DOI: 10.47979/aror.j.89.3.443-482

Nasal Preinitials in Tangut Phonology

Xun Gong

Abstract

Gong Hwang-cherng proposed that the Tangut language has a distinction between short and long vowels. To date, however, no reliable correlates have been found regarding the actual phonological nature of the distinction. A careful examination of Chinese loanwords in Tangut and Sino-Tangut pronunciation reveals that the "vowel length" distinction should be revised to that of the presence vs. absence of a nasal preinitial. The pair [illegible]$_{3806}$ "weed" vs. [illegible]$_{2138}$ "tomb," borrowed respectively from Chinese 蒲 *bu* and 墓 *muH* (the latter from a Northwest-type reflex with **mb-*), hitherto reconstructed as $bu^{ʁ1}$ {bu^1} vs. $buu^{ʁ2}$ {buu^2}, should be revised to $bu^{ʁ1}$ vs. $mbu^{ʁ2}$. The reconstructed nasal preinitial not only has a close typological parallel in Modern West Rgyalrongic, but is equally reflected in other sources of evidence, most strikingly Sanskrit transcription and *fǎnqiè*. The revision solves a large number of problems in the historical phonology of Tangut, though not without raising some new ones, especially in connection with the treatment of Proto-Rgyalrongic preinitials before nasals.

Keywords

Linguistics | Tangut | Northwestern Medieval Chinese | Qiangic languages

* *Xun Gong (minus273cn@gmail.com) is currently an Assistant at Institut für Sprachwissenschaft, Universität Wien.*

** I wish to thank the anonymous reviewers for their useful comments. This essay is dedicated to the memory of Pān "the sea slug" Léi 潘雷 (1988–2020), a fellow lover of Jerzy Petersburski and Kajio Shinji.

1. CONTEXT

1.1 Paired Columns Of Rhymes And The "Vowel Length" Distinction

Gong Hwang-cherng 龔煌城, in his 1994 article[1], proposed that Tangut[2] has a distinction between short and long vowels. Given that Tangut is written in a non-phonetic writing system, in which, much as in Chinese, a character denotes a syllable-morpheme whose phonetic nature must be deduced through a conjunction of more or less tangential evidence, a hypothesis about Tangut phonology such as the one at hand should be properly understood as consisting of two sub-hypotheses:

- *Categorization and phonemicity*: A partition of the Tangut characters into disjoint categories, such that syllables denoted by characters in one category share some common phonological feature distinguishing them from syllables in other categories.
- *Phonological substance*: An identification of the concrete phonological distinction that underlies the difference between the categories.

In order to properly understand the vowel length hypothesis, we start by examining Gong Hwang-cherng's reasoning and other potential evidence from the perspectives of both categorization and phonological substance.

In terms of categorization, Gong Hwang-cherng's hypothesis is rooted in

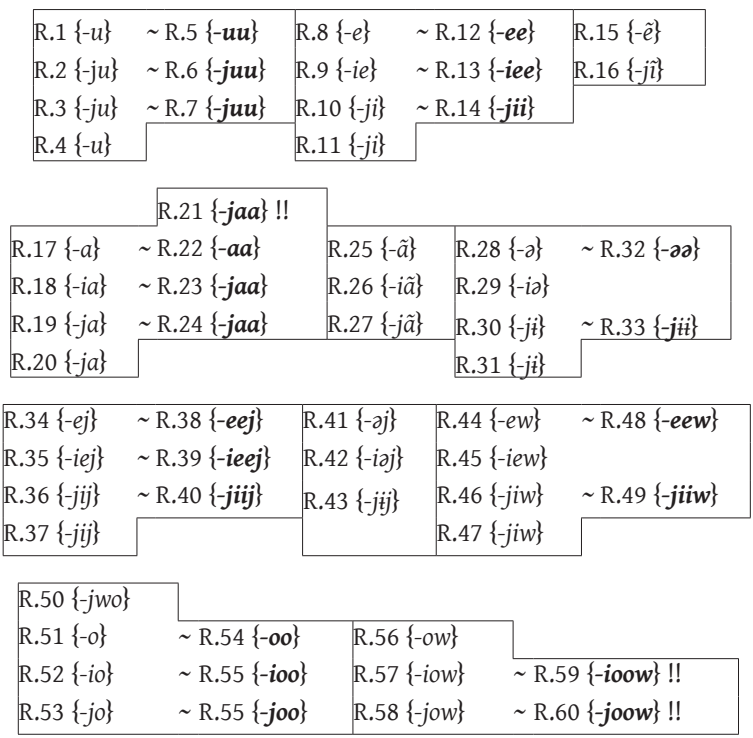
---

[1] Hwang-cherng Gong, "A Hypothesis of Three Grades and Vowel Length Distinction in Tangut," *Journal of Asian and African Studies* 46–47 (1994): 305–14.

[2] Tangut characters are annotated with their Lǐ number, referring to the numbering system of the second edition of the *Tangut-Chinese Dictionary* (2008). Transcriptions given inside curly braces conform to Gong Hwang-cherng's reconstruction. Outside curly braces, transcriptions always take into account the uvularization hypothesis —see Xun Gong, "Uvulars and uvularization in Tangut phonology," *Language and Linguistics* 21, no. 2 (2020): 175–212—and, unless otherwise clear from the context, also the nasal preinitial hypothesis proposed in this essay. A fully annotated example is 𗥃$_{1892}$ *mmi*^{1} {*mjii*^{1}} "house."
In this study, I annotate Chinese syllables in Early Middle Chinese transcribed in a slightly modified version of the system used in William H. Baxter, *A Handbook of Old Chinese Phonology.* Berlin: Mouton de Gruyter, 1992. Notably, *o* is changed to *ʌ*, and medial *-j-* to *-i-*. Other sounds are transcribed in an IPA-like fashion: among Baxter's alternative orthographies, *æ*, *ɛ*, *ɨ* are preferred to *ae*, *ea*, *ɨ*. Retroflex stops are written *ʈ*, *ʈh*, *ɖ*, *ɳ*, retroflex sibilants *ʈʂ*, *ʈʂh*, *dʐ*, *ʂ*, *ʐ* and palatal sibilants *tɕ*, *tɕh*, *dʑ*, *ɲ*, *ɕ*, *ʑ*. Similarly, I use *j* for *y*, *ŋ* for *ng*, *ʔ* for ', and *ɣ* for *h*. The transcription is explicitly of an indicative nature. Only the initials should be understood as participating in the arguments. Chinese initials are also annotated in their customary representation as single Chinese characters, *p-* 幫, etc., in order to facilitate reading for those accustomed to other reconstructions.

his discovery of the "paired sequences" of rhymes.[3] Consider, for example, the group of rhymes from R.8 to R.14. Foreign-language transcription evidence shows that they share, more or less, the same main vowel. In Tangutological parlance, they belong to the same 攝 *shè*. Moreover, as the same 1994 article demonstrates, there is another dimension within the same *shè*—namely that of grades (等 *děng*), represented in Gong Hwang-cherng's system as {*-e*}: {*-ie*} : {*-ji*}. The "paired sequences" refer to the observation that exactly the same series of rhymes, ordered by grade, seems to exist twice, juxtaposed to each other. Immediately after R.8 {*-e*}, R.9 {*-ie*}, and R.10/11 {*-ji*} come R.12 {*-ee*}, R.13 {*-iee*}, and R.14 {*-jii*}, which are, in almost every respect of foreign-language transcription, virtually equivalent to their respective counterparts in R.8–R.11.

Figure 1: Paired columns of rhymes in the major cycle of Tangut rhymes (R.1–R.60)



Note: Column-2 rhymes in boldface; "!!" marks rhymes assigned to column 2 in Gong Hwang-cherng's system, as reflected in Gong (2003) and all versions of the *Tangut-Chinese Dictionary* (Lǐ Fànwén 1997, 2008, 2012), but considered in this essay to be column-1 rhymes.

[3] Gong Hwang-cherng, *Xīxià yǔwén yánjiù lùnwénjí* [Collected Papers on Tangut Philology] (Taipei: Academia Sinica, 2002), 147.

As Figure 1 shows, this pattern of paired sequences is repeated over and over in the native rhyme ordering system. In this essay, I refer to them as *paired columns*, enumerated as *column 1* and *column 2*. The question of phonemicity naturally comes into play. How do we know, apart from rhyme ordering in native metalinguistic resources, if the distinction between R.1–R.4 and R.5–R.7 is the same as the distinction between R.8–R.11 and R.12–R.14? Gong Hwang-cherng typically addresses such questions through the method he labels *phonological alternation*[4], which he did not apply to paired columns. Nevertheless, applying his method to this phenomenon indeed shows a tight connection. For example, [illegible]$_{0716}$ {*śjii*^{1}} is the stem A of a verb "to butcher" whose stem B is [illegible]$_{4571}$ {*śjoo*^{1}}. The stem A is in rhyme R.14, belonging to the second column; the stem B, in rhyme R.55, also belongs to the second column. On the other hand, an alternating verb, such as "to eat," whose stem A [illegible]$_{4517}$ {*dzji*^{1}} belongs to R.10, a rhyme in the first column, has a stem B [illegible]$_{4547}$ {*dzjo*^{1}} with the rhyme R.53 in the first column too. This shows that the distinction between R.8–R.11 : R.12–R.14 is indeed analogous to the distinction between R.51–R.53 : R.54–R.55, thereby suggesting that they reflect a basic phonemic distinction of the language.

Concerning the phonological substance, Gong Hwang-cherng reconstructs the distinction between column 1 and column 2 as one of vowel length. Column-1 rhymes are reconstructed with short vowels, column-2 with the long vowels. Hence, [illegible]$_{3806}$ "weed," having the column-1 rhyme R.1 (1.1), is reconstructed as *bu*$^{\text{ʁ}1}$ {*bu*^{1}}, whereas [illegible]$_{2138}$ "tomb," which belongs to the column-2 rhyme R.5 (1.5–2.5), is reconstructed as *buu*$^{\text{ʁ}2}$ {*buu*^{2}}.

Gong Hwang-cherng's vowel length hypothesis relies on one single observation: "rhymes representing Chinese loanwords [i.e., rhymes with nasalized vowels] have only a short vowel sequence and no corresponding long vowel sequence."[5] In this statement, he is referring to the rhyme sequences R.15–R.16 and R.25–R.27, reconstructed in most reconstruction systems with the nasal vowels {*ĩ* / *ẽ*} and {*ã*}. As can be seen in Figure 1, these rhyme sequences are not divided into paired columns. Since these rhyme sequences mostly involved loanwords from Chinese, Gong Hwang-cherng considers that the lack of paired

4 Gong Hwang-cherng, "Phonological Alternations in Tangut," *Bulletin of the Institute of History and Philology* 59, no. 3 (1988): 783–834; Gong, "The Phonological Reconstruction of Tangut Through Examination of Phonological Alternations," *Bulletin of the Institute of History and Philology* 60, no. 1 (1989): 1–45; Gong, "Xīxiàyǔ de yīnyùn zhuǎnhuàn yǔ gòucífǎ" [Phonological Alternations and Derivational Morphology in Tangut], *Bulletin of the Institute of History and Philology* 64, no. 4 (1993): 935–68; Gong, "A Hypothesis of Three Grades and Vowel Length Distinction in Tangut."

5 Gong Hwang-cherng, *Xīxià yǔwén yánjiù lùnwénjí*, 150.

columns can be imputed to the absence of vowel length distinction in Chinese. Hence, he hypothesizes the paired column distinction as one of vowel length.

This line of argument is weak. To date, in both loanword materials and etymological comparison, the literature has not revealed any external correlates of the paired columns.

