, the right hand component is clearly U+188BF 𐞪 (4 strokes) in all editions (see Fig. 9).

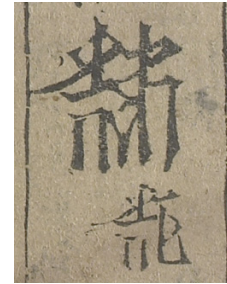
**Fig. 9: L5990 [*niō*] in editions of *Homophones***



A 16A72



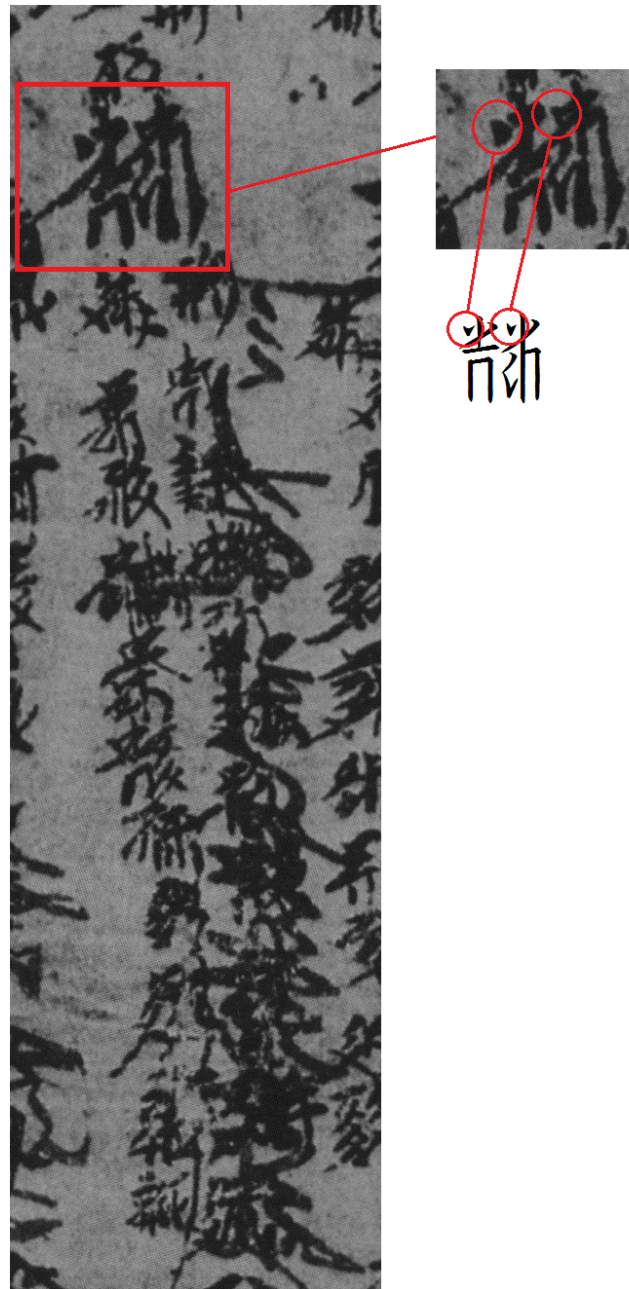
B2 17A47



D 17A47

However, in the manuscript *Combined Edition of Homophones and Sea of Characters* (*Tóngyīn Wénhǎi Bǎoyùn Hébiān* 同音文海寶韻合編), where the character is read as *ku*, the right hand component is U+18927 𐞫 (5 strokes) (see Fig. 10).

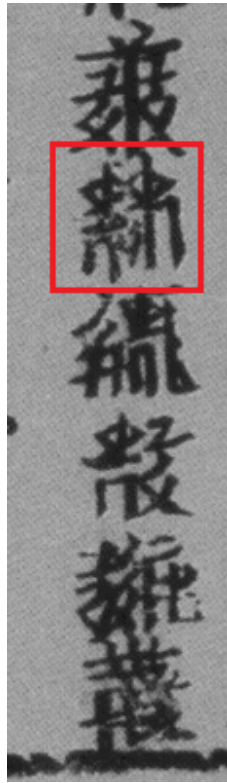
Fig. 10: L5990 [ku] in *Combined Edition of Homophones and Sea of Characters*



A 13.111

Although the right hand component is not entirely clear in this manuscript text, the character composition given beneath the head character states that the character is constructed from the left side of 'rice' and the bottom part of 'sprout' 𪎭 𪎮 𪎯 𪎰. This indicates that the right hand of this character (*ku*) is U+18927 𪎯 (bottom component of 𪎰), not U+188BF 𪎮 as shown in *Homophones* for the character read *niō*. This same glyph form (𪎯) also occurs in one of the *Odes*, where from context it would seem to represent the *ku* reading ("millet, barley and wheat") (see Fig. 11).

Fig. 11: L5990 [ku] in *Odes* No. 5 folio 6a



In light of this evidence, we propose encoding a new character with the 𪛗 glyph shape. The existing U+18307 𪛗 should be unchanged in glyph shape and source reference.

### 3. Unencoded Characters in Pearl in the Palm

The Tangut-Chinese bilingual glossary, *Pearl in the Palm* (Chinese *Fān Hàn héshí zhǎngzhōngzhū* 番漢合時掌中珠; Tangut <sup>2</sup>mi<sub>4</sub> <sup>1</sup>zar<sub>1</sub> <sup>1</sup>dzen<sub>4</sub> <sup>1</sup>bu<sub>4</sub> <sup>1</sup>paq<sub>4</sub> <sup>2</sup>gu<sub>1</sub> <sup>2</sup>niq<sub>4</sub> 𐞇𐞏𐞐𐞑𐞒𐞓), is one of the most well-known Tangut texts, and was the key to deciphering the Tangut script. It is important that all the Tangut characters in this text are encoded. However, several Tangut characters used as phonetic glosses for Han ideographs in this text are not included in modern Tangut dictionaries, and are therefore not currently encoded. Three such characters are discussed below.

The omission of these characters in modern dictionaries may in part be due to less attention having been paid by modern scholars to the Tangut phonetic glosses of Chinese characters than to the Chinese phonetic glosses of Tangut characters. For example, in Nishida's 1964 transcription of *Pearl in the Palm* he entirely omits the Tangut phonetic glosses for Chinese characters (see Fig. 12), and so Nishida's work is not further cited below.

Fig. 12: Nishida 1964 vol. 1 p. 220 (Folio 33 Col. 6)

336	𐞇𐞏	𐞇𐞐
	賊	𐞇𐞐移則
	tsiər	Bzi
	𐞇𐞑	𐞇𐞒 鐵
	reins	bit

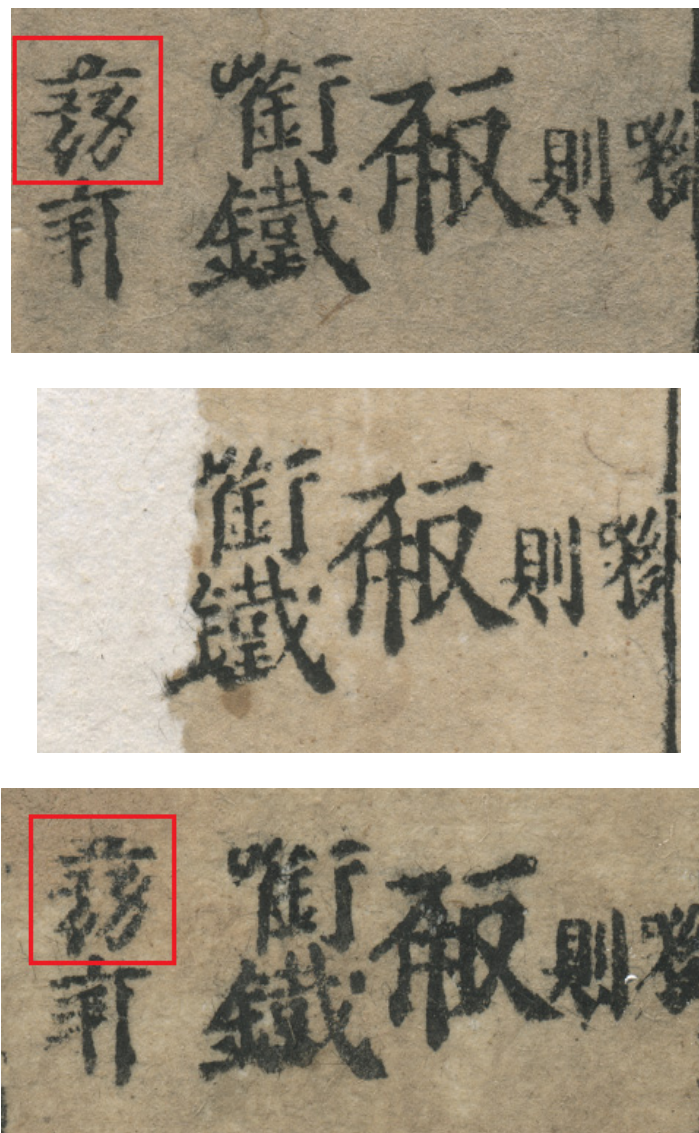
*Tangut phonetic glosses omitted (cf. Fig. 18)*

The characters discussed below may be variants of encoded characters which could be represented as standardized variation sequences, or they may be distinct characters which are candidates for encoding. However, they are not proposed for encoding at present as further research is required, and feedback from other Tangut experts is requested.

