

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 six additional Tangut ideographs
Source: Andrew West and Viacheslav Zaytsev
Status: Individual Contribution
Action: For consideration by JTC1/SC2/WG2 and UTC
Date: 2017-09-07

1. Introduction

This is a proposal to encode six additional Tangut ideographs as shown in the table below.

Table 1: Proposed Characters

Code Point	Glyph	IDS	Radical/ Strokes	Source
U+187F2	𐰪	𐰪 𐰪 𐰪 𐰪 𐰪 𐰪 𐰪	42.15	Tangut manuscript
U+187F3	𐰫	𐰫 𐰫 𐰫 𐰫 𐰫 𐰫 𐰫	185.12	
U+187F4	𐰬	𐰬 𐰬 𐰬 𐰬 𐰬 𐰬 𐰬	73.11	IOM Tang. 27
U+187F5	𐰭	𐰭 𐰭 𐰭 𐰭 𐰭 𐰭 𐰭	383.18	
U+187F6	𐰮	𐰮 𐰮 𐰮 𐰮 𐰮 𐰮 𐰮	79.14	
U+187F7	𐰯	𐰯 𐰯 𐰯 𐰯 𐰯 𐰯 𐰯	79.19	

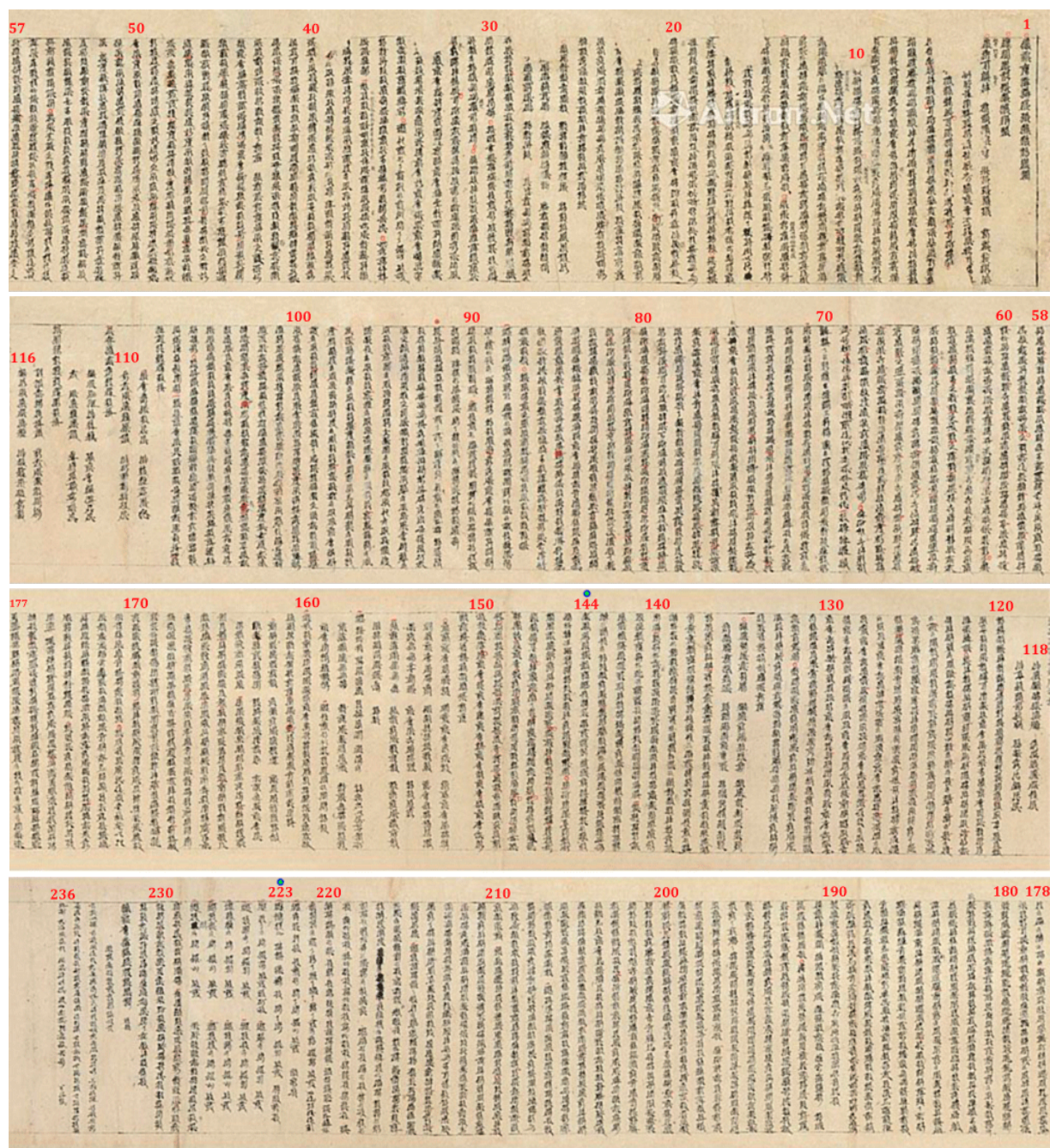
Two of these characters (U+187F2 and U+187F3) are attested in a recently discovered Tangut manuscript translation of a Tibetan Tantric Buddhist text written in 1258 by Drogön Chögyal Phagpa (1235–1280). We are grateful to Jerry You (Yóu Chéngyǔ 游程宇) for providing us with information about this manuscript.

The other four characters (U+187F4 through U+187F7) are attested in a printed Tangut collection of the writings of three generations of White Cloud Sect masters that is held at the Institute of Oriental Manuscripts [IOM] of the Russian Academy of Sciences (Saint Petersburg, Russia), pressmark Tang. 27.

2. Two New Characters in a Buddhist Manuscript Scroll

A Yuan dynasty Tangut manuscript scroll of unknown provenance, in the ownership of an anonymous private collector, was sold at auction in Beijing on 4 June 2017 for a price of RMB 2,300,000 (approximately USD 338,000), about three times the estimate of RMB 750,000–850,000 (approximately USD 110,000–125,000).¹

Fig. 1: Lot 2368 of the June 2017 Beijing Auction



Note: We have added column numbers for convenience.

¹ Lot 2368 of the Spring 2017 Art Auction held by the Beijing Taihejiacheng Auction Company. See <http://www.thjc.cn/web/auctionShow/viewAuctionItem?auctionItemId=86972&fromPage> for details.

The manuscript is 27 × 312 cm in dimension, and consists of 236 columns of 23–26 characters per column.