Gong Hwang-cherng's partition of Tangut rhymes into column-1 and column-2 rhymes gained wide acceptance among Tangutologists; his theory that this distinction reflects one of vowel quantity less so. Gong Hwang-cherng himself, in an interview conducted by Jackson T.-S. Sun, stated his lack of certainty concerning the actual value of "vowel length."[6] Later reconstruction schemes, such as those of Arakawa Shintarō 荒川慎太郎 and Marc Miyake, recognize both the validity of paired columns and the tenuity of the vowel quantity theory by marking column-2 rhymes with the semantically vacuous prime symbol -ˊ.[7]

1.2 Chronological Layers in Chinese-to-Tangut Transcription Materials

Four types of material are essential for the reconstruction of Tangut phonology[8]: native dictionaries compiled by Tangut scholars, transcription/loanwords, comparison between Tangut and Burmo-Qiangic languages, especially the closely related modern Rgyalrongic languages[9], and internal reconstruction based on what Gong Hwang-cherng terms "phonological alternations."[10]

In this essay, I am drawing in particular on Chinese-to-Tangut transcription materials; in other words, Tangut words borrowed from Chinese as

---

6 Jackson T.-S. Sun, "Gōng Huángchēng Yuànshì tán Xīxiàyǔ yánjiù" [Academician Gong Hwang-cherng on Tangut research], *Shengyun luncong* 13 (2004): 7–9.

7 See Shintarō Arakawa, "Kazō taionshiryō kara mita Seikago no seichō" [A Study on Tangut Tones from Tibetan Transcription Materials], *Gengogaku Kenkyū* 17–18 (1999): 27–44; Marc Miyake, "Complexity from Compression: a Sketch of Pre-Tangut," in *Tanguty v Central'noj Azii: sbornik statej v čest' 80-letija prof. E. I. Kyčanova* [Tanguts in Central Asia: A Collection of Articles Marking the 80th Anniversary of Prof. E. I. Kychanov], ed. Irina Popova (Moscow: Oriental Literature, 2012), 244–61.

8 For recent introductions on the sources for Tangut phonology, see Chung-pui Tai. *Xīxiàwén fójīng cánpiàn de Zàngwén duìyīn yánjiù* [A Study of Tibetan Phonological Transcription in Tangut Buddhism Fragments] (PhD thesis, Chinese Academy of Social Sciences, 2008), 5–10; Shintarō Arakawa, "Seikagoon fukugen no tame no kakushu shiryō" [Sources for the Reconstruction of Tangut Phonology], *Rekishi to chiri* 629 (2009): 27–35; Guillaume Jacques, *Esquisse de phonologie et de morphologie historique du tangoute* (Leiden: Global Oriental, 2014), 1–8; Xun Gong, "Uvulars and Uvularization in Tangut Phonology," *Language and Linguistics* 21, no. 2 (2020): 176–80.

9 See Guillaume Jacques, *Dictionnaire Japhug-chinois-français, Version 1.1*. (2016); Yunfan Lai et. al., "Tangut as a West Gyalrongic Language," *Folia Linguistica Historica* 41, no. 1 (2020): 171–203.

10 See Gong, "Phonological Alternations in Tangut"; etc.

well as Tangut transcription of Chinese. These materials form part of Sino-Tangutica[11]—namely, the totality of transcriptional and lexical materials that arose in the language contact between Tangut and Chinese. Apart from Chinese-to-Tangut materials, which we will examine in some detail, we also have at our disposal materials in the opposite direction, from Tangut to Chinese, of which the best-known example is the language textbook *Pearl in the Palm* ([illegible] {*mji*^{2}*zar*^{1} *ŋwuu*^{1}*dzjɨj*^{1} *bju*^{1}*pjạ*^{1}*gu*^{2}*njị*^{2}}, 番漢合時掌中珠 *Fānhàn Héshí Zhǎngzhōngzhū*), in which Tangut words and phrases are phonetically transcribed in Chinese.

Sino-Tangutica has been essential to the reconstruction of Tangut phonology from the very beginning of the enterprise. Its familiarity, however, should not diminish its interest. The Chinese-to-Tangut material is particularly important for an often overlooked feature: its internal divergence into different chronological strata, which shed light on sound changes both within Tangut itself and in the source Chinese varieties.

The majority of Chinese-to-Tangut evidence can be subsumed into one of two categories:

- I use the term *established borrowings* to refer to the cases analyzed in Gong Hwang-cherng's groundbreaking article, "Chinese loanwords in the Tangut language."[12] They concern words of Chinese origin that had either been assigned a dedicated Tangut character or had otherwise been identified as such in native character dictionaries. As Gong has shown, this corpus, which dates to the mid-11th century CE, already shows a degree of internal divergence. In particular, there is an older layer, corresponding to an older stage of the Chinese language, closer to Early Middle Chinese and a newer layer, basically resembling the Late Sino-Tangut pronunciation.
- I use the term *Late Sino-Tangut pronunciation* to refer to the system(s) of pronouncing Chinese characters as reflected in proper names and spe-

[11] The term *Sino-Tangutica* designates that which in Chinese is called *Xià-Hàn duìyīn cáiliào* 夏漢對音材料. The conventional translation of *duìyīn* 對音as *transliteration* / *transcription* mischaracterizes the situation, since we are not, outside of language textbooks like the *Pearl in the Hand*, dealing with transcriptions per se. Instead, the sociolinguistic situation much resembles that of English words in Hindi-Urdu or Japanese, where bilingualism, at least in terms of vocabulary, is prevalent; where almost any Chinese word can be borrowed in Tangut; but where there is nevertheless a heavy adaptation to the target-language phonology. The ambiguous term *Sino-Tangutica* better captures this ill-defined middle ground between transcription, code-switching, and borrowing.

[12] Gong, Hwang-cherng, "Xīxiàyǔ zhōng de Hànyǔ jièci" [Chinese Loanwords in the Tangut Language], *Bulletin of the Institute of History and Philology* 52, no. 4 (1981): 681–780.

cialist vocabulary of Chinese origin in a wide range of Tangut-language literature, especially legal and technical literature, as well as translations from Chinese originals, mainly from the mid-12th century CE onwards, in conjunction with the Tangut transcription of Chinese in the *Pearl in the Hand*.

In this essay, I mainly rely on Gong Hwang-cherng's forementioned article (1981c) as my source for established borrowings, and on Gong Hwang-cherng's article from 1991 [13] — an analysis of proper names and bureaucratic titles of Chinese origin in the *Lèilín* (類林, [illegible] {*djij*^{1}*bo*^{1}})— for Late Sino-Tangut pronunciation.

Two major consonant shifts demarcate Mediaeval Héxī 河西 Chinese, the dialect(s) of Mediaeval Northwestern Chinese in heavy contact with Tangut, from its precursor, Early Middle Chinese (EMC):

- Héxī **mb*- < EMC *m*- 明: in Mediaeval Northwestern Chinese dialects in general, EMC nasals turned, either allophonically or definitely, into prenasalized voiced stops[14,15];
- Héxī **ph*- < EMC *b*- 並: in Mediaeval Héxī Chinese, EMC voiced stops turned into their voiceless aspirated counterparts.[16]

As Gong Hwang-cherng demonstrates in his article "Chinese loanwords in the Tangut language," two chronological layers can be distinguished within the corpus of established loanwords. In an earlier layer, Tangut initials reflect the original EMC forms; in a later layer, Tangut initials reflect later Héxī reflexes. The Late Sino-Tangut pronunciation, unsurprisingly, also reflects the chrono-

13 Gong, Hwang-cherng, "Lèilín Xīxiàwén yìběn Hàn-Xià duìyīnzì yánjiù" [A study on the Tangut transcription of Chinese in the Tangut translation of the Lei-lin], in *Kǎogǔ yǔ lìshǐ wénhuà* (Qìngzhù Gāo Qùxún Xiānshēng Bāshí Dàshòu Lùnwénjí) [Anthropology and Historical Cultures: Festschrift on the Occasion of Kao Chü-hsün's 80th Birthday], ed. Wen-hsün Sung, vol 2 (Taipei: Cheng Chung, 1991), 185–223.

14 The term "stops," *par abus de langage*, designates both stops and affricates in this essay.

15 See Henri Maspero, "Le dialecte de Tch'ang-ngan sous les T'ang," *Bulletin de l'Ecole française d'Extrême-Orient* 20, no. 2 (1920): 29–36; Luo Charngpeir, *Táng Wǔdài Xīběi Fāngyīn* [Northwestern Dialect of Táng and Five Dynasties Period] (Pei-p'ing: Academia Sinica, 1933), 29–30.

16 See Gong, Hwang-cherng, "Shíèr shìjì mò Hànyǔ de Xīběi fāngyīn (shēngmǔ bùfèn)" [A Northwestern Dialect of Chinese at the End of the 12th Century, Part 1: Initials]. *Bulletin of the Institute of History and Philology* 52, no. 1 (1981): 37–78; Lǐ Fànwén, *Sòngdài Xīběi fāngyīn:* Fānhàn Héshí Zhǎngzhōngzhū *duìyīn yánjiū* [The Northwestern Dialect of Chinese During the Sòng Period: A Study of the Transcription Practices in the *Pearl in the Palm*] (Beijing: Zhongguo shehui kexue chubanshe, 1994).

logically later stage. The stages of sound change reflected in the chronological layers of Chinese-to-Tangut evidence are shown in Table 1.

Table 1: Chronological layers of Chinese-to-Tangut evidence.

	Chinese borrowing into Tangut / Tangut transcription of Chinese		
	Established borrowings		Late Sino-Tangut
EMC Initial	Earlier layer	Later layer	
p- 幫	Tangut {*p-*}		
	𗅁$_{3803}$ {*pjij*^{1}} < 邊 *pen*		𗦳$_{4196}$ {*pio*^{1}} < 包 *pæw*
ph- 滂	Tangut {*ph-*}		
	𗂧$_{4007}$ {*pha*^{1}} < 破 *phaH*		𗥃$_{2489}$ {*phej*^{2}} < 沛 *phajH*
b- 並	Tangut {*b-*}	Tangut {*ph-*} < Héxī **ph-*	
	𗾞$_{0876}$ {*bã*^{1}} < 盤 *ban*	𗰗$_{3648}$ {*phiej*^{2}} < 稗 *bɛH*	𗥃$_{2667}$ {*phu*^{1}} < 部 *buwX*
m- 明	Tangut {*m-*}	Tangut {*b-*} < Héxī **mb-*	
	𗡶$_{1130}$ {*mjij*^{1}} < 糜 *mie*	𗍫$_{4964}$ {*bioo*^{1}} < 貓 *mæw*	𗍘$_{2736}$ {*biaa*^{2}} < 馬 *mæX*

1.3 Scope and Structure of this Essay

Gong Hwang-cherng reconstructed "long vowels," or, in the noncommittal terms of the present essay, column-2 rhymes, systematically in the major cycle (R.01–R.60). He also reconstructed "long vowels" for several rhymes in the second minor cycle (R.80–R.98) and the totality of the third minor cycle (R.99–R.103). Only the rhymes of the first minor cycle (R.61–R.79) do not show any phenomenon of paired columns. The scope of this essay, however, is restricted to Tangut rhymes in the major cycle.

This choice follows, most of all, from the fact that the Sino-Tangut materials, which lie at the fulcrum of the argument, are found almost exclusively in the major cycle. This disproportionate concentration also applies, to a lesser degree, to the other sources examined in this essay. Preliminary research, moreover, shows that the rhymes assigned to column 2 by Gong Hwang-cherng in the minor cycles do not show the same behavior with respect to transcriptional and etymological data as column 2 rhymes in the major cycle, suggesting that those rhymes can be considered an entirely different phenomenon, if indeed they can be regarded as one single class at all. I consider what Gong Hwang-cherng assigned to column 2 in the second and third minor cycles to be unrelated to the subject of this essay and relegate discussion of the nature of these rhymes to future papers.