### 3.1 Folio 33 Col. 6 𪛗

The character 𪛗 occurs once in *Pearl in the Palm*, on folio 33 col. 6, as the Tangut phonetic gloss for the Chinese character *xián* 𪛗 'bit [for a horse]'. Three copies of a woodblock printed edition of this text that are held at the Institute of Oriental Manuscripts (IOM) in Saint Petersburg are shown below. Two of the manuscripts show the character, but it is missing due to damage in Ex 2. Ex 2 is the source for the 'A Edition' (甲種本) and Ex 1 is the source for the 'B Edition' (乙種本) of the facsimile reproductions published in *Heishuicheng Manuscripts Collected in Russia* vol. 10 (1999).

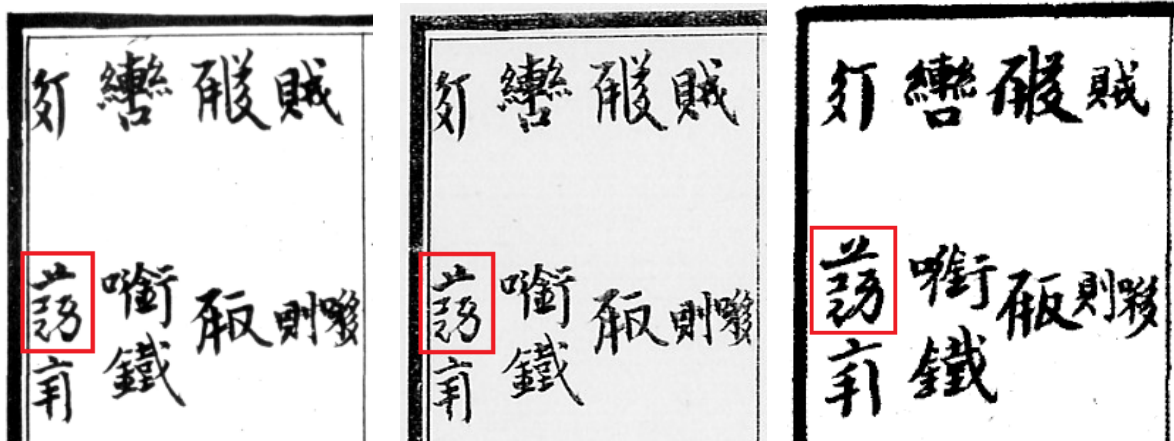
Fig. 13: *Pearl in the Palm* Tang 13 Ex 1, Ex 2 and Ex 3 folio 33 col. 6



*Images reproduced with permission of the Institute of Oriental Manuscripts*

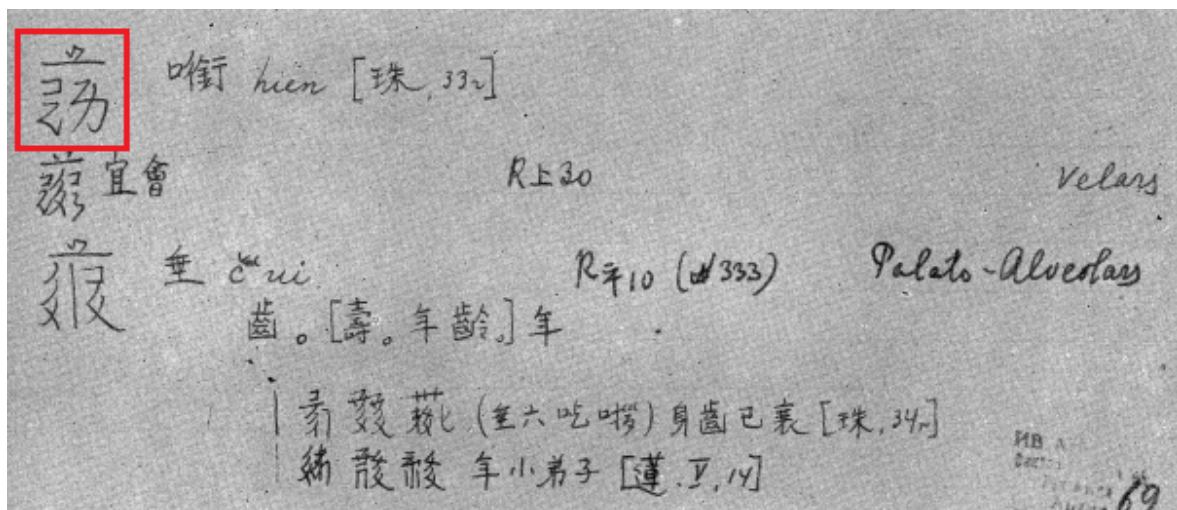
In 1924 Luo Fucheng 羅福成 published a lithographic print of a hand-written copy of the *Pearl in the Palm* (this was reprinted with appended tables of corrections and omissions in 1935). In 1963 Luo Fucheng's younger brother, Luo Fubao 羅福葆, published another hand-written copy made by Luo Fucheng in 1957. These editions all show the character as 𪛗 (see Fig. 14).

Fig. 14: *Pearl in the Palm* (1924, 1935 and 1963 editions) folio 33 col. 6



N. A. Nevsky's draft Tangut dictionary, published posthumously in 1960, also shows this character as 𪛗, referencing *Pearl in the Palm* folio 33 (see Fig. 15).

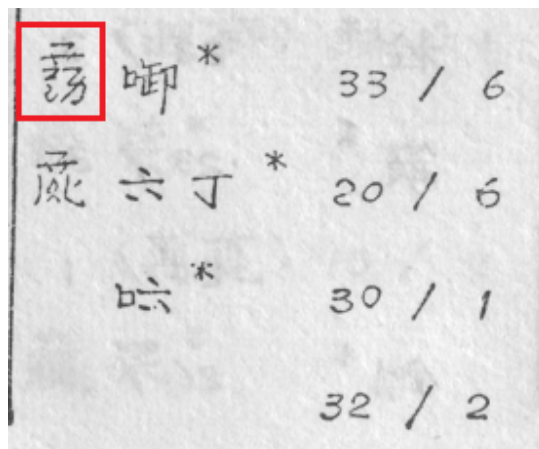
Fig. 15: Nevsky 1960 vol. 1 p. 292



Although the page of the *Pearl in the Palm* showing this character (folio 33) has been reproduced in facsimile several times (Kwanten 1982 p. 221; Huang et al. 1989 pp. 68, 140 and 151; Li Fanwen 1994 pp. 370 and 374; *Heishuicheng Manuscripts Collected in Russia* 1999 vol. 10 pp. 17 and 35), this character seems to have been overlooked in the dictionaries of Nishida, Kychanov and Li Fanwen.

It seems that recent Chinese scholars at least have considered 𪛗 in the text to be an error for 𪛗 (U+17C54, Li Fanwen No. 1502). This interpretation is first seen in the 1989 edition of *Pearl in the Palm* edited by Huang et al., where the character 𪛗 is silently corrected to 𪛗, and the corresponding Chinese character xián 𪛗 is silently corrected to xián 𪛗 (see Fig. 16).

Fig. 16: Huang et al. 1989 p. 192



In Li Fanwen's 1994 study of the *Pearl in the Palm*, he indicates 𪛗 as a correction for 𪛗 (see Fig. 17), and he corrects 𪛗 to 𪛗 in the transcriptions on pp. 80 and 440 (see Fig. 18 and Fig. 19), and in the index on p. 465.