The text of this manuscript is entitled *ngiwei* ᠨᠭᠢᠪᠡᠢ ᠨᠢᠨ ᠮᠢᠣ ᠵᠤ ᠲᠰᠢᠨ ᠫᠤᠨ ᠯᠳᠢᠭ ᠰᠢᠭᠠᠩ ᠰᠢᠭᠠᠨ ᠰᠢᠭᠠᠨ ᠰᠢᠭᠠᠨ (title translated into Chinese by modern scholars as *Xǐ jīngāng xiànzhèng rúyì bǎo* 喜金剛現證如意寶) “Hevajra’s Wish-Fulfilling Gem of Abhisamaya”, and it has been identified as a previously unknown Tangut translation of the first part of a Tibetan text relating to the Buddhist deity Hevajra (*dpal kye rdo rje'i mngon rtogs yid bzhin nor bu* དཔལ་ལྷོ་རྡོ་རྗེ་འཕྱོ་རྟོགས་ཡིད་བཞིན་ནོར་བུ་) which was made in 1258 by Drogön Chögyal Phagpa (1235–1280), fifth patriarch of the Sakya school of Tibetan Buddhism and Imperial Preceptor under Kublai Khan.²

This text has been studied by Li Ruoyu 李若愚 in an academic article (September 2016) and in his PhD dissertation (April 2017). Li has identified two new characters (𐰇 and 𐰈) used in this text as an abbreviation for the Sanskrit word *vajra*.

The character 𐰇 occurs twice in a mantra on column 223 (Fig. 2, Fig. 3, Fig. 5 and Fig. 7), and the character 𐰈 occurs a single time in a mantra on column 144 (Fig. 4 and Fig. 8). In two out of the three cases the character is used in an abbreviation for the name Hevajra. Note that in the title of the Tangut text the name Hevajra is translated as 𐰇𐰈 (Chinese *Xǐ jīngāng* 喜金剛), but within mantras given in the text the name Hevajra is transcribed phonetically as 𐰇𐰈 (xa mba ndzi rīa) or using the abbreviations 𐰇 or 𐰈.

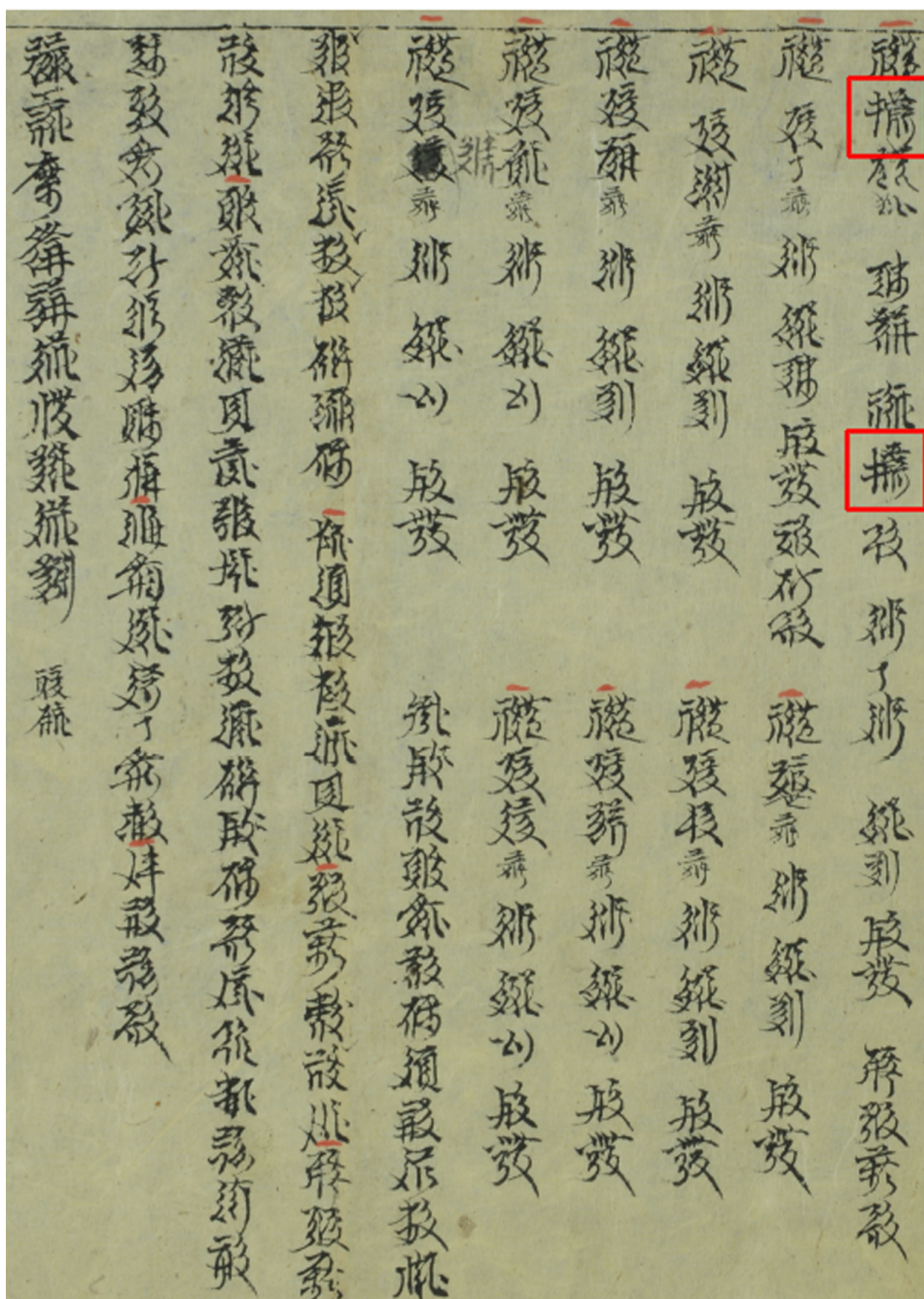
The character 𐰇 is constructed from the left side of 𐰇 (mba = va-), the top and left of 𐰈 (ndzi = -j-), and the left side of 𐰈 (rīa = -ra), the three parts together transcribing *vajra* (see Fig. 6 for Li Ruoyu’s explanation). This character is very similar in construction to the character 𐰈 which was proposed for encoding at U+187F0 in WG2 N4724. 𐰈 also stands for *vajra*, and has the same right side as 𐰇, but the left side is derived from the left side of 𐰈 (mba). That is to say, 𐰈 and 𐰇 are two different ways of representing *vajra* as a portmanteau abbreviation, but with different homophonic source characters for the left side element representing the va- part of *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 the three characters 𐰇𐰈𐰈 (see Fig. 13) and four times using the three characters 𐰇𐰈𐰈 (see Fig. 14), thus indicating that 𐰈 and 𐰇 are equivalent abbreviations for *vajra*.

According to Li Ruoyu (see Fig. 10), the character 𐰈 is an idiosyncratic creation for the word *vajra*, composed from the top and left of 𐰈 (ndzi = -j-), and the left side of 𐰈 (rīa = -ra). However, without the left side component 𐰇 or 𐰈 representing the va- of *vajra* this character construction is incomplete. Although it is possible that this is a deliberate alternative form for *vajra*, as it only occurs once it is more likely that 𐰈 is a scribal error for 𐰇. Even though 𐰈 may be an error, we propose encoding it as it is required by scholars studying this important Buddhist text.