After this introductory section, Section 2 discusses the behavior of Tangut paired columns in Sino-Tangut transcription and loanword materials and proposes the hypothesis that in syllables with voiced stop initials, the concrete nature of column-2 rhymes is akin to prenasalized voiced stops. Section 3 shows how this hypothesis is to be generalized across initial types: column-2 rhymes are proposed to indicate a nasal preinitial in Tangut phonology. Before moving on to the conclusion, Section 4 discusses how the nasal preinitial hypothesis interacts with, is supported by, or otherwise improves the treatment of other sources for the reconstruction of Tangut phonology, notably *fǎnqiè* evidence in native dictionaries as well as comparative evidence in modern Rgyalrongic languages.

2. PAIRED COLUMNS OF RHYMES IN SINO-TANGUT MATERIALS

2.1 Column-2 Rhymes and Prenasalized Initials in Mediaeval Hexi Chinese

Huáng Kǎn 黃侃 famously said that the essence of philology lies in "uncovering" *fāmíng* 發明, i.e. of hidden connections between well-known materials, and not "discovering" *fāxiàn* 發現, i.e. of new materials. In a twist reminiscent of the eccentric, a robust correlate to the paired columns, which has long eluded the search of Tangutologists, reveals itself in an all too familiar place. As this section will demonstrate, the key to understanding the nature of Tangut "vowel length" lies in the Chinese-to-Tangut evidence.

A starting point for examining this question is Gong Hwang-cherng's article from 1991, which contains a useful table[17] containing all the Chinese syllables transcribed or borrowed into Tangut in the *Lèilín*, ordered by the Tangut rhymes of the target syllables. Once one examines specifically the column-2 rhymes enlisted to render Chinese words, one could not fail to notice that Tangut syllables there mostly render Chinese syllables with the Middle Chinese nasal initials 明 *m-*, 泥 *n-*, 娘 *ɳ-*, and 疑 *ŋ-*.

Among the rhymes containing Chinese-to-Tangut syllables in *Lèilín*, thirteen major-cycle rhymes are reconstructed by Gong Hwang-cherng with "long vowels"—that is, assigned as column-2 rhymes. Among these thirteen rhymes, eleven of them, shown in Table 2, almost exclusively transcribe Chinese syllables with EMC nasal initials. The remaining two rhymes, R.21 and R.59, exhibit

[17] The parts of the table reproduced in this essay come from the reprint Gong, Hwang-cherng, *Xīxià yǔwén yánjiù lùnwénjí* 西夏語文研究論文集 [Collected Papers on Tangut Philology] (Taipei: Academia Sinica, 2002), 454–57.

the diametrically opposite behavior: as shown in Table 3, they exclusively transcribe Chinese syllables with non-nasal initials. Insofar as we understand the column 1/2 distinction as a unitary phonemic distinction, we can safely reject them as column-2 rhymes, given their completely column-1 behavior; their actual nature will be discussed in forthcoming papers.

Table 2. Tangut column-2 rhymes that predominantly transcribe Chinese nasal initials in Gong's *Xīxià yǔwén yánjiù lùnwénjí*, 454–57.

Rhyme	GHC reconstruction	MC nasal initials	MC non-nasal initials
R.5 (1.5–2.5)	{*-uu*}	奴 *nu* 罗 *nuX* 牟 *miuw* 謀 *miuw* 母 *muwX* 母 *muwX* 穆 *miuwk* 五 *ŋuX* 伍 *ŋuX* 慕 *muH* 漠 mak 茂 *muwH* 謀 *miuw* 嫫 *mu*	
R.7 (1.7–2.6)	{*-juu*}	玉 *ŋiʌwk* 圉 *ŋiʌX* 語 *ŋiʌX* 漁 *ŋiʌ* 虞 *ŋiu* 御 *ŋiʌH* 女 *ŋiʌX*	蜀 *dzʌwk* 余 *jʌ* 禹 *ɣiuX* 餘 *jʌ* 于 *ɣiu* 瑀 *ɣiu* 瑜 *ju* 羽 *ɣiuX* 猶 *juw*
R.12 (1.22–2.11)	{*-ee*}	默 *mʌk* 墨 *mʌk*	
R.14 (1.14–2.12)	{*-ii*}	儀 *ŋie* 毅 *ŋiijH* 凝 *ŋiŋ* 密 *mit* 宓 *mit* 靡 *mieX* 汨 *mek* 糜 *mie*	
R.22 (1.22–2.19)	{*-aa*}	末 *mat* 熬 *ŋaw* 鼻 *ŋawH* 納 *nʌp*	
R.23 (2.20)	{*-iaa*}	牙 *ŋæ* 顏 *ŋæn* 蠻 *mæn* 茆 *mæwX* 馬 *mæX* 雅 *ŋæX*	晏 *ʔænH*
R.24 (1.23–2.21)	{*-jaa*}	鄴 *ŋiæp* 輦 *lienX*	
R.33 (1.32–2.29)	{*-jii*}	岌 *ŋip*	
R.38 (1.37–2.34)	{*-ej*}	艾 *ŋajH* 内 *nwʌjH*	哀 *ʔʌj*
R.54 (1.52–2.45)	{*-oo*}	穆 *miuwk* 摩 *ma* 莽 *maŋX* 茂 *muwH*	
R.55 (1.53–2.46)	{*-ioo*, *-joo*}	樂 *ŋæwH* 岳 *ŋæwk*	

Table 3. Tangut column-2 rhymes that exclusively transcribe Chinese non-nasal initials in Gong's *Xīxià yǔwén yánjiù lùnwénjí*, 454–57.

Rhyme	GHC reconstruction	MC nasal initials	MC non-nasal initials
R.21 (1.21–2.18)	{*-aa*}		葛 *kat* 照 *tɕewH* 少 *ɕewH* 昭 *tɕew* 藁 *kawX* 瑤 *jew* 陶 *jew* 遼 *lew* 鷯 *lew* 邵 *dʑewH* 紹 *dʑewX*
R.59 (1.57)	{*-ioow*}		叔 *ɕuwk* 屬 *dʑʌwk* 蜀 *dʑʌwk* 筑 *ʈiuwk*

The significance of this becomes clear once it is taken into account that the Tangut syllables used to render Chinese nasals have voiced stop initials instead. They reflect the Héxī reflexes of EMC nasals as prenasalized voiced stops:

- Tangut {*b-*} for Héxī **mb-* < EMC *m-* 明: [Tangut]$_{2736}$ {*biaa*^{2}} for 馬 *mæX*, 蠻 *mæn*, 茆 *mæwX*
- Tangut {*d-*} for Héxī **nd-* < EMC *n-* 泥: [Tangut]$_{5250}$ {*daa*^{2}} for 納 *nʌp*
- Tangut {*dź-*} for Héxī **ndʐ̩-* < EMC *ɳ-* 娘: [Tangut]$_{4706}$ {*dźjuu*^{2}} for 女 *ɳiʌX*
- Tangut {*g-*} for Héxī **ŋg-* < EMC *ŋ-* 疑: [Tangut]$_{3590}$ {*gjii*^{1}} for 儀 *ŋi*, 毅 *ŋiijH*, 凝 *ŋiŋ*

The Sino-Tangut transcription in *Lèilín* reveals an affinity between the column-2 rhymes of Tangut and prenasalized voiced stops in Mediaeval Héxī Chinese. Could these supposedly "long vowel" rhymes indicate some kind of prenasalization? This question will be further discussed in §2.2. In the meanwhile, the exceptions need to receive a brief examination.

Two kinds of exceptions to this generalization exist. The first category, more apparent than real, concerns the Chinese syllables with weak, zero, or zero-like, initials: *ʔ-* 影, *ɣi-* 云, or *j-* 以, which are transcribed in Tangut with the initial consonant *g-*. This can be reasonably accounted for by an internal change in source Chinese dialects, from weak initials into *^{n}*g-* (影喻入疑 *yǐng-yù rù yí*):

- [Tangut]$_{4031}$ {*gjuu*^{2}} renders the Chinese syllable 于 *ɣiu* with a weak initial, but also Chinese syllables with the expected 疑 *ŋ-* initial such as 玉 *ŋiowk* and 御 *ŋiʌH*.
- [Tangut]$_{0775}$ {*gjuu*^{1}} transcribes the Chinese syllables 于 *ɣiu*, 羽 *ɣiuX*, 猶 *juw*, 禹 *ɣiuX*, and 瑀 *ɣiuX* with weak initials, but also Chinese syllables with the expected 疑 *ŋ-* initial such as 玉 *ŋiowk*, 御 *ŋiʌH*, 虞 *ŋiu*, 語 *ŋiʌX*, 圉 *ŋiʌX*, and 漁 *ŋiʌ*.

- 𘞃$_{3323}$ {*ŋiaa²*} "goose" (the reconstruction of which should be revised to {*giaa²*}, cf. §4.2.2) transcribes the Chinese syllable 晏 *ʔænH* with a weak initial, but also Chinese syllables with the expected 疑 *ŋ-* initial such as 雅 *ŋæX*, 玡 *ŋæH*, 顏 *ŋæn*, and 牙 *ŋæ*.
- 𗯆$_{1009}$ {*geej²*} transcribes the Chinese syllable 哀 *ʔʌj* with a weak initial, but also the Chinese syllables with the expected 疑 *ŋ-* initial such as艾 *ŋajH*.

The case of 𗗙$_{3745}$ {*dzjaa²*}, which transcribes 輦 *lienX*, might also reflect variant pronunciation in the source Chinese dialect. The Tangut form reflects a likely non-standard pronunciation **ɳienX*, which had since become mainstream in contemporary Chinese dialects, cf. Modern Běijīng *niǎn*, Suzhou *ȵiɪ*˥, etc.

There are, however, two indisputably genuine exceptions, where the Tangut form does not have a voiced initial consonant. They are left unaccounted for in this essay.

- 𗉁$_{0151}$ {*śjuu¹*} transcribes the Chinese syllable 蜀 *dʐowk*. However, 蜀 *dʐowk* is also transcribed in the *Lèilín* as 𗉃$_{4425}$ {*śju¹*} and 𘟙$_{5297}$ {*śioow¹*}, the latter of which belongs to rhyme R.59, mistakenly assigned by Gong Hwang-cherng to column 2.
- 𘒗$_{5821}$ {*·juu¹*} renders the Chinese syllables 瑜 *ju*, 餘 *jʌ*, 余 *jʌ*. However, these syllables are also transcribed in the *Lèilín* as 𗦫$_{3519}$ {*·ju¹*} (餘) and 𗏇$_{1778}$ {*·ju²*} (瑜).

In conclusion, in the corpus of the Sino-Tangut transcription in *Lèilín*, Tangut column-2 rhymes almost exclusively contain Tangut syllables with a voiced stop initial, which render Chinese syllables with Northwest-type prenasalized voiced reflexes of EMC nasals.

2.2 Paired Columns in Chinese Loanwords in Tangut

This affinity between Tangut column-2 rhymes and prenasalized voiced stops in Sino-Tangut seems rather telling. Could column-2 rhymes actually express not long vowels, but the presence of prenasalization on the initials instead? Only by recourse to a contrastive scenario with minimal or near-minimal pairs could we determine the exact nature of this contrast.