Fig. 17: Li Fanwen 1994 p. 33

350	𪛗	𪛗	[𪛗]	33	6	4	1	769	𪛗	𪛗	[𪛗]	33	6	4	1
351	𪛗	𪛗	[𪛗]	33	6	4	2	770	𪛗	𪛗	[𪛗]	33	6	4	2
352	𪛗	𪛗	[𪛗]	33	6	4	9	771	𪛗	𪛗	[𪛗]	33	6	4	9
353	𪛗	𪛗	[𪛗]	33	6	4	9	772	𪛗	𪛗	[𪛗]	33	6	2	9

Corrections to 𪛗 in the 1924 and 1963 hand-copied editions

Fig. 18: Li Fanwen 1994 p. 80: 336

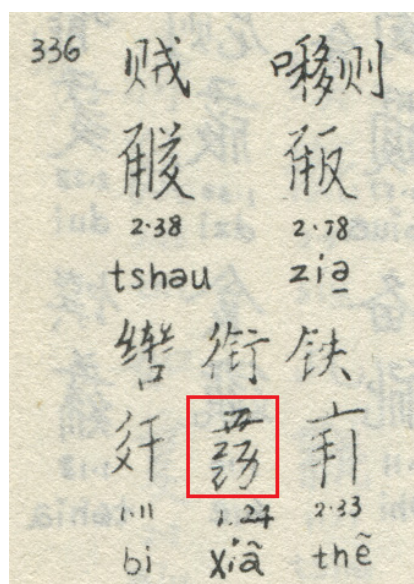
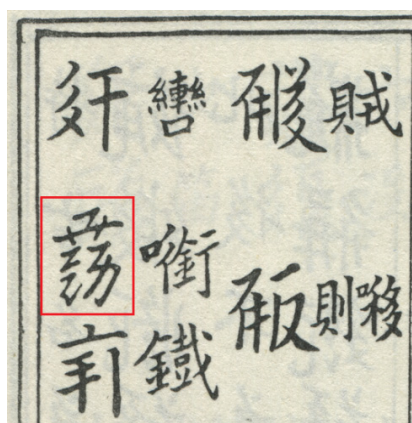


Fig. 19: Li Fanwen 1994 p. 440



Hand copy by Lú Tóng 盧桐?

At first sight the correction of 𪗊 to 𪗋 seems to be reasonable because 𪗋 is used to write the Chinese loanword *hān* 憨 ‘foolish’, and 憨 and 𪗊 would have been homophones or near homophones in the Chinese of the time (both approximately *xam*). Therefore it is plausible that 𪗋 could have been used to phonetically gloss the Chinese character 𪗊.

However, the correction of the written character in *Pearl in the Palm* may be unsound, as the components 𪗊 and 𪗋 are not normally interchangeable (WG2 N4722, the preliminary list of proposed variation sequences for 178 Tangut ideographs, does not include a single example of variation between 𪗊 and 𪗋). It is possible that 𪗋 is a special transcription character for 𪗊 used deliberately by the author of *Pearl in the Palm*. Marc Miyake has pointed out to us that L4940 𪗋 is the source for the 𪗊 component in several transcription/name characters, and so may also be the source of 𪗊 in 𪗋.



### 3.2 Folio 7 Col. 3 𪛗

The character 𪛗 occurs once in *Pearl in the Palm*, on folio 7 col. 3 of the 'B Edition' (乙種本), as the Tangut phonetic gloss for the Chinese character *shé* 蛇 'snake' (see Fig. 20). This character is missing due to damage on the corresponding page of the 'A Edition' (甲種本) (see Fig. 21).

Fig. 20: *Pearl in the Palm* 'B Edition' folio 7 col. 3

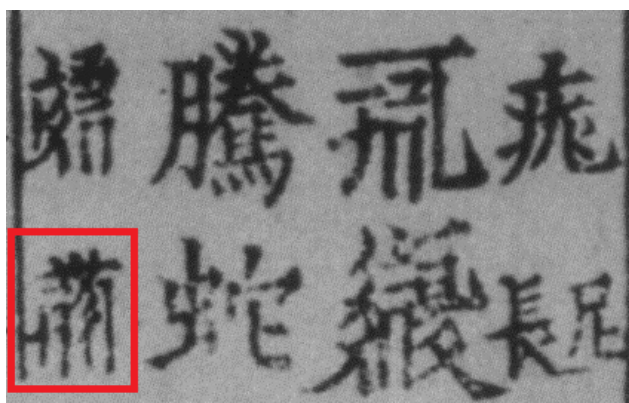
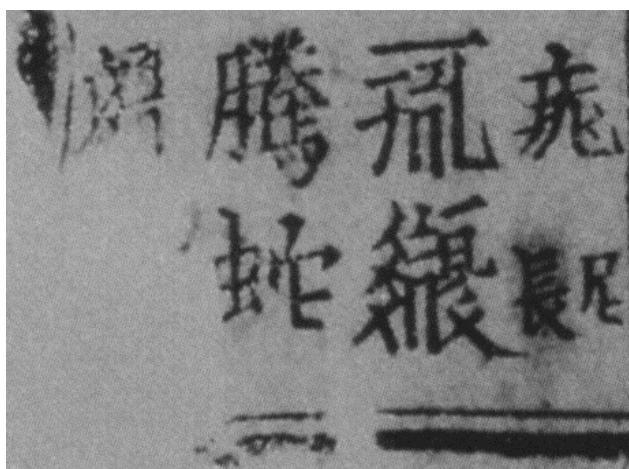
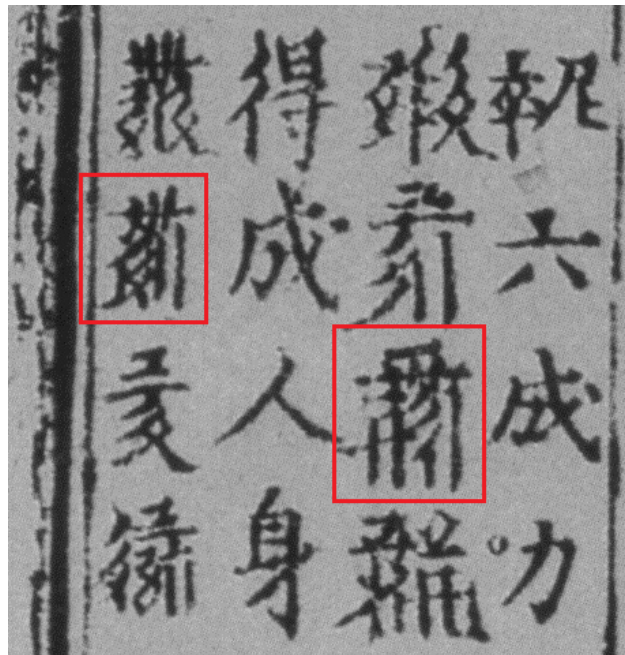


Fig. 21: *Pearl in the Palm* 'A Edition' folio 6 col. 6



This character is similar to U+17D46 𪛗 (L0113), but with 𪛗 instead of 𪛗. 𪛗 occurs as a main Tangut character on folio 19 col. 6 (see Fig. 22), where it is phonetically glossed with the Chinese character *chéng* 成; and the Chinese character *chéng* 成 in the same entry is phonetically glossed with U+176D0 𪛗 (L4156). This latter character is in turn used as the phonetic gloss for the Chinese character *shé* 蛇 on folio 16 col. 1. Therefore 𪛗 and 𪛗 could both be used to phonetically gloss *shé* 蛇. In fact, in Li Fanwen's transcriptions of 'Edition A' folio 6 he replaces the missing Tangut gloss to *shé* 蛇 with 𪛗 (see Fig. 23 and Fig. 24), although this does not accord with the occurrence of 𪛗 in the 'B Edition'.

Fig. 22: *Pearl in the Palm* folio 19 col. 6



薊 glossed as 成, and 成 glossed as 薊

Fig. 23: Li Fanwen 1994 p. 48: 066

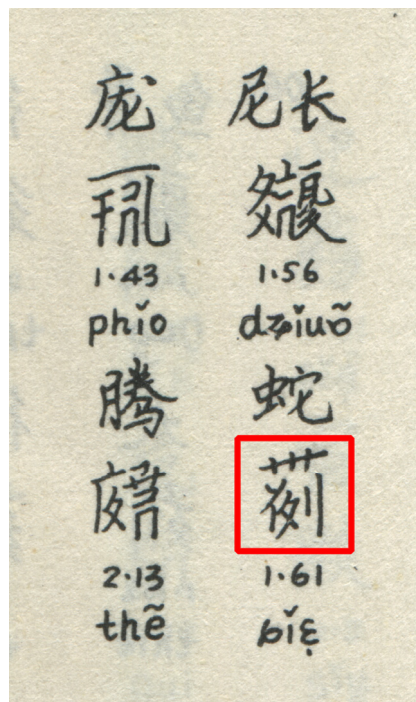
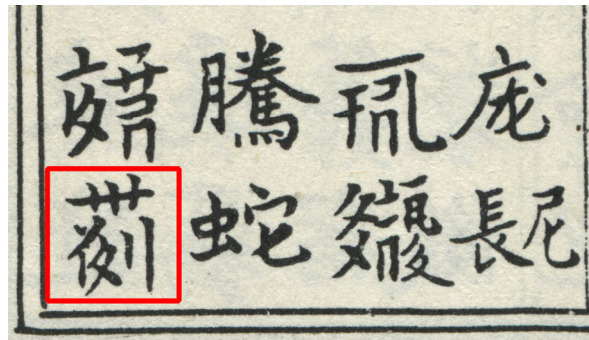


Fig. 24: Li Fanwen 1994 p. 386



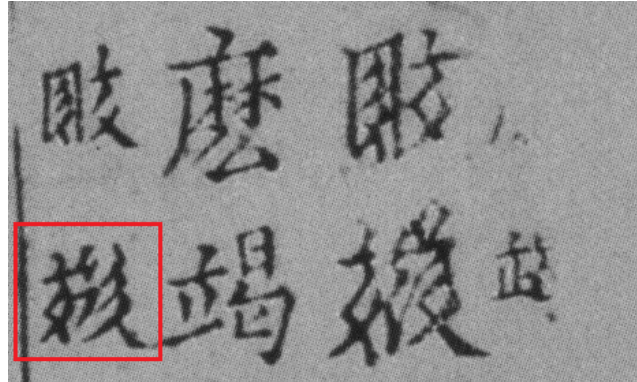
Hand copy by *Lú Tóng* 盧桐?