² Tibetan text published in Drogön Chögyal Phagpa 2007 pp. 413–451. See Fig. 11 and Fig. 12.

Fig. 2: Columns 223–232 of the original manuscript showing 櫛

232 231 230 229 228 227 226 225 224 223



Column 223: *om vajrakartari hevajrāya hūṃ hūṃ hūṃ phaṭ svāhā*

祿擗覯 𢇛𢇛 祿擗𢇛 𢇛𢇛 𢇛𢇛 𢇛𢇛 𢇛𢇛

Fig. 3: Close-up of Fig. 2

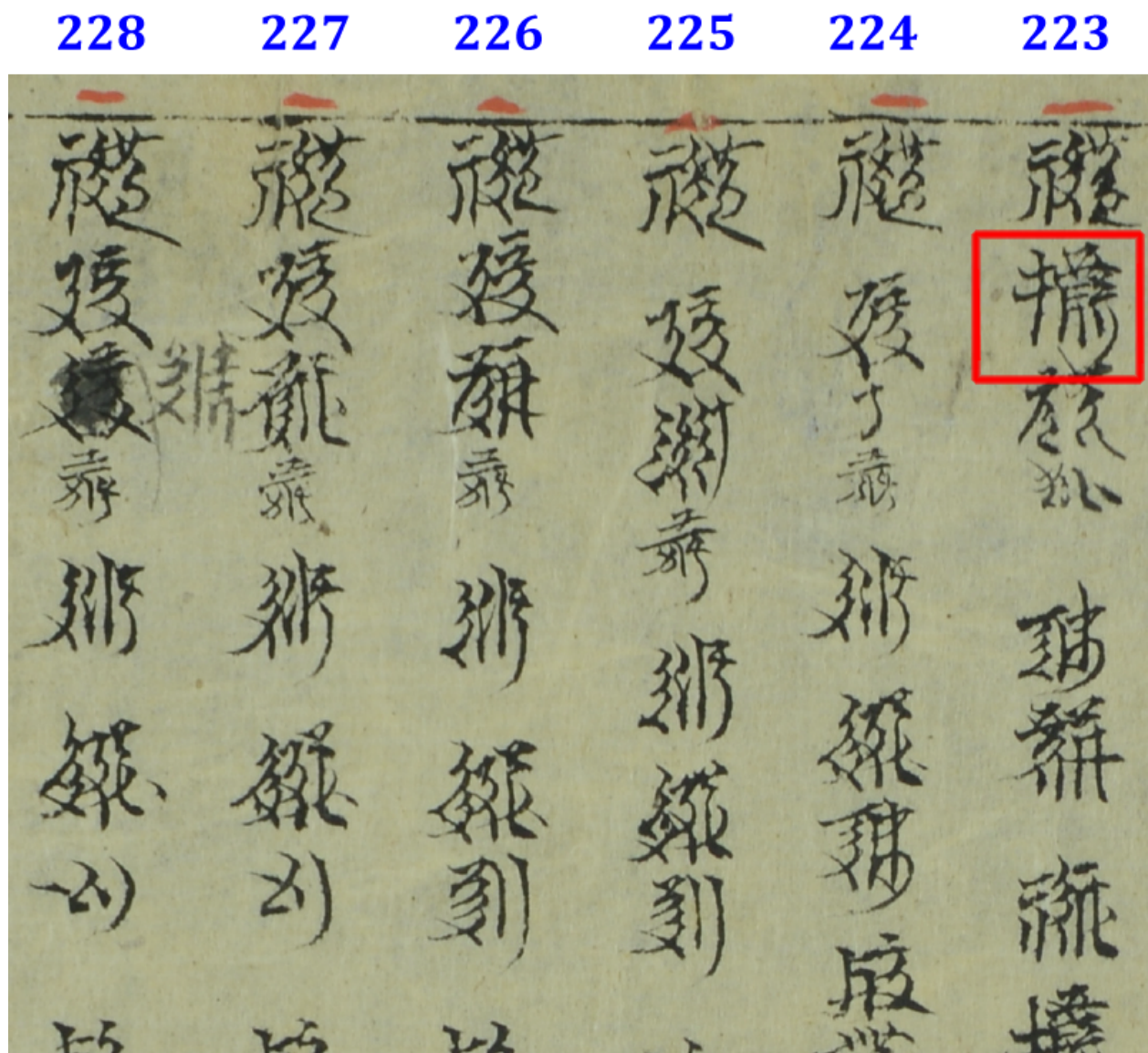
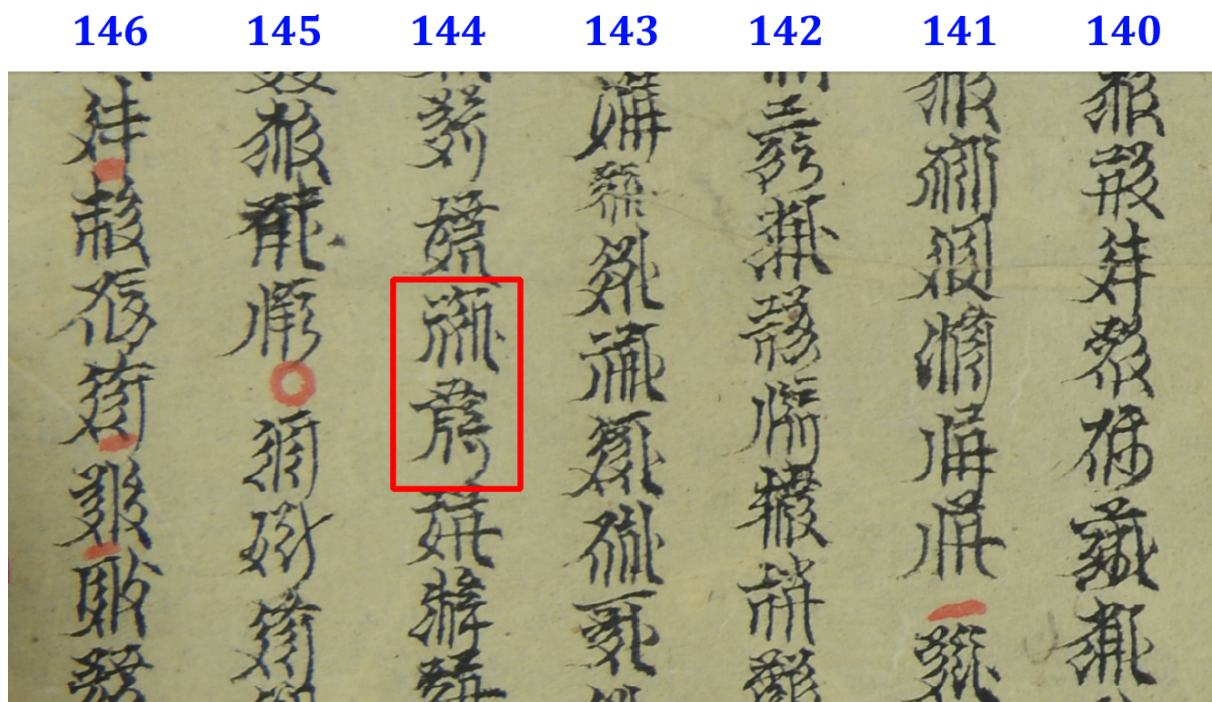


Fig. 4: Columns 140–146 of the original manuscript showing 祇禱



Column 144: *om bhagavām mahākāru[ni]ka hevajra saparivāra ...*

祿接甌薺煥叢甌薺[薺]甌甌薺薺薺薺薺薺薺薺...