We immediately encounter a problem: no attested or major reconstructed variety of Middle or Late Mediaeval Chinese contrasts plain voiced *b-* with prenasalized *mb-*. We can then remind ourselves that loanwords in the target language freeze the source form at the precise time and place of borrowing. As we saw in §1.2, plain voiced *b-* exists in older forms of Chinese continuing EMC

b- 並, while prenasalized *mb-* exists in Mediaeval Héxī Chinese < EMC *m-* 明. A synchronically non-existent *b-:mb-* contrast can be collaged, so to speak, from the different chronological strata of borrowings. In other words, a Tangut syllable with the initial *b-* used to render something Chinese could be an example of either of the following two cases:

- If the Tangut initial *b-* renders a Chinese syllable in the older layer of established Chinese loanwords, in which case we speak of an *old voiced stop*, the source form would have the original plain voiced value of the EMC voiced stop *b-* 並.
- If the Tangut initial *b-* occurs in one of the newer layers of Tangut-to-Chinese material, such as the newer layer of established Chinese loanwords or Late Sino-Tangut, in which case we speak of a *new voiced stop*, the source form would have the prenasalized Héxī reflex *mb-* of the EMC nasal *m-* 明 instead.

If the actual nature of Tangut paired columns of rhymes does involve Northwest-type prenasalization, one would expect loanwords with old voiced stops (borrowed from EMC *b-* 並, etc.) to occur exclusively in column-1 rhymes, and loanwords with new voiced stops (borrowed from Héxī **mb-* < *m-* 明, etc.) to occur exclusively in column-2 rhymes. As we shall soon see, this hypothesis is confirmed by an exhaustive investigation of established Chinese loanwords in Tangut with voiced stop initials.

2.2.1 Tangut Reflexes of Chinese Borrowings with Old Voiced Stops

We first examine all established Chinese loanwords in Tangut covered in Gong Hwang-cherng's "Xīxiàyǔ zhōng de Hànyǔ jièci" that belong to a rhyme in the major cycle (R.1–R.60) with an old voiced stop, i.e., those that have a Tangut voiced stop initial which renders a voiced stop initial in the Chinese source. We expect them all to belong to Tangut column-1 rhymes.

First, we examine the Tangut syllables in {*b-*} borrowed from Chinese etyma in EMC *b-* 並. There are four of them. As expected, all four belong to column-1 rhymes.

- Tangut [Tangut]$_{0876}$ {*bã*^{1}} "tray, plate" is borrowed from Chinese 盤 *ban* "id."
- Tangut [Tangut]$_{3806}$ {*bu*^{1}}, in the disyllable [Tangut][Tangut] {*bu*^{1}*lọ*^{1}} "weed," is borrowed from Chinese 蒲 *bu* "cattail."
- Tangut [Tangut]$_{1971}$ {*bia*^{2}} "to crawl, to creep" is borrowed from Chinese 爬 *bæ* "id."

- Tangut [Tangut]$_{1508}$ {*bej*^{1}} "to lose and flee in a war" is borrowed from Chinese 敗 *bæiH* "to lose, to fail."

Four Tangut syllables in {*d*-} are borrowed from Chinese etyma in EMC *d*- 定. Three belong to column-1 rhymes as expected, but there is one exception.

- Tangut [Tangut]$_{3098}$ {*djɨj*^{2}} "to stop, to rest" is borrowed from Chinese 停 *deŋ* "id."
- Tangut [Tangut]$_{2833}$ {*djɨj*^{2}} "tranquility, certainly" is borrowed from Chinese 定 *deŋH* "id."
- Tangut [Tangut]$_{0712}$ {*du*^{2}} (rhyme 2.4), in the disyllable [Tangut][Tangut] {bə^{2}du^{2}} "襪肚 *wàdù*, a kind of scarf worn as loincloth" is borrowed from Chinese 肚 *duX* "belly."
- Tangut [Tangut]$_{2799}$ {*dwəə*^{1}} "protruding, concave" is considered by Gong Hwang-cherng to be borrowed from Chinese 凸 / 突 *duʌt* "id." The Tangut word belongs to a column-2 rhyme, and thus shows an exceptional correspondence. One might be tempted to dismiss this exception as a non-borrowing. This course should nonetheless not be taken too lightly, given that the rhyme correspondence, Tangut {-*ə*} for Chinese 没 -*uʌt*, is corroborated in late Sino-Tangut pronunciation: [Tangut]$_{5233}$ {*phə*^{1}} transcribes Chinese 渤 *ph- < *buʌt*.

Two Tangut syllables in {*dź*-} are borrowed from Chinese etyma in EMC *ḍ*- 澄. As expected, both belong to column-1 rhymes.

- Tangut [Tangut]$_{0443}$ {*dźjo*^{1}} "long" is borrowed from Chinese 長 *ḍiaŋ* "id."
- Tangut [Tangut]$_{4411}$ {*dźjwã*^{1}} "rafter" is borrowed from Chinese 椽 *ḍiwen* "id."

Four Tangut syllables in {*g*-} are borrowed from Chinese etyma in EMC *g*- 羣. As expected, all four belong to column-1 rhymes.

- Tangut [Tangut]$_{5503}$ {*gju*^{1}} "canal, ditch" is borrowed from Chinese 渠 *giʌ* "id."
- Tangut [Tangut]$_{0005}$ {*gjow*^{1}} "to win" is borrowed from Chinese 强 *giaŋ* "strong."
- Tangut [Tangut]$_{3879}$ {*gju*^{2}} "utensil, container" is borrowed from Chinese 具 *giuH* "utensil, tool."
- Tangut [Tangut]$_{5501}$ {*gju*^{2}} "tool" is borrowed from Chinese 具 *giuH* "utensil, tool."

Five Tangut syllables in {*dz*-} are borrowed from Chinese etyma in EMC *dz*- 從. As expected, all five belong to column-1 rhymes.

- Tangut [illegible]$_{2982}$ {*dzwa*^{1}} “short in stature” is borrowed from Chinese 矬 *dzwa* “id.”
- Tangut [illegible]$_{1604}$ {*dzjɨj*^{1}} “money” is borrowed from Chinese 錢 *dzien* “id.”
- Tangut [illegible]$_{5097}$ {*dzwej*^{1}} “crime, *agha*” is borrowed from Chinese 罪 *dzwʌjX* “id.”
- Tangut [illegible]$_{2549}$ {*dza*^{1}} “mixed” is borrowed from Chinese 雜 *dzap* “id.”
- Tangut [illegible]$_{4170}$ {*dza*^{1}} “to chisel” is borrowed from Chinese 鑿 *dzak* “id.”

Finally, two Tangut syllables in {*dź-*} are borrowed from Chinese etyma in EMC *dʐ-* 崇. As expected, both belong to column-1 rhymes.

- Tangut [illegible]$_{0561}$ {*dźio*^{2}} “to help” is borrowed from Chinese 助 *dʐiʌH* “id.”
- Tangut [illegible]$_{3407}$ {*dźiow*^{2}} “official report” is borrowed from Chinese 狀 *dʐiaŋH* “id.”

There are twenty-three established loanwords with old voiced stops. Twenty-two among them belong to column-1 rhymes. Only one exception belongs to a column-2 rhyme, which is however unexplained.

2.2.2 Tangut Reflexes of Chinese Borrowings with New Voiced Stops

We now turn our attention to established major-cycle Chinese loanwords with a new voiced stop, i.e., those borrowed from Chinese etyma with a nasal initial. First, we examine the Tangut syllables in {*b-*} borrowed from Chinese etyma with Héxī **mb-*, reflecting a Northwest-type outcome of EMC *m-* 明. There are two of them. As expected, both belong to column-2 rhymes.

- Tangut [illegible]$_{4964}$ {*bioo*^{1}} “cat” is borrowed from Chinese 貓 *mæw* “id.”
- Tangut [illegible]$_{2138}$ {*buu*^{2}} “tomb” is borrowed from Chinese 墓 *muH* “id.”

One Tangut syllable in {*d-*} is borrowed from a Chinese etymon with Héxī **nd-* < EMC *n-* 泥. As expected, it belongs to a column-2 rhyme.

- Tangut [illegible]$_{2637}$ {*duu*^{1}} “slave” is borrowed from Chinese 奴 *nu* “id.”

One Tangut syllable in {*dź-*} is borrowed from a Chinese etymon with Héxī **ndʐ-* < EMC *ɳ-* 娘. As expected, it belongs to a column-2 rhyme.

- Tangut [illegible]$_{4706}$ {*dźjuu*^{2}} “woman” is borrowed from Chinese 女 *ɳiʌX* “id.”

Two Tangut syllables in {*g*-} are considered by Gong Hwang-cherng to be borrowed from Chinese etyma with Héxī **ŋg*- < EMC *ŋ*- 疑. One, however, is likely not to be a loanword. The other is exceptional, belonging to a column-1 rhyme.

- Tangut 𗂅$_{0750}$ {*gjii*^{1}} "to chew, to hold by teeth" is considered by Gong Hwang-cherng to be borrowed from Chinese 齧 *ŋet* "to bite."[18] I previously argued that this is an inherited word, cognate to Japhug *kɤ-nɤŋka*.[19]
- Tangut 𗆴$_{1478}$ {*gjĩ*^{1}} "to examine, to check" is borrowed from Chinese 驗 *ŋiemH* "id." This exception is expected, given that nasalized rhymes do not have a distinction of paired sequences.

There are five safe Chinese loanwords with a new voiced stop. Four among them belong to column-2 rhymes. Only one exception belongs to a column-1 rhyme, which can be explained by the rhyme of the syllable. These examples show the same behavior as the Late Sino-Tangut pronunciation, shown in Table 2.

2.3 Conclusion

The evidence discussed in §2.1 and §2.2 suggests a radical revision of the reconstruction of Tangut paired columns of rhymes. Rather than treating the distinction between column 1 and column 2 as one between short and long vowels, it is far more natural to view it as one between the absence and presence of an initial nasal element.

Table 4 shows near-minimal pairs and other contrastive examples between the different chronological strata of Sino-Tangut materials. A column-1 syllable like 𗼃$_{3806}$ {*bu*^{1}} is borrowed from EMC 蒲 *bu*; a column-2 syllable like 𗿼$_{2138}$ {*buu*^{2}} is borrowed from Héxī 墓 **mbuH* < EMC *muH*, with a Héxī prenasalized stop **mb*- corresponding to an EMC nasal. This immediate correspondence allows us to project the Chinese situation straight onto Tangut: column-2 rhymes actually indicate prenasalization, which is absent in column-1 rhymes.

18 Gong, "Xīxiàyǔ zhōng de Hànyǔ jiècì."

19 Xun Gong, *Le rgyalrong zbu, une langue tibéto-birmane de Chine du Sud-ouest : une étude descriptive, typologique et comparative* (PhD thesis, INALCO, 2018), 302–3, cf. §4.3.6.

Table 4. Proposed revision of Gong Hwang-cherng reconstruction with prenasalization.

Column-1 rhymes short vowels → absence of prenasalization			Column-2 rhymes long vowels → presence of prenasalization		
Tangut character	revision	Chinese source	Tangut character	revision	Chinese source
$_{3806}$ {*bu*^{1}}	*bu*^{ʁ1} → *bu*^{ʁ1}	蒲 *bu*	$_{2138}$ {*buu*^{2}}	*buu*^{ʁ2} → *mbu*^{ʁ2}	墓 Héxī **mb-* < *muH*
$_{0712}$ {*du*^{2}}	*du*^{ʁ2} → *du*^{ʁ1}	肚 duX	$_{2637}$ {duu^{1}}	*duu*^{ʁ1} → *ndu*^{ʁ1}	奴 Héxī **nd-* < *nu*
$_{0443}$ {*dźjo*^{1}}	*dźo*^{1} → *dźo*^{1}	長 *ḍiaŋ*	$_{4706}$ {*dźjuu*^{2}}	*dźuu*^{2} → *ndźu*^{2}	女 Héxī **nḍẓ-* < *ŋiʌX*
$_{3879}$ {*gju*^{2}}	*gu*^{2} → *gu*^{2}	具 *giuH*	$_{4031}$ {*gjuu*^{2}}	*guu*^{2} → *ŋgu*^{2}	御 Héxī **ŋg-* < *ŋiʌH*

3. HYPOTHESIS: COLUMN-2 RHYMES HAVE A NASAL PREINITIAL

The marked contrast between older and newer Chinese loanword sources of Tangut voiced stop initials, shown in Table 4, heavily implies that initial prenasalization is the distinguishing element that sets apart column-2 from column-1 rhymes. However, column-2 rhymes occur not only with voiced stop initials, but with other types of initials too. It is therefore necessary to sketch a complete theory of column-2 syllables with different initial types.