There are two possibilities: A) 薊 is a mistake for 薊; or B) 薊 is a special character used to phonetically gloss *shé* 蛇, constructed from the top of L4156 薊 and bottom of L0113 薊.

### 3.3 Folio 7 Col. 1 𪛗

The character 𪛗 occurs once in *Pearl in the Palm*, on folio 7 col. 1 of the 'A Edition' (甲種本), as the Tangut phonetic gloss for the Chinese character *jié* 竭 [a character here used in the transcription of the Sanskrit name of a mythical sea-creature called a *Makara*] (see Fig. 25).

Fig. 25: *Pearl in the Palm* 'B Edition' folio 7 col. 1



Li Fanwen transcribes this character as U+1803A 𪛗 (L4003) on p. 48 (see Fig. 26), but as U+18039 𪛗 (L3897) on p. 389 (see Fig. 27). He also gives 𪛗 as the correction for 𪛗 on p. 15. It is possible that Li Fanwen's interpretation of 𪛗 as a mistake for 𪛗 is correct, but it is also possible that it is a special character used deliberately as a phonetic gloss for *jié* 竭.

Fig. 26: Li Fanwen 1994 p. 48: 071

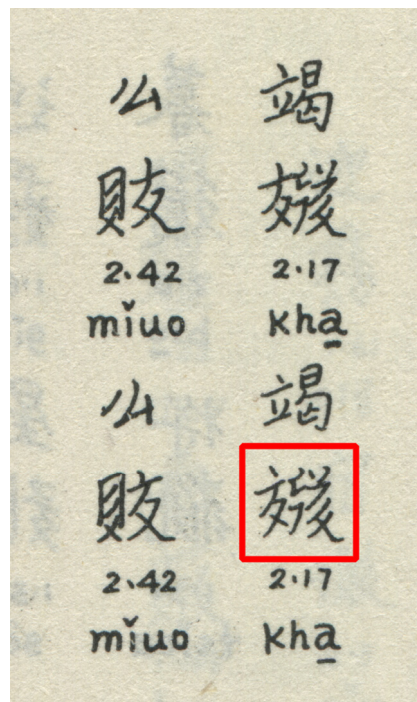
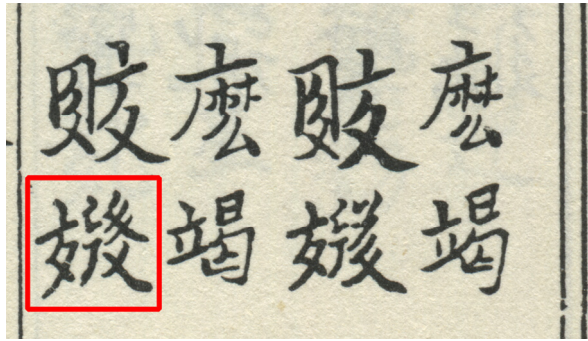


Fig. 27: Li Fanwen 1994 p. 389



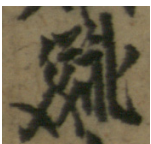
Hand copy by *Lú Tóng* 盧桐?

#### 4. Unidentified Characters in Kychanov 2010

This section is provided for information only.

In Prof. Kychanov’s study and translation of chapter 6 of the *New Laws [of the Tangut State]* 禩𑖅 (Chinese *Xīn Fǎ* 新法), published in 2010, he lists four unidentified (Russian нз. = неизвестный знак) Tangut characters (see Fig. 32, Fig. 33, Fig. 34 and Fig. 35 in the Appendix). We initially considered proposing these for encoding, but examination of the source manuscript where they occur (see Fig. 36–38 or Fig. 39–40 in the Appendix), suggests that they are not new characters at all, but are misanalysed forms of existing characters. The manuscript is written in a semi-cursive hand, and the characters are not very clear, but we believe that they should be read as shown in Table 2.

**Table 2: Unidentified Characters in Kychanov 2010**

Pressmark/ Position	Manuscript Glyph	Kychanov Reading	Proposed Reading	Meaning of Character
Tang. 56/4 04.04.06				17F6A (L2122) “city wall and moat”
Tang. 56/4 04.07.12				17F3B (L2627) “land” In this context 𑖅𑖅 means “territory”
Tang. 56/4 05.06.10				17AD1 (L2786) “livestock” In this context 𑖅𑖅 means “domesticated and wild animals”
Tang. 56/4 13.01.17				17FA9 (L2800) “county”

We do not consider it appropriate to encode misanalysed characters such as these as characters. However, that leaves the question of how these and other misanalysed character forms in modern academic sources should be represented by scholars who want to discuss works that use them. One possibility is to define standardized variation sequences, but in some cases the difference between the correct and the misanalysed form is so great that the one form cannot easily be considered to be a variant of the other form (e.g. 𑖅 vs. 𑖅 where there are no matching components).

## 5. Source References

The following entries should be appended to TangutSrc.txt (ISO/IEC 10646) and TangutSources.txt (Unicode):

U+187ED	kTGT_MergedSrc	L2012-6075
U+187ED	kRSTUnicode	195.14
U+187EE	kTGT_MergedSrc	L2012-6076
U+187EE	kRSTUnicode	308.14
U+187EF	kTGT_MergedSrc	L2012-6077
U+187EF	kRSTUnicode	415.13
U+187F0	kTGT_MergedSrc	WG2N4724-1
U+187F0	kRSTUnicode	308.17
U+187F1	kTGT_MergedSrc	WG2N4724-2
U+187F1	kRSTUnicode	415.11

L2012        Lǐ Fàn wén 2012 (Abridged Tangut-Chinese Dictionary)

There are no secondary sources for U+187F0 and U+187F1, so this document could be used as the source reference for these characters, as shown above.

## 6. Bibliography

- Huáng Zhènghuá 黄振华, Niè Hóngyīn 聂鸿音, Shǐ Jīnbō 史金波 (eds.). *Fān Hàn héshí zhǎngzhōngzhū* 番汉合时掌中珠. Yinchuan: Ningxia renmin chubanshe, 1989.
- Колоколов, В. С. [V. S. Kolokolov] and Кычанов, Е. И. [E. I. Kuchanov]. *Китайская классика в тангутском переводе* [Chinese classics in Tangut translation]. Moscow: Nauka, 1966.
- Кычанов, Е. И. [E. I. Kuchanov]. *Каталог тангутских буддийских памятников Института востоковедения Российской Академии Наук* [Catalogue of Tangut Buddhist monuments at the Institute of Oriental Studies of the Russian Academy of Sciences]. Kyoto: Kyoto University, 1999.
- Кычанов, Е. И. [E. I. Kuchanov]. “«Новые законы» тангутского государства. Глава VI. Законы и нормы поведения” [“The New Laws” of the Tangut State: Chapter 6, Laws and rules of conduct]. *Письменные памятники Востока* [Written Monuments of the Orient] 2 (13), 2010: 5–31. [http://www.orientalstudies.ru/rus/images/pdf/PPV\\_2010\\_2-13\\_01\\_kuchanov.pdf](http://www.orientalstudies.ru/rus/images/pdf/PPV_2010_2-13_01_kuchanov.pdf)
- Kwanten, Luc. *The Timely Pearl: A 12th Century Tangut-Chinese Glossary*. Vol.1. The Chinese Glosses. Bloomington, IN: Research Institute for Inner Asian Studies, Indiana University, 1982.
- Lǐ Fànwén 李范文. *Sòngdài xīběi fāngyīn: “Fān Hàn héshí zhǎngzhōngzhū” duìyīn yánjiū* 宋代西北方音——《番汉合时掌中珠》对音研究 [The Northwest Chinese Dialect at Song Times: Transliteration in Fanhan Heshe Zhangzhongzhu (ZZZ)]. Beijing: Zhongguo shehui kexue chubanshe, 1994.
- Lǐ Fànwén 李范文. *Xià-Hàn zìdiǎn* 夏漢字典 [Tangut-Chinese dictionary]. Beijing: Zhongguo shehui kexue chubanshe, 1997. ISBN 7-5004-2113-3
- Lǐ Fànwén 李范文. *Xià-Hàn zìdiǎn* 夏漢字典 [Tangut-Chinese dictionary (revised ed.)]. Beijing: Zhongguo shehui kexue chubanshe, 2008. ISBN 978-7-5004-2113-9
- Lǐ Fànwén 李范文. *Jiǎnmíng Xià-Hàn zìdiǎn* 簡明夏漢字典 [Abridged Tangut-Chinese dictionary]. Beijing: Zhongguo shehui kexue chubanshe, 2012. ISBN 978-7-5161-1544-2
- Ли Фань-вэнь [Lǐ Fànwén 李范文]. “Е.И. Кычанов и изучение документов из Хара-Хото” [E. I. Kuchanov and the study of documents from Khara-Khoto]. In Irina Porova (ed.), *Тангуты в Центральной Азии: сборник статей в честь 80-летия проф. Е.И.Кычанова* [Tanguts in Central Asia: a collection of articles marking the 80th anniversary of Prof. E. I. Kuchanov] pp. 206–212. Moscow: Oriental Literature, 2012. ISBN 978-5-02-036505-6