Note that all the manuscript examples write the component 扌 in 搨 and 搨 as 扌, with an extra horizontal stroke, but this is an obvious scribal error which we do not follow in the proposed code chart glyphs.

卷尾部分








[222] 𢇛 𢇜 𢇝 𢇞 𢇟 𢇠 𢇡 𢇢 𢇣 𢇤 𢇥 𢇦 𢇧 𢇨 𢇩 𢇪
𢇫。𢇬 𢇭 𢇮。

[illegible][illegible]





Page 7

校注：

(1) 心咒 (𐰇𐰺𐰍), 西夏字面义为“心真”, 对应藏文 snying po(精华, 藏, 要旨), 佛经常译作“心咒”。

(2) vajra, 西夏文作 。该字不见于西夏字书, 系译经人临时自造的合体俗字, 实为    (vajra) 三字的省写, 以  字的左半部加上  字的上半部和下半部的左边, 另外加上  字的左半边构成。

(3) 根本心咒 (𐰇𐰺𐰍), 西夏字面义为“近心真”, 对应藏文 nye ba'i snying po, 在佛经中常译作“根本心咒, 本尊心咒”。

Li Ruoyu's explanation that the unencoded character  is an abbreviation for    (vajra) and that it is constructed from elements of those three characters.

**Fig. 9: Li Ruoyu 2017 p. 74: Transcription of Tibetan text, and
Free translation into Chinese (for Col. 144)**

baṃ zhes brjod pas rang la thim/ de nas/ sma rī dril bu 'dzin pa mang po spros/ dril bu'i sgra
thams cad la khyab par bsam zhing/ dril bu'i phyag rgya byas la/ om bha ga vāṃ ma hā kā ru ṇi ka
he vajra sa pa ri vā ra ghasma rī va śaṃ ku ru hoḥ zhes brjod pas ye zhes pa rnam rang la mnyes
bzhin du bzhugs par bsam mo//

思自心中射出^[1]多个 gaurī 母，皆持铁钩^[2]，以钩触诸智尊^[3]心，挽于近旁^[4]，身结铁钩
印，语^[5] “ye hye hi bha ga vān ma hā kā ru ṇi ka he vajra sa pa ri vā ra”，于后二处^[6]亦合，诵
咒 “gau rī ā karṣa ya ja”，使诸智尊至于己处。次思射出多个红色 caurī 母，皆持罽索^[7]，以
罽索触诸智尊颈，挽于近旁，以诵咒 “cau rī pra be śa ya hūṃ”，至于胜妙殿。次思射出多个
黄色 vetālī 母，皆持铁链^[8]，触诸智尊之足，入于己身，结铁链印，以诵咒 “ve tā lī pandha baṃ”
入于己身。复次思射出多个 ghasmarī 母，皆持铃^[9]，铃声遍及一切，结铃印，诵咒 “om bha
ga vāṃ ma hā kā ru ka he vajra^[10] sa pa ri vā ra ghasma va śaṃ ku ru ho”，使智尊欢居于己。

注释：

Fig. 10: Li Ruoyu 2017 p. 75 Note 10 for p. 74 (see Fig. 9 above)

[10] vajra, 西夏文作 “𐵇𐵓”。该字不见于西夏字书，系译经人临时自造的合体俗字，实
为 “𐵇𐵓” (jra) 二字的省写，以 “𐵇” 字的上半部和下半部的左边，另外加上 “𐵓” 字的
左半边构成。

Fig. 11: Drogön Chögyal Phagpa 2007 p. 432: Tibetan text (for Col. 223)

ཕའ་སྐྱ་རྒྱ། ཞེས་པ་ནི་ཙུ་བའི་སྐྱགས་སོ། །ཨོ་དེ་མ་པི་ཙུ་བའི་ཧྱུ་ཧྱུ་ཧྱུ་ཕའ་སྐྱ་རྒྱ།
 སྒྲིང་བོའོ། །ཨོ་བཟླ་ཀར་རི་རེ་བཟླ་ཡ་ཧྱུ་ཧྱུ་ཧྱུ་ཕའ་སྐྱ་རྒྱ། ཉེ་བའི་སྒྲིང་བོའོ། །ཨོ་
 ཡུམ་(《ཡུ》+ཨ་ཡུམ་ ཨེ་ཨྱ། ཨྱ་ཨྱ། རི་ཨྱ། ཨེ་ཨྱ། ཨེ་ཨྱ། ཨོ་ཧྱུ་ཕའ་སྐྱ་རྒྱ། ཡུམ་གྱི་ཙུ་བའི་སྐྱགས་
 སོ།།)ཨོ་ཧྱུ་ཕའ་སྐྱ་རྒྱ། ཡུམ་ : གྱིའོ(《ཡུ》གྱི་སྒྲིང་བོའོ)། །ཨོ་ཡུམ་གྱི་ཧྱུ་ཕའ་སྐྱ་རྒྱ།
 ཨོ་ཡུམ་ཙུ་ཧྱུ་ཕའ་སྐྱ་རྒྱ། ཨོ་ཡུམ་བུ་ཧྱུ་ཕའ་སྐྱ་རྒྱ། ཨོ་ཡུམ་གྱི་ཧྱུ་ཕའ་སྐྱ་རྒྱ། ཨོ་ཡུམ་བུ་
 ཧྱུ་ཕའ་སྐྱ་རྒྱ། ཨོ་ཡུམ་ཤི་ཧྱུ་ཕའ་སྐྱ་རྒྱ། ཨོ་ཡུམ་ལོ་ཧྱུ་ཕའ་སྐྱ་རྒྱ། ཨོ་ཡུམ་རྩི་ཧྱུ་ཕའ་སྐྱ་
 རྒྱ། ལའོར་གྱི་ཧྱུ་མོ་རྣམས་གྱི་སྐྱགས་སོ། །ཚིག་གཞན་གྱིས་བར་མ་ཆད་བར་