Directly generalizing prenasalization is precluded by the existence of column-2 syllables with a nasal initial, such as $_{0330}$ {*mjiij*^{1}} "dream" or $_{4902}$ {*ŋwuu*^{1}} "speech." If we understand prenasalization as a timing effect of soft palate raising, nasals, by definition, cannot be prenasalized. Revising the reconstruction of $_{2138}$ {*buu*^{2}} from *buu*^{ʁ2} to some kind of *mbu*^{ʁ2}, but for syllables with a nasal initial, does not make any sense unless understood as one of the following possibilities:

- An initial consonant cluster of which the first element is a nasal consonant. Thus, $_{0330}$ {*mjiij*^{1}} "dream," reconstructed as *mee*^{1} under the uvularization hypothesis, is to be revised as *mme*^{1}, with an initial "geminate" consonant cluster of *mm-*;
- A minor syllable in a sesquisyllabic phonotactics. For example, $_{0330}$ {*mjiij*^{1}} is to be understood as *m̩me*^{1}, with a demi-syllabic preinitial *m̩* preceding a syllable *me*^{1}.

The difference between these two treatments is neither knowable in principle nor consequential with regard to other aspects of Tangut synchronic and diachronic linguistics. Following general usage in Sino-Tibetan, East Asian, and

Mainland Southeast Asian linguistics, syllables previously reconstructed with a "long vowel" should be analyzed as having a *nasal preinitial.*

The interaction of the nasal preinitial hypothesis with the problem of the reconstruction of the Tangut voiced series needs to be addressed briefly. Gong Hwang-cherng proposes that this voiced stop series of initial consonants should be reconstructed as non-prenasalized plain voiced consonants.[20] Guillaume Jacques, on the other hand, demonstrates that the Tangut voiced series come from the Pre-Tangut prenasalized voiced stop, a result that raises the possibility that even during the Tangut empire period, the voiced series remained prenasalized.[21] The nasal preinitial hypothesis favors the non-prenasalized value of plain voiced consonants, as the contrast between [Tangut]$_{3806}$ {*bu*^{1}} and [Tangut]$_{2138}$ {*buu*^{2}} would be less awkward as one between *bu*^{ʁ1} and *mbu*^{ʁ2} than one between ^{n}*bu*^{ʁ1} and *m*^{n}*bu*^{ʁ2}. The latter scenario is not impossible, as Swahili features a contrast between trisyllabic /m.bu.ni/ "coffee bush" and disyllabic /^{n}bu.ni/ "ostrich," but presupposes a sesquisyllabic treatment.[22]

One final question pertains to the identity of the nasal preinitial. Etymologically (see especially §4.3.1–3), the nasal preinitial predominantly derives from earlier dental **n*-. However, before stops and nasal, there are strong reasons to consider the nasal preinitial as homorganic. Thus, the closest typological parallel is the nasal preinitial of modern West Rgyalrongic languages, likely the closest relatives to Tangut[23], such as Khroskyabs[24] and Geshiza[25]. In these languages, the nasal preinitial *N*- is homorganic before stops: Khroskyabs *ŋgə́*, Geshiza *ŋgæ* "nine." However, before other types of consonants, it can surface as a dental *n*- instead: Khroskyabs *nvə̂*, Geshiza *nvə* "soft."

Based on the Sino-Tangut data, etymological comparisons, and typological parallels, I propose the following distribution before different initial types:

- Before stops or affricates, a homorganic nasal preinitial is reconstructed: *mp*-, *nd*-, *ŋk*-, *ɴɢ*-, *ntsh*-, *ndź*-... For example, [Tangut]$_{2138}$ {*buu*^{2}} "tomb," reconstructed under the uvularization hypothesis[26] as *buu*^{ʁ2}, should be revised into *mbu*^{ʁ2}.

20 Gong Hwang-cherng, "Voiced Obstruents in the Tangut Language," *Bulletin of the Institute of History and Philology* 52, no. 1 (1981): 1–16.

21 Jacques, *Esquisse de phonologie et de morphologie historique du tangoute,* 36–37.

22 Taeko Maeda, *The Mora and the Syllable in KiMvita (Mombasa Swahili) and Japanese* (PhD thesis, SOAS, 2001), 162.

23 Lai et al., "Tangut as a West Gyalrongic Language."

24 Lai, Yunfan, *Grammaire du khroskyabs de Wobzi* (PhD thesis, Université Paris 3, 2017), 48–50.

25 Sami Honkasalo, *A Grammar of Eastern Geshiza: A Culturally Anchored Description* (PhD thesis, University of Helsinki, 2019), 169–72.

26 Gong, "Xīxiàyǔ zhōng de Hànyǔ jiècì."

- Before nasals, Khroskyabs dialects show a dental *n*-.[27] In Tangut, however, the mixed *fǎnqiè* behavior (§4.2.1) and etymological origin in *ɣ- and *ʁ- (§4.3.4) favors an interpretation as a homorganic nasal preinitial. The preinitialed nasal initials can be labeled *geminate nasals*: *mm*-, *nn*-, *ŋŋ*-, *ɴɴ*-. Thus, the reconstruction of 𗼃$_{5702}$ {*mjaa¹*} "ulcer" should be revised from *maa¹* to *mma¹*.
- Before sibilants, the preinitial *n*- is reconstructed: *ns*-, *nś*-, *nz*-, *nź*-. For example, the reconstruction of 𗅋$_{0716}$ {*śjii¹*} "to butcher" should be revised from *śii¹* to *nśi¹*, cf. Khroskyabs *ɲɕî*.
- Before laterals, the preinitial *n*- is reconstructed: *nl*-, *nlh*-. For example, the reconstruction of 𘃡$_{5522}$ {*ljiij²*} "to wait" should be revised from *lee²* to *nle²*, cf. Khroskyabs *njé*.
- Before glides, the nasal preinitial *n*- is reconstructed[28]: *n·w*-, *n·j*-. For example, the reconstruction of 𗄐$_{0320}$ {*·wəə¹*} should be revised from *wəəʶ¹* to *n·wəʶ¹*, cf. Geshiza *nvə*.
- Few instances of attested Tangut syllables in column-2 rhymes have initials *x*-, *ɣ*-, and *·*- (zero initial). Purely as a notation, one could write *nx*-, *nɣ*-, and *n·*- for such cases. The only examples of such syllables with *nx*-, etc. will be argued in §4.2.3 to be spurious, lacking the nasal preinitial in reality.

The proposed reconstruction will be presented again in Table 5 in the conclusion.

4. DISCUSSION

In this section,[29] I will discuss the interaction of the nasal preinitial hypothesis with other sources of evidence on Tangut phonology, starting with other transcription materials (§4.1), followed by the *fǎnqiè* practice in native Tangut rhyme books (§4.2), and concluding in a comparison with modern Rgyalrongic and other Burmo-Qiangic languages (§4.3).

27 Lai, *Grammaire du khroskyabs de Wobzi,* 48–50.

28 Given that the Gong Hwang-cherng reconstruction does not admit a reliable distinction between initial glides *w*-, *j*- and zero initial followed by glide medial *·w*-, *·j*-, Gong Hwang-cherng's zero initial symbol *·*- can be repurposed as an orthographical separator, thus *n·w*-, *n·j*-. Also, with regard to the initial *j*- question, it is assumed that any zero-initial Grade III syllable with a nasal preinitial has an initial yod.

29 Tangut-language literature referred to in this section, unless otherwise indicated, came from the Tangut collection of the Institute of Oriental Manuscripts of the Russian Academy of Sciences, St. Petersburg, published as Institute of Oriental Manuscripts of the Russian Academy

4.1 Nasal Preinitials in the Other Transcription Materials

Tibetan transcription evidence is rather limited in utility, as most cases of the Tibetan nasal preinitial འ <’-> occur before voiced stops indiscriminately in both column-2 rhymes with a Tangut nasal preinitial and column-1 rhymes without one.[30] As discussed in the previous section, there are reasons to prefer both a prenasalized and non-prenasalized value for Tangut voiced stops. The nasal preinitial hypothesis is only moderately in favor of the latter hypothesis. Accordingly, under either belief, this situation can be analyzed as reflecting some residual prenasalization of the Tangut voiced series or reflecting a deprenasalization in the Bde dialect of Tibetan in parallel to the Tangut sound change.[31]

However, it is still noteworthy that two of the only three examples of <’-> preceding a voiceless aspirated stop in the Tibetan transcription[32] involve Tangut column-2 rhymes:

- [Tangut]$_{1223}$ *mpho²* {*phjoo²*} “to combine,” transcribed as འཕོ <’pho>.
- [Tangut]$_{3228}$ *ntho¹* {*thjoo¹*} “wonderful,” transcribed as འཐོ <’tho>.

The one exception is [Tangut]$_{5993}$ *qhaʁ¹* {*kha¹*} “locative particle,” transcribed as འཁ <’kha>. However, this could be understood as one of the orthographical devices used to transcribe in Tibetan letters the Tangut uvular initial.[33]

Sanskrit transcriptions, on the other hand, provide a much more solid basis for discussion. A particularly interesting point to consider is a transcription practice discovered by Arakawa Shintarō[34] in the Tangut version of *Mahāmāyūrīvidyārājñī*.[35] From this collection of mantras, Arakawa documented

of Sciences, Institute of Ethnology and Anthropology of the Chinese Academy of Social Sciences, eds. *Écáng Hēishǔichéng wénxiàn* 俄藏黑水城文獻 / *Pamjatniki pis’mennosti iz Chara-choto chranjaščiesja v Rossii*, 29 vols, 1996–2019. In those cases, they are referenced with the inventory number “Inv. N°.”

30 As one could check in Tai, *Xīxiàwén fójīng cánpiàn de Zàngwén duìyīn yánjiù*, 133–5.

31 I propose labeling the mediaeval dialect(s) of Tibetan under direct language contact with Tangut as the Bde dialect, in the same way that its Chinese counterpart is often labeled the Héxī dialect of Chinese, since the part of the territory of Western Xià that had formerly belonged to the Tibetan empire largely coincides in extent with the Tibetan military governorate of Bde (*bde·khams* / *bde·gams*), both centered on the city of Liángzhōu 凉州.

32 Tai, *Xīxiàwén fójīng cánpiàn de Zàngwén duìyīn yánjiù*, 136.

33 Gong, “Uvulars and Uvularization in Tangut Phonology,” 200.

34 Arakawa, *Seikago tsūin jiten*, 115.

35 The Tangut title [Tangut] is translated from the Tibetan *rig·sngags kyi rgyal·mo rma·bya chen·mo*. The Sanskrit text of the *Mahāmāyūrīvidyārājñī* is cited from Takubo (*Bon-*

the existence of six *nasal-CV characters*. It turns out that five of the proposed nasal-CV characters belong to a column-2 rhyme and therefore, under the nasal preinitial hypothesis, have nasal preinitials:

- [illegible]$_{5792}$ *nda*^{1} {*djaa*^{1}}, used to transcribe *ratnakara**ṇḍa**ke*.
- [illegible]$_{1512}$ *mba*^{ʁ1} {*baa*^{1}}, used to transcribe *a**mba**re a**mba**rāvati*.
- [illegible]$_{5418}$ *ndu*^{1} {*djuu*^{1}}, used to transcribe *vi**ndu**pati*.
- [illegible]$_{5388}$ *mbo*^{ʁ2}{*boo*^{2}}, used to transcribe *ka**mbu***.
- [illegible]$_{3889}$ *mbe*^{2} {*bjiij*^{2}}, used to transcribe *du**ṃbe** dodu**ṃbe***.