- Luó Fúbǎo 羅福葆. *Bǎnàběn Fān Hàn héshí zhǎngzhōngzhū* 百衲本番漢合時掌中珠. [1963?].
- Luó Fúchéng 羅福成 (ed.). *Fān Hàn héshí zhǎngzhōngzhū yìjuàn* 番漢合時掌中珠一卷. [Tianjin]: Yi'antang jingjipu [1924].
- Luó Fúchéng 羅福成 (ed.). *Fān Hàn héshí zhǎngzhōngzhū yìjuàn* 番漢合時掌中珠一卷. [Tianjin]: Yi'antang jingjipu [1935].
- Murata Jirō 村田治郎 (ed.), with contributions of Nishida Tatsuo 西田龍雄 [et al.]. *Kyoyōkan* 居庸關 [Chü-yung-kuan: the Buddhist arch of the fourteenth century A. D. at the pass of the Great Wall northwest of Peking]. Vol. I: Text. [Kyoto]: Faculty of Engineering, Kyoto University, 1957.
- Nevsky, N. A. (Н. А. Невский), *Тангутская филология: Исследования и словарь (Tangutskaja filologija: Issledovanija i slovar')* [Tangut philology: researches and dictionary]. Moscow: Oriental Literature Publishing House, 1960.
- Nishida Tatsuo 西田龍雄. *Seikago-no kenkyū: Seikago-no saikōsei-to seikamoji-no kaidoku* 西夏語の研究: 西夏語の再構成と西夏文字の解讀 [A study of the Hsi-Hsia language: reconstruction of the Hsi-Hsia language and decipherment of the Hsi-Hsia script]. Volume I. Tokyo: Zayūhō kankōkai 座右宝刊行会, 1964.
- Shǐ Jīnbō 史金波, Wèi Tóngxián 魏同賢, and E. I. Kychanov (chief eds.). *É cáng Hēishuǐchéng wénxiàn* 俄藏黑水城文獻 [Heishuicheng Manuscripts Collected in Russia]. Vol. 9. Shanghai: Shanghai guji chubanshe, 1999. ISBN 7-5325-2628-3
- Shǐ Jīnbō 史金波, Wèi Tóngxián 魏同賢, and E. I. Kychanov (chief eds.). *É cáng Hēishuǐchéng wénxiàn* 俄藏黑水城文獻 [Heishuicheng Manuscripts Collected in Russia]. Vol. 10. Shanghai: Shanghai guji chubanshe, 1999. ISBN 7-5325-2648-8
- Shǐ Jīnbō 史金波, Wèi Tóngxián 魏同賢, and E. I. Kychanov (chief eds.). *É cáng Hēishuǐchéng wénxiàn* 俄藏黑水城文獻 [Heishuicheng Manuscripts Collected in Russia]. Vol. 11. Shanghai: Shanghai guji chubanshe, 1999. ISBN 7-5325-2653-4