Sanskrit transliteration
for Tangut (Col. 223):

om vajrakartari hevajrāya hūṃ hūṃ hūṃ phaṭ svāhā

Tibetan:

oM badz+ra karta ri he badz+rA ya hU~M hU~M hU~M phaT swA hA/

ཨོ་བཟླ་ཀར་རི་རེ་བཟླ་ཡ་ཧྱུ་ཧྱུ་ཧྱུ་ཕའ་སྐྱ་རྒྱ།

Fig. 12: Drogön Chögyal Phagpa 2007 p. 425: Tibetan text (for Col. 144)

བར་བསམ་ཞིང་། ལྷགས་སྒྲོག་གི་ཕྱག་རྒྱ་བྱས་ལ། བེ་དྲ་ལྷེ་བརྒྱ་བེ་ཞེས་བརྗོད་
 བས་རང་ལ་ཐིམ། དེ་ནས་། སྐྱ་རྩེ་དེའི་ཁུ་འཛིན་པ་མང་བོ་སྒྲོས། དེའི་ཁུ་ལྷོ་
 ཐམས་ཅད་ལ་ཁུ་བར་བསམ་ཞིང་། དེའི་ཁུ་ལྷོ་ཕྱག་རྒྱ་བྱས་ལ། ཨོ་རྩ་ག་ལྷོ་
 (《ཞ》ལྷོ)མ་དྲ་ཀླ་རུ་ཅི་ཀ་རེ་བཟླ་ས་པ་རི་ལྷ་ར་ལྷོ་རྩེ་མ་ཤི་ཀླ་རུ་རྩེ་ཞེས་
 བརྗོད་པས་ཡེ་ཤེས་པ་རྣམས་རང་ལ་མཉེས་བཞིན་དུ་བཞུགས་པར་བསམ་མོ། །

Sanskrit transliteration
for Tangut (Col. 144):

om bhagavāṃ mahākāru[ni]ka hevajra saparivāra ...

Tibetan:

oM b+ha ga wAM* ma hA kA ru Ni ka he badz+ra sa pa ri wA ra ...

ཨོ་རྩ་ག་ལྷོ་མ་དྲ་ཀླ་རུ་ཅི་ཀ་རེ་བཟླ་ས་པ་རི་ལྷ་ར་

* Zha lu Monastery edition: wAn ལྷོ་ instead of wAM ལྷོ་, i. e. b+ha ga wAn ལྷོ་ག་ལྷོ་ = Sanskrit bhagavān

Fig. 13: Juyong Pass East Wall cols. 8–12: rubbing and Nishida's transcription
 (Murata Jirō (ed.) *Kyoyōkan* 居庸關 (1957) rub. IV & p. 182)



𑖀 𑖀 𑖀 𑖀 𑖀 𑖀
 *le bha *ra *ga — *ya
 vajra-kāya-

𑖀 𑖀 𑖀 = *vajra*

Fig. 14: Juyong Pass East Wall cols. 17–21: rubbing and Nishida's transcription
(Murata Jirō (ed.) *Kyoyōkan* 居庸關 (1957) rub. IV & p. 183)



Column 18:

□ 𑖀𑖩 𑖀𑖩 𑖀𑖩 𑖀𑖩 𑖀𑖩
*ga *ri bhe *m^wa *ʔdzi *ra
ya- garbhe vajra-

Column 19:

𑖀𑖩 𑖀𑖩 𑖀𑖩 𑖀𑖩 𑖀𑖩 𑖀𑖩 𑖀𑖩
*ʔdzi *ri *ni *m^wa *ʔdzi *ra *mi
jrini, vajram

𑖀𑖩 𑖀𑖩 = *vajra*

3. Four New Characters in IOM Tang. 27

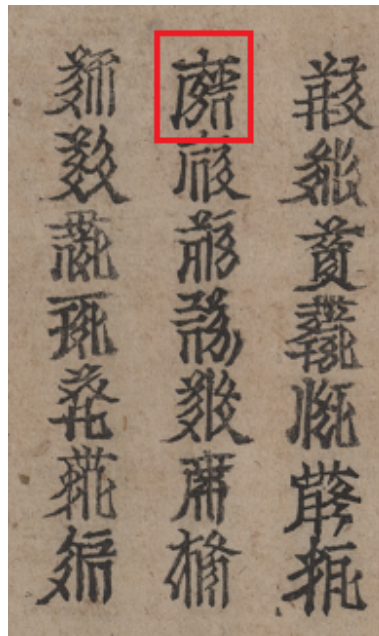
Tang. 27 held at the Institute of Oriental Manuscripts of the Russian Academy of Sciences (Saint Petersburg, Russia) is a Tangut text printed using moveable type (Tangut *ndi siwu* 佩繼 “living characters”) during the Western Xia period. It comprises a collection of the writings of three generations of White Cloud Sect masters entitled *sə ʃiei ndźwɪ swew ngwu ʃiou ɿwə* 散胤彙彙祔祔祔 (title translated into Chinese by modern scholars as *Sānshìshǔ míngyán jíwén* 三世屬明言集文 or *Sāndài xiāngzhào yán wénjí* 三代相照言文集) “Collected Writings of the Shining Speech of Three Generations”.

This is an important text for the study of Buddhism under the Western Xia, and has been studied by Prof. Arakawa Shintarō (2001), Prof. Evgeny Ivanovich Kychanov (2004), Prof. Sun Bojun (2011), and other scholars. Sun Bojun 孫伯君 recently pointed out to us that there are four unencoded Tangut characters in the text, which she requests should be encoded. These four characters all have a unique graphic construction, and cannot convincingly be taken to be corrupt or mistaken forms for any existing encoded Tangut character.

3.1 U+187F4 𐰪𐰸

The character 𐰪𐰸 occurs once, on folio 2A col. 5 (see Fig. 15). We do not know of any published study that discusses it.

Fig. 15: Tang. 27 folio 2A cols. 4–6: 𐰪𐰸

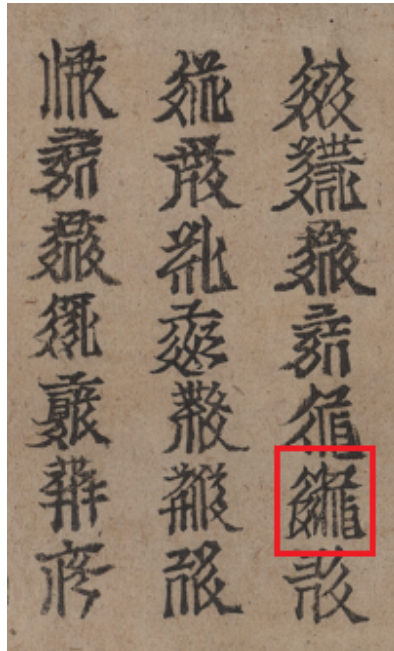


The character bottom middle and right components occur as the bottom part of *pha* 𐰪𐰸 “conceited”. The character occurs before *tsja* 𐰪𐰸 “to give as a present”. It is possible that 𐰪𐰸 is a mistake for *tsja* 𐰪𐰸 “presently, in a moment”.