Arakawa also considers the column-1 syllable [illegible]$_{5743}$ *gi*^{1} {*gji*^{1}} to be a “nasal-CV” character.[36] This character appears in the segment [illegible][37][illegible] *phə*^{2}*rir*^{2}*gi*^{1}*rir*^{2}*ka*^{1}*·je*^{2} {*phji*^{1}*rjir*^{2}*gji*^{1}*rjir*^{2}*kjaa*^{1}*·jij*^{2}}, used to transcribe *bhṛ**ṅgā**rikāya*, which poses too much of an irregularity to warrant a conclusion one way or the other. There is, however, another possible candidate of a “nasal-CV” character for Sanskrit *ṅgi* / *ṅgī* in the *Mahāmāyūrīvidyārājñī*. In the same *dhāraṇī*, the word *uṭṭiṅgiri*[38] is transcribed as [illegible] *·wu*^{ʁ}*?ti*^{2}*ŋgi*^{2}*rir*^{2} {*·wu?tji*^{2}*gjii*^{2}*rjir*^{2}}. The syllable [illegible]$_{4868}$ *ŋgi*^{2} {*gjii*^{2}} transcribes the segment *ṅgi*, thereby constituting

bun Kujaku myōōkyō / Ārya-Mahā-Māyūrī Vidyā-Rājñī. Tokyo: Sankibō busshorin, 1972), and the Tangut text from Wáng Jìngrú (“Fómǔ Dàkǒngquè Míngwángjīng Xià-Fàn-Zàng-Hàn hébì jiàoshì” [A Comparative Study of the Tangut, Sanskrit, Tibetan and Chinese Editions of the *Mahāmāyūrīvidyārājñī*], *Xīxià yánjiù* 1, ed. Wáng Jìngrú (Pei-p’ing: Academia Sinica, 1932): 181–250.)

36 Arakawa, *Seikago tsūin jiten*, 115.

37 The character [illegible]$_{3948}$ {*kjaa*^{1}} belongs to the rhyme R.21 (1.21–2.18), whose status as a “long vowel” rhyme assigned by Gong Hwang-cherng is mistaken, cf. §2.1.

38 Based on a majority of Sanskrit mss., all written in the Rañjanā script, Takubo Shūyo reads *ḍaṭaṅgini* for this word, which he proposes to emend to *datiṅgani*. (see Takubo, *Bonbun Kujaku myōōkyō / Ārya-Mahā-Māyūrī Vidyā-Rājñī*, 46.) The Sanskrit texts that served as the basis for Takubo’s edition, however, do not agree with most of the mediaeval Chinese and Tibetan testimonies—a point that Takubo failed to note.The Chinese translations read 欝鄧徒拯反祁利, i.e., *uṭṭiṅgiri* or *uḍḍiṅgiri* (Sanghabala 僧伽婆羅 tr., ~ 520CE, Taishō 0984), 嗢徵祇哩, i.e., *uṭṭiṅgiri* (Yìjìng 義淨 tr., ~700CE, Taishō 0985), and 嗢徵上儗霓以反抳, i.e., *uṭṭiṅgiṇi* (Amoghavajra 不空 tr., ~760CE, Taishō 0982). The standard Tibetan version (Derge 0559, Peking 0178) reads ཨུཥྟི་གི་ནི *uṣṭigini*; the re-Sanskritized form, a rather awkward one since Skt *-ṣṭ-* would have yielded an aspirated *-ṭṭh-* instead, points to an underlying shape similar to that of Amoghavajra. Takubo also cited Serge Oldenbourg’s earlier 1899 Sanskrit edition. The nature of this reference must have been rather bibliographical, however, since Oldenbourg, too, reads *uṭṭiṅgari*—see Ol’denburg, Sergej, “Otryvki kašgarskix i sanskritskix rukopisej iz sobranija N. Ḟ. Petrovskago [Some Kashgarian and Sanskrit Manuscripts from N. F. Petrovsky’s Collection]”, *Zapiski vostočnago otdelenija imp. russ. arx. obščestva* 11 (1899), 252. Oldenbourg’s Sanskrit ms., written in a mixed Gupta script with “Kashgar” (= South Turkestan Gupta?) elements (p. 208), must lie rather close to the ancestral text of the Chinese and Tibetan editions.

a possible "nasal-CV" character not discussed in Arakawa.[39] This syllable is also a column-2 syllable as expected. Thus, all currently known reliable "nasal-CV" characters have nasal preinitials under the nasal preinitial hypothesis.

By way of summary, although the Tibetan evidence favors the nasal preinitial hypothesis only slightly, the consistent column-2 status of "nasal-CV" characters in the Sanskrit-to-Tangut transcription of the *Mahāmāyūrīvidyārājñī* strongly supports the nasal preinitial hypothesis.

4.2 Nasal Preinitials and *Fǎnqiè* Evidence

4.2.1 *Fǎnqiè* Behavior of Syllables with Nasal Preinitials

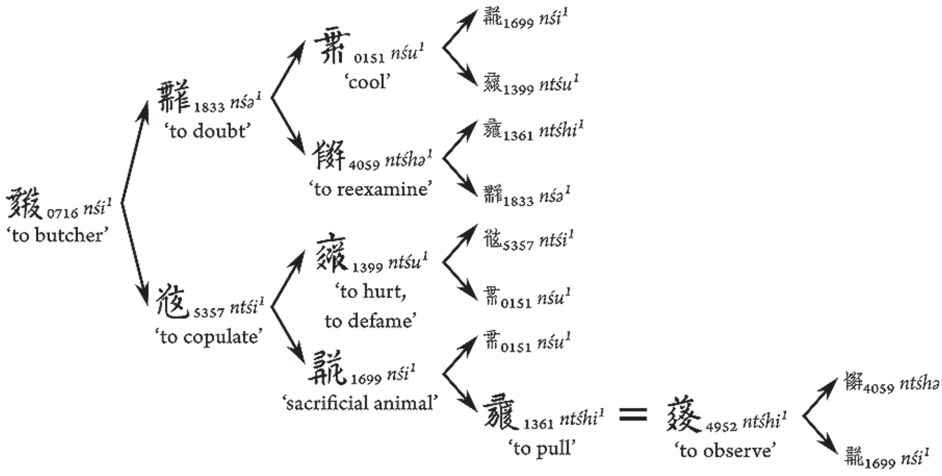
If Tangut column-2 rhymes indeed involve a nasal preinitial, we may expect that native speakers conceive of the presence of the nasal preinitial as part of the initial consonant rather than rhyme. We can catch a glimpse of the native-speaker psychology concerning syllable structure from the practice of *fǎnqiè* 反切 in native dictionaries, a type of phonetic spelling which functions by phonologically segmenting a syllable into the initial, indicated by the initial speller (反切上字 *fǎnqiè shàngzì*), and the rhyme, indicated by the rhyme speller (反切下字 *fǎnqiè xiàzì*). Thus, we would predict that the initial speller of a column-2 rhyme should itself belong to a column-2 rhyme.

It is beyond the scope of this article to attempt an exhaustive examination of all *fǎnqiè xìlián* 反切系聯 sequences in Sofronov in the light of the nasal preinitial hypothesis.[40] I content myself with presenting in Figure 2 all the *fǎnqiè* ancestors of the character 豬$_{0716}$ *nśi*^{1} {*śjii*^{1}} "to butcher." As we can see in Figure 2, all its *fǎnqiè* ancestors of both sides themselves belong to a column-2 rhyme, reconstructed as having a nasal preinitial.

[39] Arakawa, *Seikago tsūin jiten*, 115.

[40] Sofronov, Mikhail V., *Grammatika tangutskogo jazyka* [Grammar of the Tangut Language], vol. 2 (Moscow: Nauka, 1968).

Figure 2: All fǎnqiè ancestors of the character $_{0716}$ *nśi*^{1} {*śjii*^{1}} "to butcher."



Note: Smaller characters indicate loops: they point to characters that already exist in this figure.

This generalization of *matched presence of nasal preinitials*—that both *fǎnqiè* spellers of a syllable with nasal preinitials have nasal preinitials—is not without exceptions. "Geminate" nasal initials (*mm-*, *nn-*, *ŋŋ-*, *NN-*), in particular, can show *unmatched* spellers. For example, the character $_{4902}$ *NNwu*^{ʁ1} {*ŋwuu*^{1}} "speech" is spelt $_{0226}$ *Nwu*^{ʁ1}{*ŋwu*^{1}} + $_{5625}$ *nthwu*^{ʁ1}{*thwuu*^{1}}. $_{4902}$ *NNwu*^{ʁ1} {*ŋwuu*^{1}} belongs to the column-2 rhyme 1.5 in *Wénhǎi*. The column-2 character of this word is also supported by the fact that it is indicated as non-homophonous to column-1 $_{0226}$ *Nwu*^{ʁ1}{*ŋwu*^{1}} in *Homophones* and other dictionaries. Under the nasal preinitials hypothesis, $_{4902}$ *NNwu*^{ʁ1} {*ŋwuu*^{1}} "speech" does have a nasal preinitial.

Thus, in the case of nasal-initial syllables like $_{4902}$ *NNwu*^{ʁ1} {*ŋwuu*^{1}}, the unmatched spellers $_{0226}$ *Nwu*^{ʁ1}{*ŋwu*^{1}} + $_{5625}$ *nthwu*^{ʁ1}{*thwuu*^{1}} are resolved in preference to the rhyme speller $_{5625}$ *nthwu*^{ʁ1}{*thwuu*^{1}}. This *fǎnqiè* behavior is probably unsurprising as *mm-* resembles *m-* acoustically in a way that *mb-* does not resemble *b-*. In §4.2.3, We shall see that this is not always the case with other unmatched spellers.

It is worth pondering, as one anonymous reviewer suggests, whether the fact that the rhyme speller for a column-2 syllable is itself column-2 implies that the distinction of parallel columns does not only involve nasal preinitials but is also associated with some vocalic feature. The current evidence does not allow us to decide the question either way, as the absence of vocalic cor-

relates is plausible given the formal rules of *fǎnqiè*. If there is some kind of vocalic correlate, however, I would lean toward weak nasalization rather than Gong Hwang-cherng's hypothesis of vowel length, given the general direction of compression in which Tangut phonology has developed.[41]

4.2.2 Revising the Tangut Reconstruction of 嶽$_{3323}$ *ɴɴʃa*^{ʁ2} {*ŋiaa*^{2}} "Goose"

The rule of matched presence of nasal preinitials contributes to our growing understanding of the behavior of Tangut *fǎnqiè*, which in turn can help determine the validity of *fǎnqiè* spellings in sources. An example is the character 嶽$_{3323}$, in the disyllable 綴$_{3672}$嶽$_{3323}$ {*bã*^{1}-} "goose." Its pronunciation given in all versions of the *Tangut-Chinese Dictionary*[42] is *ɴɴʃa*^{ʁ2} {*ŋiaa*^{2}}. It is justified by the following *fǎnqiè* spelling given in the dictionary entry: 巍$_{2777}$ *Ne*^{ʁ}*wr*^{1} {*ŋewr*^{1}} + 巂$_{5766}$ *nśʃa*^{ʁ2} {*śiaa*^{2}}. This *fǎnqiè* spelling is doubly suspicious given that:

- 巍$_{2777}$ {*ŋewr*^{1}} is a retroflex syllable, which usually only enters into a *fǎnqiè* relationship with other retroflex syllables. 巍$_{2777}$ {*ŋewr*^{1}} belongs to Sofronov"s *fǎnqiè xìlián* series velar-20, which contains the following characters in Sofronov (1968:81), which are all retroflex: 鼮$_{1357}$ *Nwi*^{ʁ}*r*^{1} {*ŋwer*^{1}} "to equate," 頑$_{0333}$ *Nwi*^{ʁ}*r*^{2} {*ŋwer*^{2}} "keen," 苶$_{4423}$ *Nu*^{ʁ}*r*^{1} {*ŋur*^{1}} "head," 垚$_{4871}$ *Nə*^{ʁ}*r*^{1} {*ŋər*^{1}} "hill," 豙$_{1423}$ *Nwə*^{ʁ}*r*^{1} {*ŋwər*^{1}} "seven," 皺$_{0510}$ *Nwə*^{ʁ}*r*^{1} {*ŋwər*^{1}} "emperor," 矜$_{1219}$ *Nwi*^{ʁ}*r*^{2} {*ŋwer*^{2}} "slack."
- Similarly, as §4.2.1 argues, prenasalized syllables mostly enter into a *fǎnqiè* relationship with other prenasalized syllables, though in the case of *ŋ-* the rule is not as strict.