7. Appendix

Fig. 28: Китайская классика в тангутском переводе (1966) p. 145

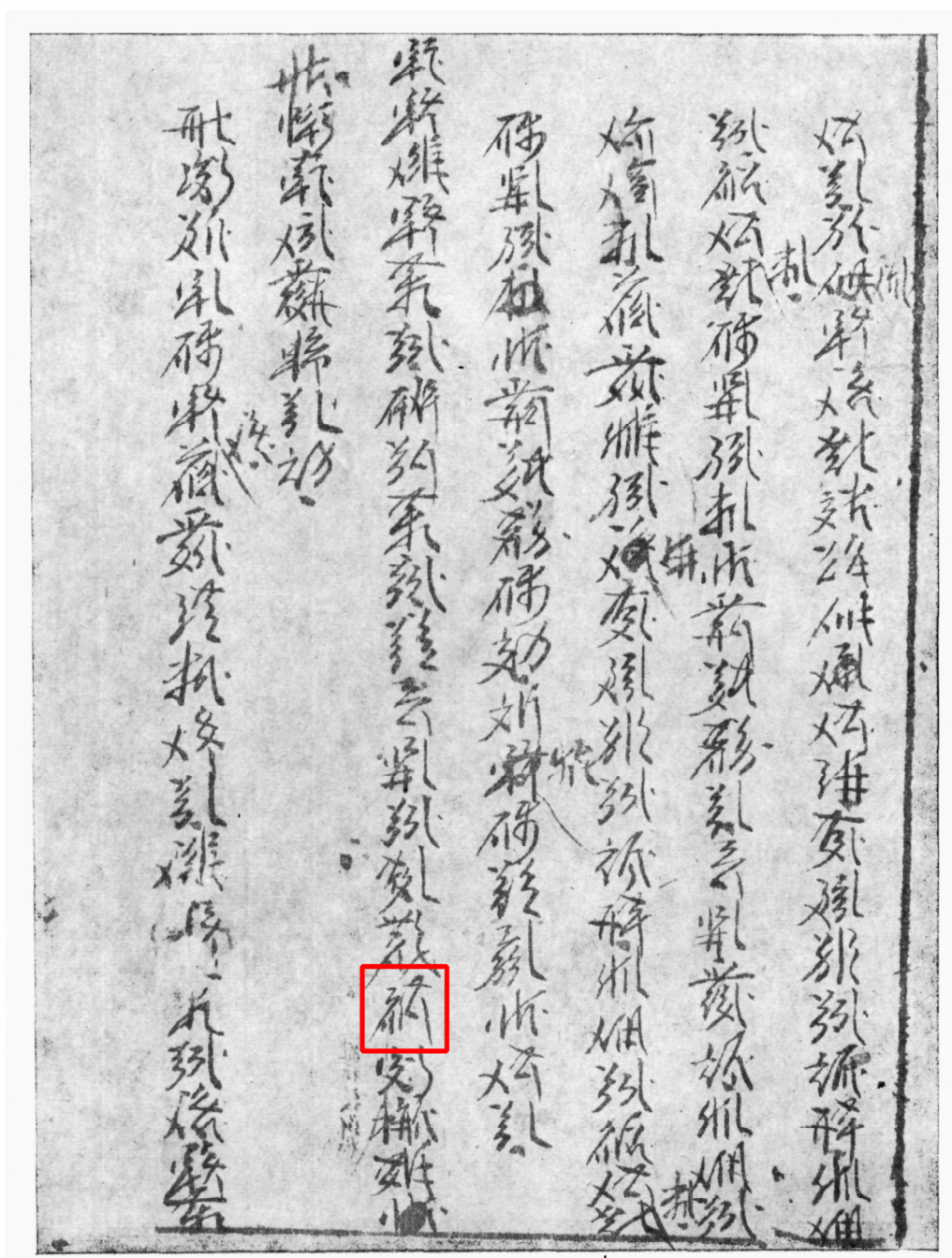


Fig. 29: Китайская классика в тангутском переводе (1966) p. 150

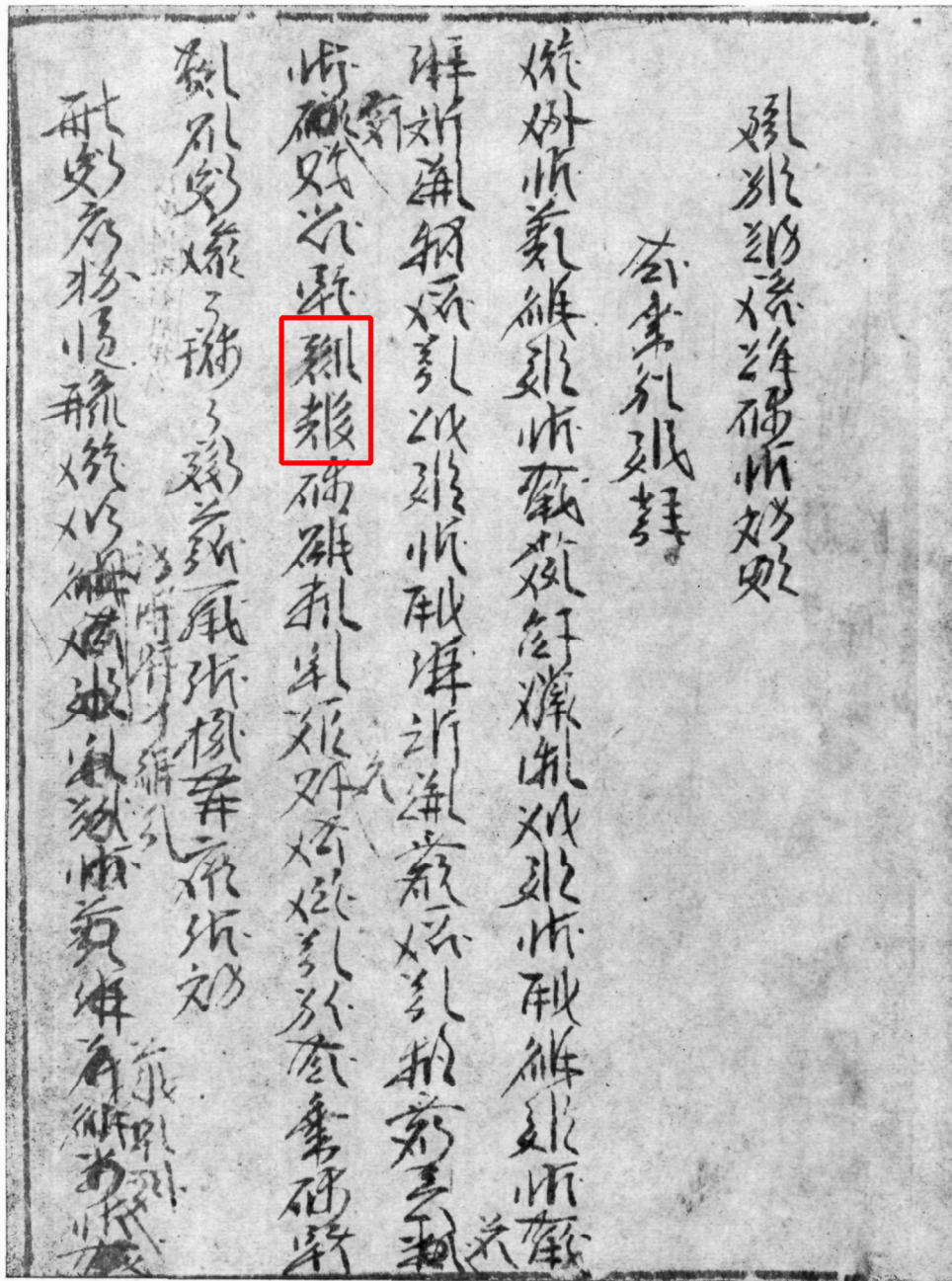
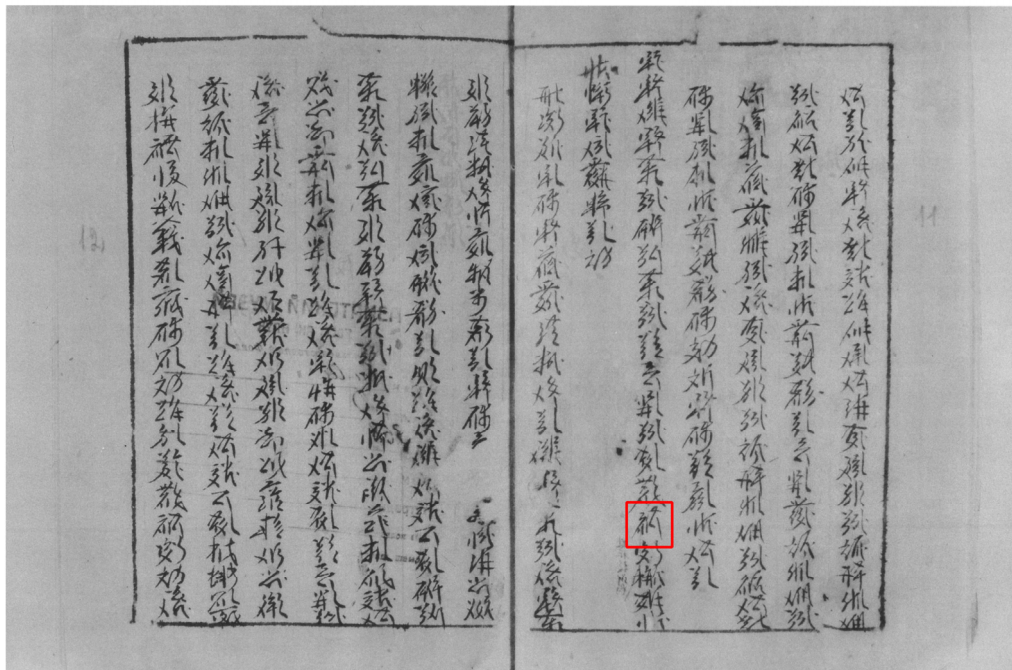


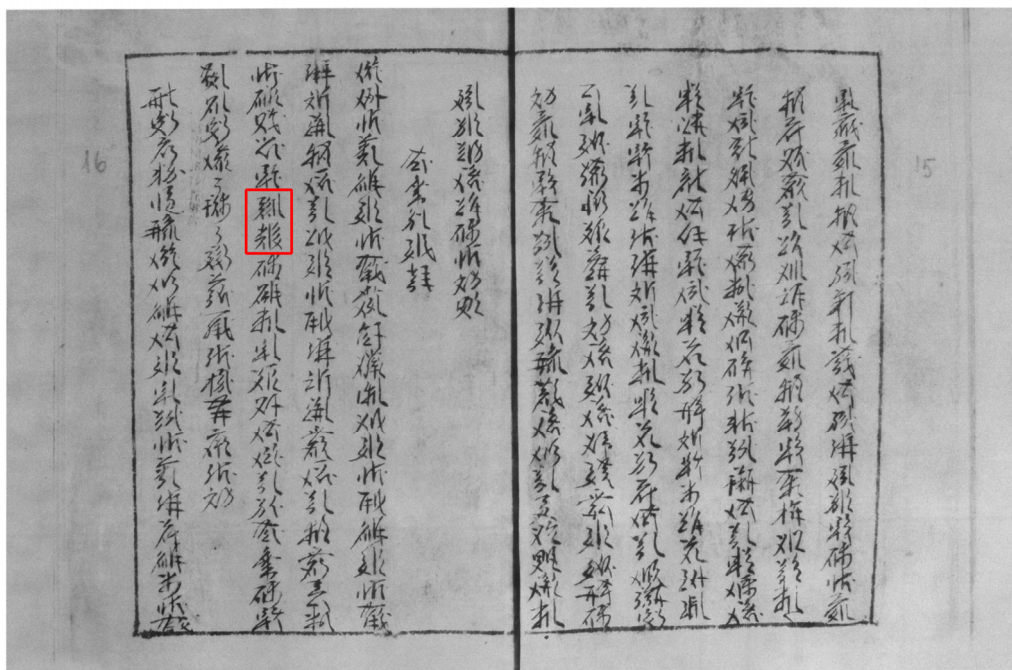
Fig. 30: Heishuicheng Manuscripts Collected in Russia vol. 11 (1999) p. 8: 45-7



俄 圖.No.2627 孝經傳 (45-7)

禱

Fig. 31: Heishuicheng Manuscripts Collected in Russia vol. 11 (1999) p. 12: 45-11



俄 圖.No.2627 孝經傳 (45-11)

禱 and 禱

**Fig. 32: Kychanov 2010 p. 15**

**[Статья 4]**

Повсюду *сину* (纒須, K1)... (нз. 1 纒), стражники-обозники (甞鞞, K2), одинаковы с *гвоздями* кандидатов на службу. Среди лиц данной категории обозникам дается скидка на 300 связок монет. Остальные, получив категорию, должны отправиться на службу в армию по действующим правилам. Если такие не получили категории, то получают полную скидку и должны служить обозниками (K3).