3.2 U+187F5 𪛵

The character 𪛵 occurs once, on folio 22A col. 3 (see Fig. 16). We do not know of any published study that discusses it.

Fig. 16: Tang. 27 folio 22A cols. 3-5: 𪛵



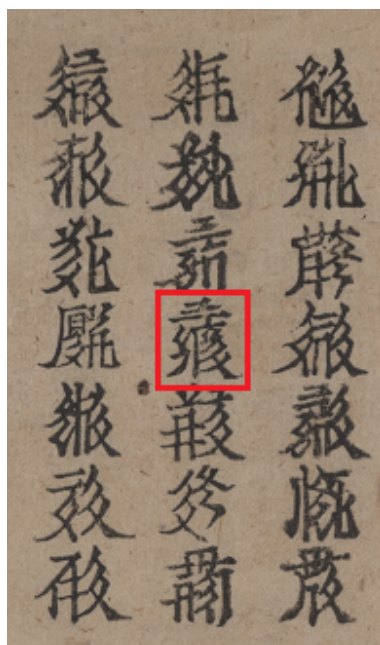
The character has the same left and right components as *po* 𪛵 “name of a ghost”. The middle and right also occur as the character *tsha* 𪛵 “ghost”. As 𪛵 occurs immediately after *·ju* 𪛵 “ghost”, it would seem to be a character meaning some sort of ghost.

3.3 U+187F6 𪛶

The character 𪛶 occurs once, on folio 29B col. 3 (see Fig. 17). We do not know of any published study that discusses it.

The character is similar in construction to character *tshie* 𪛶 “to pull, to tug”, but with the component 彡 in place of 扌. The bottom part also occurs as the character *nai* 𪛶 “lubrication”. As 𪛶 occurs immediately after *·m* 𪛶 “self”, it is possible that the character is constructed from the top and bottom right of *tshie* 𪛶 “to pull, to tug”, and the bottom left of *·m* 𪛶 “self”. Cf. the phrase 𪛶𪛶 “pulling at oneself” on folio 36A quoted in section 3.4 below. It is even possible that 𪛶 is a mistake for *tshie* 𪛶 “to pull, to tug”.

Fig. 17: Tang. 27 folio 29B cols. 2–4: 𪛗



3.4 U+187F7 𪛗

The character 𪛗 occurs twenty times on folios 36A through 38A, ten times in a poem entitled *kai tsə tʃia lə ndjie* 𪛗𪛗𪛗𪛗𪛗 “Pain of desire and *rūpa* [material objects] hindering the way” (see Fig. 18), and ten times in a poem entitled *mje ngje ndziwo lie siwə* 𪛗𪛗𪛗𪛗𪛗 “Worrying that fame and fortune destroys people” (see Fig. 19). See Table 2 for locations and images of all occurrences of the character.

Each poem consists of ten stanzas of three seven-character lines covering fifteen columns. Both poems are transcribed and translated in Arakawa 2001 (see Fig. 20 and Fig. 22). In both poems the unknown character occurs in a refrain which comprises the third line of each stanza. The refrain is slightly different in each poem, but both refrains start with the unknown character 𪛗 followed by *ndzie* 𪛗 “to teach”.

The reading of the character 𪛗 is unknown, and its meaning is uncertain. The left side of this character does not occur as a component of any other character, but is the whole of the character *tshje* 𪛗 “to pull, to tug”. The right side component 𪛗 occurs in characters relating to dried meat, cheese, and verbs relating to moisture, so it is not obvious what it represents semantically (it is possible that it is a phonetic component).

Arakawa relates 𪛗 to the character *tshje* 𪛗 “to pull, to tug” as it occurs on the line before the first occurrence of 𪛗, and he conjecturally translates the unknown character into Japanese as *indō* 引導. In Modern Japanese this word usually refers to the last words addressed to a newly-deceased person by the officiating priest at a Buddhist funeral, but we suppose that here it is used in the original Chinese sense of “to guide, to lead” or “guidance”. Arakawa renders the two characters 𪛗𪛗 as “teaching of guidance” (see Fig. 20, Fig. 21 and Fig. 22).

On the first occurrence of the unknown character in each poem it is followed by two small-sized Tangut characters, which would seem to be a gloss for the reader on this unfamiliar character: *tshêu lâ* 𐰇𐰺𐰍, literally “to damage” and “drunkenness”, which we take to mean “drunken to the point of incapacity” (“blind drunk” in colloquial English). This gloss does not seem to have any bearing on the character 𐰇 if it is interpreted as meaning “guidance”, so it is possible that other interpretations should be considered.

As the character *ndzie* 𐰇𐰺 “to teach” can also mean “master, teacher” (as a homophonic alternative for *ndzie* 𐰇𐰺 “master, teacher”), we suggest an alternative possible interpretation of the two characters 𐰇𐰺 as a sobriquet for the subject of the poem, the Drunken Master. Our tentative translation of the first stanza of each poem is given below.

First stanza of poem entitled *kai tsə tšja lə ndje* 𐰇𐰺𐰍𐰺𐰍𐰺𐰍 “Pain of desire and *rūpa* [material objects] hindering the way” (cf. Arakawa’s Japanese translation shown in Fig. 20):

𐰇𐰺𐰍𐰺𐰍𐰺𐰍𐰺𐰍
𐰇𐰺𐰍𐰺𐰍𐰺𐰍𐰺𐰍
𐰇𐰺𐰍𐰺𐰍𐰺𐰍𐰺𐰍

Grieving for somebody who renounces the world and follows the way,
At the boundary pulling at oneself like a mad dog,
Drunken Master’s hard work, what does he hope to achieve?

First stanza of poem entitled *mje ngje ndziwo lje siwə* 𐰇𐰺𐰍𐰺𐰍𐰺𐰍 “Worrying that fame and fortune destroys people” (cf. Arakawa’s Japanese translation shown in Fig. 22):

𐰇𐰺𐰍𐰺𐰍𐰺𐰍𐰺𐰍
𐰇𐰺𐰍𐰺𐰍𐰺𐰍𐰺𐰍
𐰇𐰺𐰍𐰺𐰍𐰺𐰍𐰺𐰍

Fame and fortune echo as sweet as other [things],
Emptiness and no fortune encircling me,
Drunken Master’s grief and woe, what does he hope to achieve?