This *fǎnqiè* spelling of the *Tangut-Chinese Dictionary* is obviously not taken from the *Wénhǎi*, of which only the *píng*-toned volumes have survived to this day. A review of the sources reveals that it originates from the *Combined Edition of Wénhǎi and Homophones* (original title lost; assigned the title 同音文海宝韵合编 by Hán Xiǎománg)[43], which survives chiefly in Inv. No. 4153/4781/6685/8179, a particularly challenging manuscript written in a careless semi-cursive hand

41 See Miyake, "Complexity from Compression: a Sketch of Pre-Tangut"; Gong, "Uvulars and Uvularization in Tangut Phonology," 198–9.

42 Lǐ Fànwén, *Xià-Hàn Zìdiǎn* [The Tangut-Chinese dictionary] (Beijing: Zhongguo shehui kexue chubanshe, 1997, 2008); Lǐ, *Jiǎnmíng Xià-Hàn Zìdiǎn* [The Concise Tangut-Chinese Dictionary] (Beijing: Zhongguo shehui kexue chubanshe, 2012).

43 Hán Xiǎománg, Tóngyin Wénhǎibǎoyùn hébiān *zhěnglǐ yǔ yánjiū* [Combined Edition of the Homophones and *Wénhǎi*: A Critical Edition with Extensive Commentary] (Beijing: Zhongguo shehui kexue chubanshe, 2008).

on the reverse side of another document. The dictionary entry of [Tangut]$_{3323}$ "goose" is reproduced in Figure 3.

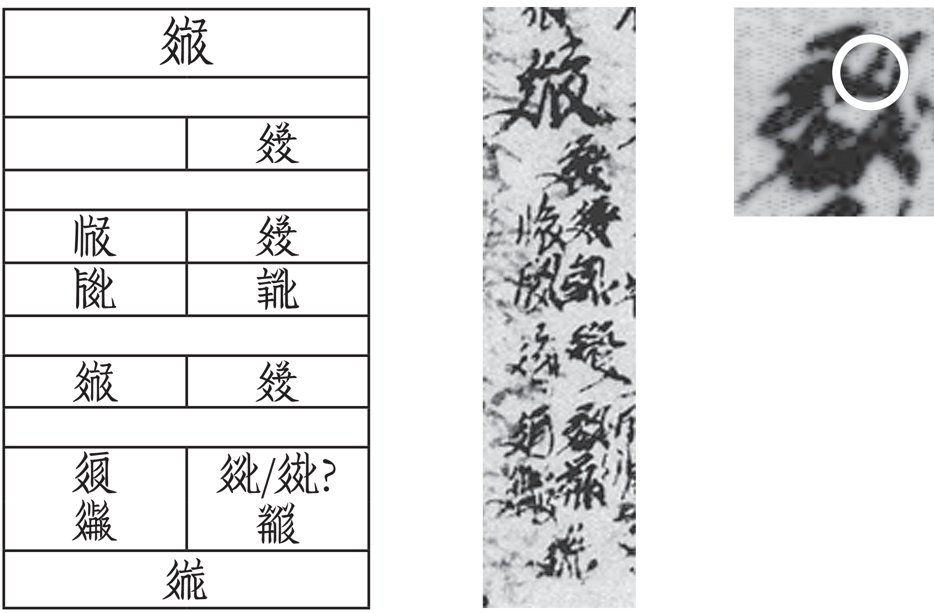


Figure 3: The dictionary entry for the character [Tangut]$_{3323}$ 'goose' in the *Combined edition of Wénhǎi and Homophones* (Inv. No. 4153/4781/6685/8179, 24-15), zooming in on the initial speller in question [cursive form] ([Tangut]? [Tangut]?), with the diagnostic ∠-shape highlighted.

The actual pronunciation of this character thus depends on the reading of the initial speller in the *Combined Edition*, namely [cursive form], which is also shown enlarged in Figure 3. Hán Xiǎománg retains the doubly dubious reading of the *Tangut-Chinese Dictionary*, namely [Tangut]$_{2777}$ *ɴe*^{ʁ}*wr*^{1} {*ŋewr*^{1}}.[44] I propose that the initial speller [cursive form] should instead be read as [Tangut]$_{2776}$ *ŋgu*^{2} {*gjuu*^{2}}. This reading is not only compatible with other sources of evidence for the pronunciation of [Tangut]$_{3323}$ but is also paleographically more convincing.

The single most diagnostic difference between [Tangut]$_{2777}$ *ɴe*^{ʁ}*wr*^{1} {*ŋewr*^{1}} and [Tangut]$_{2776}$ *ŋgu*^{2} {*gjuu*^{2}} is between the upper portions of the middle components: between [component] and [component]. As a matter of fact, the contrast between [component] and [component] is among the most robust in the notoriously chaotic semi-cursive and cursive styles of hand-

44 Hán, Tóngyīn Wénhǎibǎoyùn hébiān *zhěnglǐ yǔ yánjiū*, 144.

written Tangut. In the following examples, semi-cursive characters are taken from the manuscript version of *the Art of War*[45]; cursive characters are taken from Sūn Yǐngxīn's study of the *Eight Upavāsa Precepts*. (2015)[46].

- [illegible] is usually written with a joined ∠-shape, cf. semi-cursive [illegible] for [illegible]$_{3830}$, [illegible] for [illegible]$_{0648}$; cursive [illegible] for [illegible]$_{2559}$, [illegible] for [illegible]$_{2640}$.
- [illegible], on the other hand, always conserves its regular shape with the left dot often independent, cf. semi-cursive [illegible] for [illegible]$_{5751}$, [illegible] for [illegible]$_{4797}$, cursive [illegible] for [illegible]$_{5215}$, [illegible] for [illegible]$_{3452}$.
- Characters like cursive [illegible] for [illegible]$_{5388}$ tellingly contrast both graphical elements within a single character.

The initial speller [illegible] contains the characteristic ∠-shape indicative of the [illegible] component, and hence should be read as [illegible]$_{2776}$ *ŋgu²* {*gjuu²*} instead of [illegible]$_{2777}$ *ɴeʁwr¹* {*ŋewr¹*}. With this revision of the initial speller, the reading of the character [illegible]$_{3323}$ should be revised from *ɴɴʕaʁ²* {*ŋiaa²*} to *ɴɢʕaʁ²* {*giaa²*}.

This proposed revision can be generalized into the following conjecture: whenever a Tangut syllable in a column-2 rhyme is used to transcribe a Chinese syllable with an EMC nasal initial, the syllable is likely to have a voiced stop initial rather than a nasal initial.[47]

4.2.3 Transcription Characters with Apparent Nasal Preinitials Before Zero Initial and *x*-

Staying on the subject of R.23 (2.20), I conclude this section by discussing the characters [illegible]$_{4623}$ *n·ʕaʁ* {*·iaa²*}, [illegible]$_{0871}$ *n·ʕaʁ* {*·iaa²*}, and [illegible]$_{2856}$ *nχʕaʁ* {*xiaa²*}, which constitute the only reliable examples of column-2 syllables with the initials *·*- and *x*-/*χ*-. There are no reliable examples that start with *ɣ*-/*ʁ*-. All these characters are special characters presumably created for the purpose of transcribing Sanskrit. Their *fǎnqiè* spelling and assumed Sanskrit target of transcription are as follows:

45 Inv. N° 775, cf. Sūn, Yǐngxīn, "Xīxià yìběn *Sūnzǐ Zhuàn* kǎobǔ 西夏译本《孙子传》考补" [Further Remarks on the Tangut Translation of the *Biography of Sun Tzǔ*], *Xixia yanjiu* 6 (2010): 70–74.

46 Sūn, Yǐngxīn, "Xīxià xiěběn *Jìnzhù bā zhāijiè wén* cǎoshū guīlǜ chūtàn" [A Preliminary Investigation of Regular Features of the Cursive Writing in the Tangut Manuscript Version of the *Eight Upavāsa Precepts*]. *Ningxia shehui kexue* 188 (2015): 124–34.

47 Note that this hypothesis does not generalize to the opposite Tangut-to-Chinese direction: in the *Pearl in the Hand*, Chinese syllables with EMC nasal initials happily transcribe both nasal and voiced stop initials in Tangut.

- [Tangut]$_{0871}$ *n·ʕa*^{ʁ2} {*·iaa*^{2}} is spelt [Tangut]$_{4330}$ *·ị*^{1} {*·jị*^{1}} + [Tangut]$_{2512}$ *ntśʕa*^{ʁ2} {*tśiaa*^{2}}. It likely used to denote the Sanskrit syllable *āḥ*, as it is glossed in the *Wénhǎi* as [Tangut] "one of the four major seed syllables (*bīja*, 種子字)" (Inv N° 211 212 213:100–108).
- [Tangut]$_{4623}$ *n·ʕa*^{ʁ2} {*·iaa*^{2}} is spelt as [Tangut]$_{0434}$ *·i*^{1} {*·ji*^{1}} + [Tangut]$_{0871}$ *n·ʕa*^{ʁ2} {*·iaa*^{2}}. It refers transparently to the Sanskrit syllable *ā*, as it is graphically derived from [Tangut]$_{4541}$ *·a?* "Sanskrit syllable *a*" and [Tangut]$_{0443}$ *dźo*^{1} {*dźjo*^{1}} "long."
- [Tangut]$_{2856}$ *nχʕa*^{ʁ2} {*xiaa*^{2}} has no surviving *fǎnqiè* spelling, but it is a *fǎnqiè* character made up of [Tangut]$_{3808}$ *xu*^{1} {*xju*^{1}} + [Tangut]$_{0871}$ *n·ʕa*^{ʁ2}{*·iaa*^{2}}. I am not aware of any Sanskrit syllables actually transliterated with this character, but it would probably denote Sanskrit *hā*.