**Fig. 33: Kychanov 2010 p. 15**

**[Статья 5]**

Повсюду, когда поднимаются войска и производится мобилизация, то... (нз. 2 纒) расстояния до [дислокации] армии должны быть сообщены (K1). Мужчины гвона сами должны определить характер местности (K2) и скорость передвижения [мобилизованных]. И когда можно будет установить сроки явки [на место сбора], то они должны быть установлены. Если таковые просрочены, то старшие и младшие должностные лица, те, которые отправляли [мобилизованных], — командиры марша (甞鞞), командиры подразделений (甞鞞), направляющие (纒發), опоздавшие вместе с военнослужащими или допустившие такие упущения по службе, которые кому-то причинили вред, должны получить наказание как дезертиры (K3).

**Fig. 34: Kychanov 2010 p. 15**

**[Статья 6]**

Когда в государстве и за его пределами обнаруживаются необычные признаки действий врага, передвижения армий, [шаек] грабителей или мятежи, то для розыска тайно [посылается человек]. Если посланный человек в дороге, по которой он был послан, [сбил с пути? нз. 3 纒], не ушел туда, где вражеская армия, ведет себя необычно, правдивой информации нет, то это проявление трусости или желания причинить вред, достойное смерти (K1). Тот, кто предумышленно спасал свою жизнь, не достоин быть мужчиной. В том случае, если он вернулся, то по высочайше установленному закону подлежит смертной казни, а пославшее старшее или младшее должностное лицо...<sup>11</sup> Если имели место пристрастные действия и отправка [разведчика]

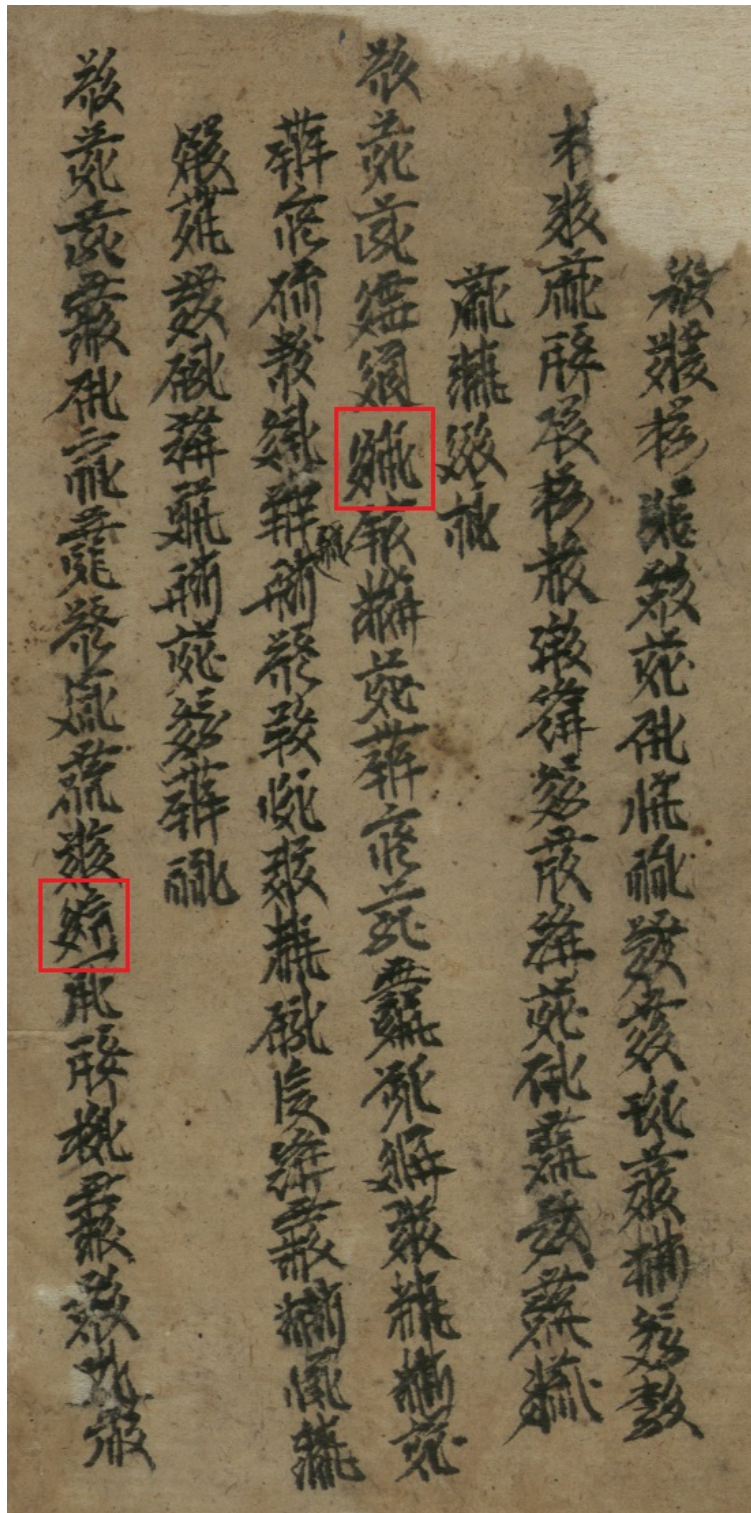
**Fig. 35: Kychanov 2010 p. 17**

**[Статья 14]**

Некто, сражаясь, лично сам проявил мужество. При обучении даются наставления другим. Прибегая к наказаниям (нз. 4 纒), следует исходить из нижеустановленного:

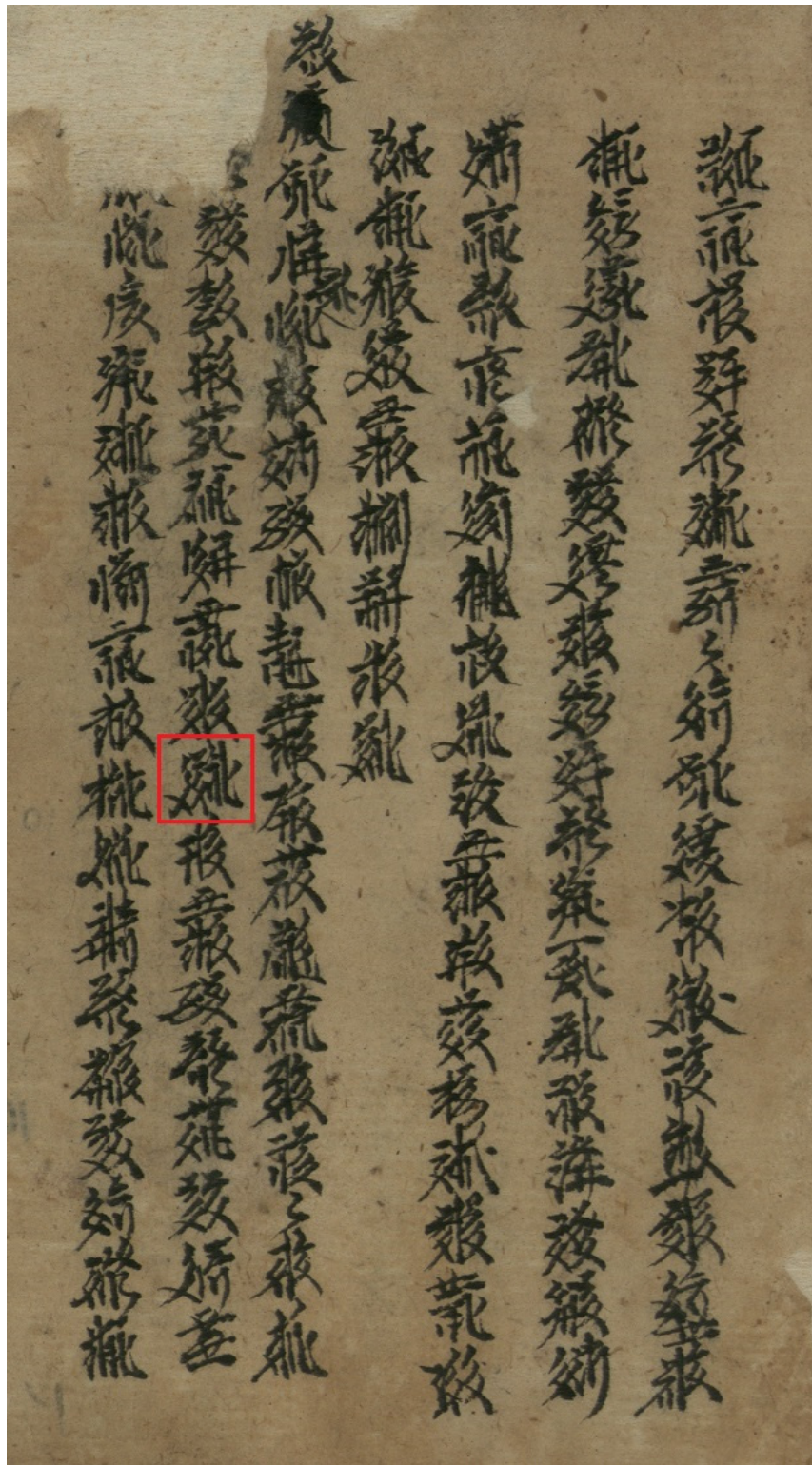
[§ 1] В сражениях у своих должны быть назначены среди пяти всадников один *ка*, над 50 всадниками один полусотник, над ста всадниками — один сотник. Когда во время сражения войска побегут, то командиры-сотники, оставившие своего командира (генерала), подлежат смертной казни. У них, вероятно, не хватило мужества (K1).