Fig. 18: Tang. 27 folio 36A
cols. 1-3: 𪛗

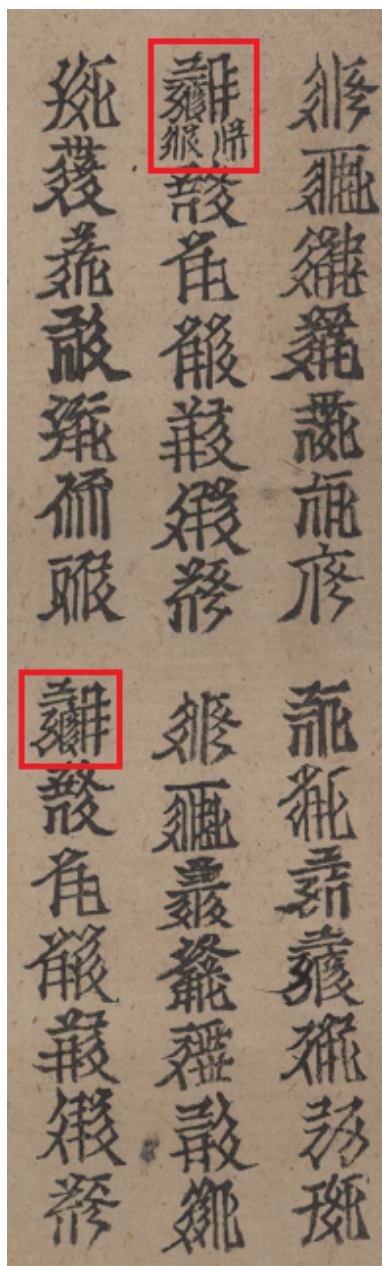


Fig. 19: Tang. 27 folio 37A
cols. 3-5: 𪛗

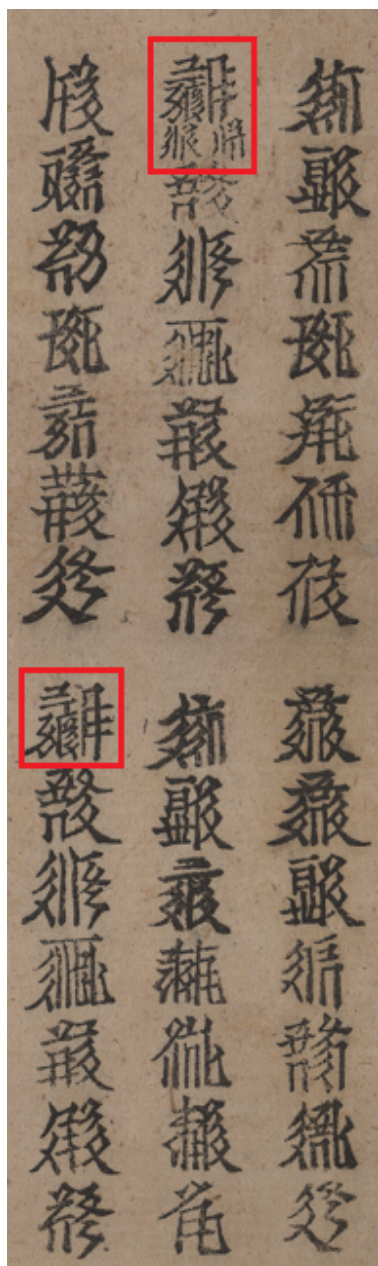
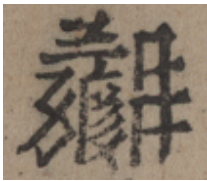



Table 2: All occurrences of 辯 on folios 36 through 38

Poem 1		Poem 2	
36A02			37A04
36A03			37A05
36A05			37A07
36A06			37B01
36B01			37B03
36B02			37B04
36B04			37B06

Poem 1		Poem 2	
36B05			37B07
36B07			38A02
37A01			38A03

Although the character 翻 occurs twenty times in the text, because it was printed using moveable type, some of the type sorts for this character were reused for printing on two or more folios. This can be clearly seen in the cases of defective characters resulting from damage to the type sort, such as the two characters at 36A02 and 37A04 which share the same broken top horizontal stroke, and the three characters at 36B07, 37B07 and 38A02 which share the same broken middle horizontal stroke. In the case of the characters at 36A02 and 37A04 it seems that the two following small characters 併 繼 were cut on the same type sort as the character 翻.

Fig. 20: Arakawa 2001 page 218: Transcription and translation of folio 36a cols. 1-2

36a-1	<p> ²ngo ²de:' 1ni: 1phl 1ca: 1jenq <u>2myeq'2</u> 悲痛なるかな、出家し^{註70} 道を行う者 </p>	<p> ²myl' 2'a 1'e: 1chi:' 1khwl: 2'aq <u>2syu</u> 境に自らを引き出し、狂犬^{註71} の如し </p>
36a-2	<p> ^{???}^{註72} ²dze:' 2ryeq'2 2gyu 2wa 2ryeq'2 <u>1wi:</u> 引導の^{註73} [毀れ酔う^{註74}] 教え、その労苦は何を願うのか </p>	<p> ²ngo ²de:' 1chwo" 1denq 2si: 1jyu <u>?lhc?</u> 悲痛なるかな、或いは（同じ）部族 の女をだまし取る（ほどに） </p>

Fig. 21: Arakawa 2001 page 218 footnotes 70-73

- ^{註70} 西夏語の語順・意味では「家を一捨てる」となっている。
^{註71} 西夏語の語順では「犬一狂った」と形容詞の後置修飾である。
^{註72} 声調・声母・韻母不明であるので推定音は ??? で表記せざるを得なかった。
^{註73} この字形は各種西夏語字書『文海』、『同音』などに確認できない。意味は前の行（36a-1）に現れる字形 𐽳 「引き出す」から推定した。
^{註74} 字形が確認しづらいが 𐽳 「酔う」（『夏漢』3547, p. 660）と判読できる。

Fig. 22: Arakawa 2001 page 220: Transcription and translation of folio 37a cols. 2-4

「名利人滅憂（名利が人を滅ぼすことを憂う）」

37a-2	<p> ²me:' 1genq 2dzwo: 2le:' 1syl' 名利が人を滅ぼすことを憂う </p>	
37a-3	<p> ²me:' 1genq 1zyi 2syu 1tha: 1'e: <u>2lyeq</u> 名利は響く如く、その甘いこと </p>	<p> 1nga 1nga 1genq 1me: 2pon 2ror <u>2de:</u> 空空^{註85} であり、利はなく、自ら廻っている </p>
37a-4	<p> ^{???} ²dze:' 2ngo ²de:' 2wa 2ryeq'2 <u>1wi:</u> 引導の[毀れ酔う^{註86}] 教え、その悲痛は何を願うのか </p>	<p> ²me:' 1genq 1dzu 1byu 1nwl 1li:q' <u>2ryeq'2</u> 名利は楽しみに随い、（あとで）知る、労苦を^{註87} </p>

4. *Source References*

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

U+187F2	kTGT_MergedSrc	WG2 N4851-1
U+187F2	kRSTUnicode 42.15	
U+187F3	kTGT_MergedSrc	WG2 N4851-2
U+187F3	kRSTUnicode 185.12	
U+187F4	kTGT_MergedSrc	WG2 N4851-3
U+187F4	kRSTUnicode 73.11	
U+187F5	kTGT_MergedSrc	WG2 N4851-4
U+187F5	kRSTUnicode 383.18	
U+187F6	kTGT_MergedSrc	WG2 N4851-5
U+187F6	kRSTUnicode 79.14	
U+187F7	kTGT_MergedSrc	WG2 N4851-6
U+187F7	kRSTUnicode 79.19	

We suggest using this document as the source reference for these characters, as shown above.

5. Bibliography

- N4724.** Andrew West, Viacheslav Zaytsev, Sun Bojun. *Proposal to encode five additional Tangut ideographs*. ISO/IEC JTC1/SC2/WG2 N4724. 2016-04-21. <http://www.unicode.org/L2/L2016/16095-n4724-tangut.pdf>
- Arakawa 2001.** Arakawa Shintarō 荒川慎太郎. “Seikashi no kyakuin ni mirareru inbo ni tsuite: ‘Sanze zoku meigen shū bun’ shoshū Seikagoshi” 西夏詩の脚韻に見られる韻母について: 『三世属明言集文』 所収西夏語詩 [About the rhymes in Tangut verses: Reanalysis of Tangut rhyming poetry in *San shi shu ming yan ji wen*]; *Kyōto daigaku gengogaku kenkyū* 京都大学言語学研究 = *Kyoto University Linguistic Research* vol. 20 (2001): 195–224.
- Drogön Chögyal Phagpa 2007.** 'gro mgon chos rgyal 'phags pa = Bāsibā 八思巴 [Drogön Chögyal Phagpa]. *sa skya bka' 'bum dpe bsdur ma las 'gro mgon chos rgyal 'phags pa'i gsung 'bum pod dang po bzhugs so* [Comparative Edition of the Collected Writings of Sakya: Collected Works of Drogön Chögyal Phagpa. Vol. 1] = *sa skya gong ma rnam lnga'i gsung 'bum dpe bsdur ma las 'gro mgon chos rgyal 'phags pa'i gsung pod dang po bzhugs* [Comparative Edition of the Collected Works of the Five Founding Fathers of Sakya: Works of Drogön Chögyal Phagpa. Vol. 1] = *Sàjiā wǔzǔ quánjí duìkān běn* 萨迦五祖全集对勘本 (17/25) [Comparative Edition of the Collected Works of the Five Founding Fathers of Sakya. Vol. 17/25]. Pe cin = Běijīng 北京: Krung go'i bod rig pa dpe skrun khang = Zhōngguó zàngxué chūbǎnshè 中国藏学出版社, 2007. (Mes po'i shul bzhag = Xiānzhé yíshū 先哲遗书; 19). ISBN 978-7-80057-846-5
- Kychanov 2004.** Кычанов, Е. И. (E. I. Kychanov). “Несколько предварительных замечаний по поводу тангутского текста «Собрание слов, передаваемых от одного к другому в трёх поколениях»” [Some Preliminary Remarks about the Tangut Text *The Collection of Words Transmitted through Three Generations from One Person to Another*]; *Письменные памятники Востока* = *Written Monuments of the Orient* no. 1 (2004, Spring–Summer): 147–159.
- Li Ruoyu 2016.** Lǐ Ruòyú 李若愚. “‘Xǐ jīngāng xiànzhèng rúyì bǎo’: Yuán dìshī Bāsibā zhùzuò de Xīxià yìběn” 《喜金刚现证如意宝》: 元帝师八思巴著作的西夏译本 [Hevajra's Ruyibao: A Tangut Version by Yuan Imperial Preceptor Ba Siba]; *Níngxià Shèhuì Kēxué* 宁夏社会科学 [Social Sciences in Ningxia] No. 5 (September 2016): 206–212.
- Li Ruoyu 2017.** Lǐ Ruòyú 李若愚. *Xīxiàwén ‘Xǐ jīngāng xiànzhèng rúyì bǎo’ kǎoshì* 西夏文《喜金刚现证如意宝》考释 [Study of the Tangut version of *Hevajra's Wish-Fulfilling Gem of Abhisamaya*]. Ph.D. dissertation. Běijīng 北京: Zhōngguó shèhuì kēxuéyuàn yánjiūshēng-yuàn 中国社会科学院研究生院 = Graduate School of Chinese Academy of Social Sciences, 2017. Advisor, Prof. Niè Hóngyīn 聂鸿音.

- Murata 1957.** 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.
- Sun Bojun 2011.** Sūn Bójūn 孙伯君. “Yuándài báiyún zōng yìkān Xīxiàwén wénxiàn zōngkǎo” 元代白云宗译刊西夏文文献综考 [General study of Tangut documents translated and printed during the Yuan dynasty by the White Cloud sect]; *Wénxiàn* 文献 [Documents] No. 2 (2011): 146–157.

6. *Tangut Sources*

Hevajra's Abhisamaya. *ngiwei 𐰇𐰺 𐰍𐰆 𐰍𐰏𐰤 𐰇𐰤 tsin 𐰇𐰏𐰤𐰠𐰏𐰤𐰠𐰏𐰤* [Hevajra's Wish-Fulfilling Gem of Abhisamaya [direct realization] = *Xǐ jīngāng xiànzhèng rúyì bǎo* 喜金剛現證如意寶]. Sold at the Spring 2017 Art Auction held by the Beijing Taihejiacheng Auction Company (Běijīng tài hé jiā chéng pāimài yǒuxiàn gōngsī 北京泰和嘉成拍卖有限公司) on 4 June 2017 (Lot 2368). Provenance and current owner unknown.

Three Generations. *sə 𐰇𐰏𐰤 ndźwɪ swew ngwu 𐰇𐰏𐰤 𐰇𐰏𐰤 𐰇𐰏𐰤 𐰇𐰏𐰤 𐰇𐰏𐰤 𐰇𐰏𐰤* [Collected Writings of the Shining Speech of Three Generations = *Sānshìshǔ míngyán jíwén* 三世屬明言集文 or *Collection of Words Transmitted From Person to Person through Three Generations* = *Sāndài xiāngzhào yán wénjí* 三代相照言文集]. Institute of Oriental Manuscripts of the Russian Academy of Sciences (Saint Petersburg, Russia). Tangut fond. Pressmark Tang 27, old inventory № 4166.

7. 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 10646³ 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 six additional Tangut ideographs
2. Requester's name:	Andrew West and Viacheslav Zaytsev
3. Requester type (Member body/Liaison/Individual contribution):	Individual contribution
4. Submission date:	2017-09-07
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:	6
3. Proposed category (select one from below - see section 2.2 of P&P document):	A-Contemporary <input type="checkbox"/> B.1-Specialized (small collection) <input type="checkbox"/> B.2-Specialized (large collection) <input type="checkbox"/> C-Major extinct <input type="checkbox"/> D-Attested extinct <input type="checkbox"/> E-Minor extinct <input type="checkbox"/> F-Archaic Hieroglyphic or Ideographic <input checked="" type="checkbox"/> G-Obscure or questionable usage symbols <input type="checkbox"/>
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.):	Jing Yongshi, Andrew West, Michael Everson
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.

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