Fig. 36: Tang. 56/4, Inv. No. 247 (old pressmark Tang. 56, old Inv. No. 827) p. 4



Kychanov 𪗇 = 𪗇 and 𪗈 = 𪗈

Fig. 37: Tang. 56/4, Inv. No. 247 (old pressmark Tang. 56, old Inv. No. 827) p. 5

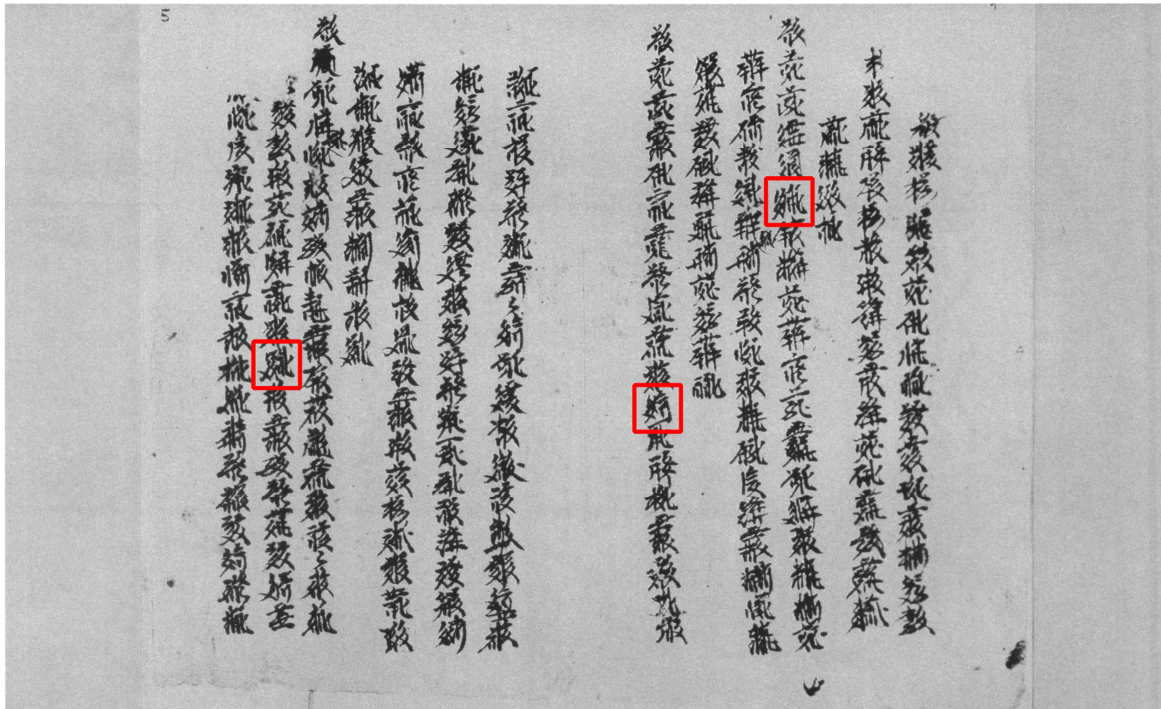


Kychanov 𪛗 = 𪛗





Fig. 39: Heishuicheng Manuscripts Collected in Russia vol. 9 (1999) p. 79: 47-4

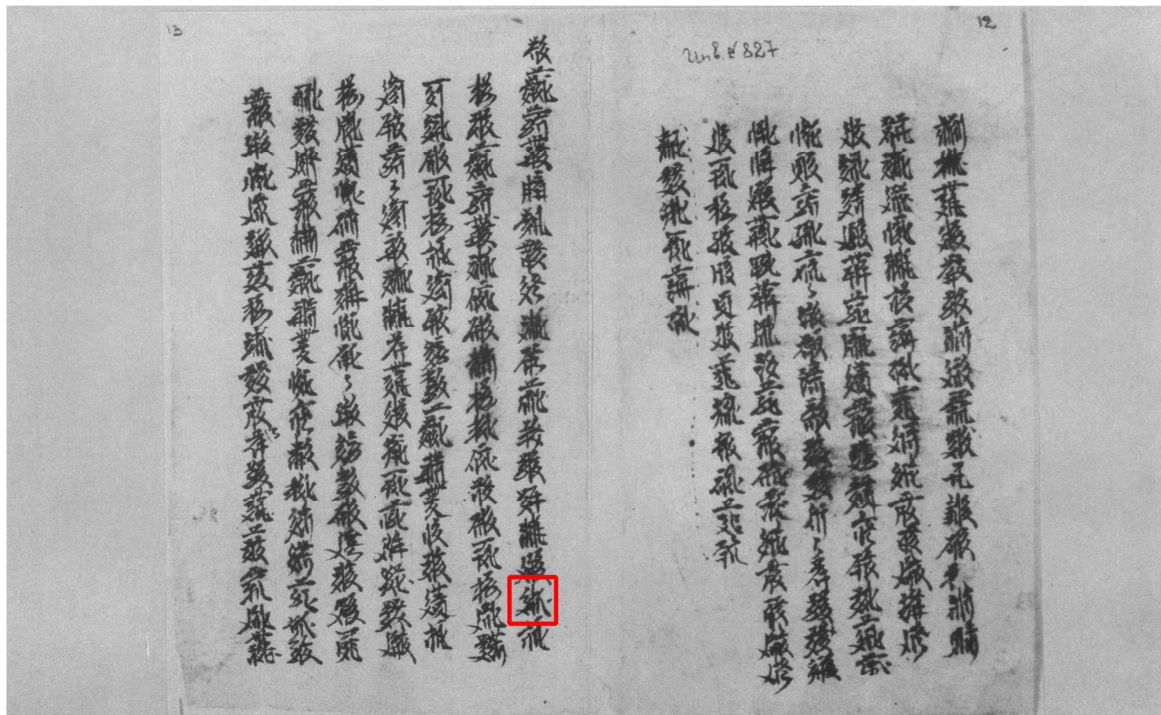


俄 Ивв.No.827 法則(乙種本)第六至第九 (47-4)

79

Kychanov 𠄎 = 𠄎, 𠄎 = 𠄎 and 𠄎 = 𠄎

Fig. 40: Heishuicheng Manuscripts Collected in Russia vol. 9 (1999) p. 81: 47-8



俄 Инв. No. 827 法則(乙種本)第六至第九 (47-8)

81

Kychanov 𪛗 = 𪛗

## 8. Proposal Summary Form

**ISO/IEC JTC 1/SC 2/WG 2  
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>1</sup>**

**Please fill all the sections A, B and C below.**

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.  
See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

### A. Administrative

1. Title: **Proposal to encode five additional Tangut ideographs**

2. Requester's name: *Andrew West, Viacheslav Zaytsev, Sun Bojun*

3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*

4. Submission date: *2016-04-21*

5. Requester's reference (if applicable):

6. Choose one of the following:

This is a complete proposal:  YES

(or) More information will be provided later:

### B. Technical – General

1. Choose one of the following:

a. This proposal is for a new script (set of characters):  NO  
Proposed name of script:

b. The proposal is for addition of character(s) to an existing block:  YES  
Name of the existing block: *Tangut*

2. Number of characters in proposal: *5*

3. Proposed category (select one from below - see section 2.2 of P&P document):

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

4. Is a repertoire including character names provided?  YES

a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?  YES

b. Are the character shapes attached in a legible form suitable for review?  YES

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?  
*Andrew West*

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):  
*Andrew West*

6. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?  YES

b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?  YES

7. Special encoding issues:

Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?  NO

### 8. Additional Information:

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

<sup>1</sup> Form number: N4102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

**C. Technical - Justification**

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