In the framework of the nasal preinitial hypothesis, *n·ʕa*^{ʁ2} as a transcription of *ā* or *āḥ* would be quite unnatural, as would *nχʕa*^{ʁ2} as a transcription of *hā*. Equally unnatural is the fact that these characters are the only examples featuring the guttural initials with the nasal preinitial *n·*-, *nɣ-*/*nʁ-*, and *nx-*/*nχ-*. Given that there are no syllables reconstructed with *·ʕa*^{ʁ2} {*·ia*^{2}} and *χʕa*^{ʁ2} {*xia*^{2}}, both problems can be eliminated by removing the nasal preinitial, i.e., revising the pronunciation of [Tangut]$_{0871}$ and [Tangut]$_{4623}$ from *n·ʕa*^{ʁ2} {*·iaa*^{2}} to *·ʕa*^{ʁ2} {*·ia*^{2}}, and that of [Tangut]$_{2856}$ from *nχʕa*^{ʁ2} {*xiaa*^{2}} to *χʕa*^{ʁ2} {*xia*^{2}}.[48]

As a concluding remark, this revision does complicate the picture of the treatment of unmatched *fǎnqiè* spellings with regard to the nasal preinitial:

- [Tangut]$_{4902}$ *wu*^{ʁ1} {*ŋwuu*^{1}} "speech" is spelt [Tangut]$_{0226}$ *ɴwu*^{ʁ1}{*ŋwu*^{1}} + [Tangut]$_{5625}$ *nthwu*^{ʁ1}{*thwuu*^{1}}. This syllable, by virtue of belonging to rhyme 1.5 (R.5) and contrasting phonologically with [Tangut]$_{0226}$ *ɴwu*^{ʁ1}{*ŋwu*^{1}} in the native dictionaries, does indeed have a nasal preinitial. In this case, the conflict between the initial speller without a nasal preinitial *ɴwu*^{ʁ1} and the rhyme speller with a nasal preinitial *nthwu*^{ʁ1} is resolved in preference of the latter.
- In the case of [Tangut]$_{0871}$ *n·ʕa*^{ʁ2} → *·ʕa*^{ʁ2} {*·iaa*^{2} → *·ia*^{2}}, spelt [Tangut]$_{4330}$ *·ị*^{1} {*·jị*^{1}} + [Tangut]$_{2512}$ *ntśʕa*^{ʁ2} {*tśiaa*^{2}}, systematic and transcriptional considerations discussed above suggest the absence of the nasal preinitial. Thus, the conflict be-

48 One anonymous reviewer raised the question whether the fact that these syllables all seem to transcribe Sanskrit long vowels could not support Gong Hwang-cherng's hypothesis of vowel length instead. I consider this an unlikely possibility. [Tangut]$_{4623}$ *n·ʕa*^{ʁ2} {*·iaa*^{2}} "Skt. *ā*" graphically containing the component [Tangut] "long", forms a pattern with other Sanskrit transcription characters with an analogous graphical formation. However, none of the other instances, for example, [Tangut]$_{1544}$ {*·ji*^{2}} "Skt. *ī*," a column-1 syllable, and [Tangut]$_{1540}$ {*·wụ*^{1}} "Skt. *ū*," a tense syllable having no paired columns distinction—belongs to a column-2 rhyme.

tween the initial speller *·i*[1] and the rhyme speller *ntśʃa*[ʁ2] is resolved in favor of the initial speller.

This lack of consistency, while worrisome, should be maintained nonetheless given the strong rationales for both treatments in the cases mentioned. There is a serious need for a wholesale reinvestigation of Tangut *fǎnqiè* behavior addressing the problems raised in this section.

4.3 Comparative Problems

Jacques' *Esquisse de phonologie et de morphologie historique du tangoute*, our primary authority on Tangut etymology, features no discussion regarding the origin of Tangut column-2 ("long-vowel") rhymes; nor did the subsequent literature address the issue. Revising the value of column-2 rhymes from vowel length to the presence of a nasal preinitial enables meaningful hypotheses to be postulated as to the origin of these rhymes. In this section, we discuss all the etymologies postulated in the *Esquisse* of Tangut words belonging to a column-2 rhyme in the major cycle,[49] as well as a few other cognates not proposed there. The comparison is made, as usual, mostly against modern Rgyalrongic languages, especially Japhug and occasionally its fellow Upper Rgyalrong languages Tshobdun and Zbu, as well as modern West Rgyalrongic languages, Khroskyabs, and an assortment of Stau-Horpa lects, likely the closest relatives to Tangut.[50,51] The Pre-Tangut and Proto-Rgyalrong(ic) forms are provisional and liable to further changes.

---

[49] However, the tentative comparison given by Jacques in the *Esquisse* between 𗍫$_{2047}$ *mmi*[1] {*mjii*[1]} "to give" and Japhug *-mbi* "id." judged as "pas certaine" (p. 37) and "très problématique" (p. 97) on account of its *sui generis* correspondence between Tangut *m-* and Japhug *mb-*, is not discussed. One anonymous reviewer suggested that 𗍫$_{2047}$ *mmi*[1] {*mjii*[1]} could be related to Geshiza *mə* "to feed" instead.

[50] Lai et al., "Tangut as a West Gyalrongic language."

[51] Japhug data is cited from the *Esquisse* and checked against the latest version of Jacques' dictionary (*Dictionnaire Japhug-chinois-français, Version 1.1.*). Zbu Rgyalrong data are cited from Gong's *Le rgyalrong zbu, une langue tibéto-birmane de Chine du Sud-ouest : une étude descriptive, typologique et comparative*. Khroskyabs data are cited from Lai's *Grammaire du khroskyabs de Wobzi.* "Stau" data, referring to the Stau dialect of Khang·gsar, are kindly provided by Guillaume Jacques, Lai Yunfan, Anton Antonov, and Lobsang Nyima (cf. the authors 2017). "Geshiza" data, referring to Eastern Geshiza of Balang, are cited from Honkasalo's *A grammar of Eastern Geshiza: A culturally anchored description.* Other Stau-Horpa lects, as well as a few forms in Zbu and Tshobdun, are cited from the *rGyalrongic Languages Database*, ed. Yasuhiko Nagano and Mariëlle Prins. Entries of the *Database* are annotated with the locality, the four-letter locality code, and the numerical entry code. Whenever there is a retranscription, the original form is also left in parentheses: Mda·mdo *kājə* (*kan'jə*, DB-dand-1993).

4.3.1 Root Dental Preinitial *n-* in Rgyalrongic Languages

Several Tangut words with a nasal preinitial show good evidence for a nasal preinitial **n-* in Pre-Tangut.

- 𗏁$_{0716}$ *nśi*¹ {*śjii*¹} "to butcher" is a cognate[52] of Japhug *-ntɕha* "to butcher," from proto-Rgyalrong **-nəɕa*. Some West Rgyalrongic languages preserve a form close to the revised Tangut reconstruction, such as Khroskyabs *ɲɕî*.
- 𗟲$_{0320}$ *n·wəʁ¹* {*·wəə*¹} "soft" is a cognate of Japhug *-mpɯ* from proto-Rgyalrong **-nəpuʁ*. Some West Rgyalrongic languages preserve a form close to the revised Tangut reconstruction, such as Khroskyabs *nvə̂* and Stau-Horpa forms such as Geshiza *nvə*, Tag·gsum *ŋvə* (DB-dasa-1714).
- 𗢭$_{3113}$ *ŋgə*¹ {*gjɨɨ*¹} "nine" is a cognate of Japhug *kɯngɯt*. While all modern Rgyalrongic evidence points to a nasal preinitial **n-*, Sino-Tibetan comparanda such as Tibetan *dgu* suggest a Pre-Proto-Rgyalrong form ***-təⁿgu*. Modern West Rgyalrongic languages, such as Khroskyabs *ŋgə́*, Geshiza *ŋgæ*, share the place assimilation hypothesized for Tangut.
- 𗗙$_{5522}$ *nle*² {*ljiij*²} "to wait" is a cognate of Japhug *-nɤjo*, Zbu *-nɐⁿdjê*, from a proto-Rgyalrong root akin to **-nəlaŋ*. Khroskyabs *njé* similarly preserve the nasal preinitial.

The *n-* preinitials in two cases are not of obvious Pan-Rgyalrongic pedigree but must be reconstructed using evidence specifically from modern West Rgyalrongic languages.

- 𗘍$_{2737}$ *nlə*¹ {*ljɨɨ*¹} "heavy" is usually reflected in modern Rgyalrongic with a preinitial *r-*: Japhug *-rʐi*,[53] Khroskyabs *rdə̂*. However, a nasal preinitial is supported by Stau-Horpa forms Mda·mdo *kãjə* (*kan"jə*, DB-dand-1993) and Khang·gsar *ˈndərə* (DB-kong-1993).
- 𗗚$_{2621}$ *nse*² {*sjiij*²} "to think" is analyzed by Jacques as cognate to Japhug *-sɯso*.[54] However, Khroskyabs *ntshə̂*, Stau and Geshiza *ntshə* seem to present a better candidate for cognacy.

---

52 Gong, *Le rgyalrong zbu, une langue tibéto-birmane de Chine du Sud-ouest*, 303–9.

53 An alternative Tangut word for "heavy, weight," 𗒘$_{0902}$ *źər*¹ {*źjɨr*¹}, leaves the slim but enticing possibility of a separate Proto-Rgyalrongic root shared by Japhug *-rʐi*.

54 Jacques, *Esquisse de phonologie et de morphologie historique du tangoute*, 180.

4.3.2 Autobenefactive *nə-

In one case, the Tangut nasal preinitial clearly derives from the Rgyalrongic autobenefactive prefix **nə-*.[55]

- 敍$_{4040}$ *ɴqhʕu*$^{\text{ʁ1}}$ {*khjuu*^{1}} "to greet"[56] is a cognate of Japhug *-qru*. Zbu and West Rgyalrongic uniformly prefer a form with the autobenefactive **nə-*: Zbu *-nəqhrə̂*, Khroskyabs *ŋk*^{h}*rûʁæ*, Stau *ɴq*^{h}*rə*.

Three other verbs have an unexplained nasal preinitial, which probably also reflects the Rgyalrongic autobenefactive prefix **nə-*. It is worth noting, however, that none of these etymologies seems particularly solid.

- 絁$_{2621}$ *nse*^{2} {*sjiij*^{2}} "to think" is analyzed by Jacques as cognate to Japhug *-sɯso*.[57] While in §4.3.1 I propose that it could also be analyzed as cognate to Khroskyabs *ntshə̂*, retaining Jacques's etymology would suggest that this word is an example of autobenefactive **nə-*.
- 豵$_{0369}$ *nthu*^{1} {*thjuu*^{1}} "to inspect" is judged by Jacques[58] to be "potentiellement [...] rapproché" to Japhug *-thu* "to ask."
- 㝊$_{5612}$ *ntshe*^{1} {*tshjiij*^{1}} "to speak" is judged by Jacques[59] to "potentiellement se comparer" to Japhug *-ti*. The correspondence, while rather poor, cannot be entirely ruled out, as for example Zbu B dialects have comparable forms such as Go·la·thang *ké-tse* (*ka'tse*, DB-gele-0904).

4.3.3 Stative **ŋa-* Before a Pre-Tangut Acute Prenasalized Voiced Initial

Two instances of the nasal preinitial, both with Tangut voiced stop initials < Pre-Tangut prenasalized voiced initials, derive from the stative prefix **ŋa-*, reflected as Japhug *a-*.

- 䋼$_{5149}$ *ndu*$^{\text{ʁ1}}$ {*duu*^{1}} "to accumulate" < Pre-Tangut **ŋ-*^{n}*du*$^{\text{ʁ}}$ is a cognate of Japhug *-ajtɯ* or alternatively *-ndɯ*, from Proto-Rgyalrong **ŋa-ləntuŋ*.
- 𢇍$_{2396}$ *ndzu*$^{\text{ʁ2}}${*dzuu*^{2}} "to sit" < Pre-Tangut **ŋ-*^{n}*dzu*$^{\text{ʁ}}$ is a cognate of Japhug *-amdzɯ*, Zbu *´-amdzo*$^{\text{ɣ}}$ "id.," from Proto-Rgyalrong **ŋa-mə*^{n}*dzuŋ*. This verb

55 Guillaume Jacques, "The Spontaneous-Autobenefactive Prefix in Japhug Rgyalrong," *Linguistics of the Tibeto-Burman Area* 38, no. 2 (2015): 271–91.
56 Concerning the rhyme of this word R.6 (1.6), cf. Miyake, "Complexity from Compression."
57 Jacques, *Esquisse de phonologie et de morphologie historique du tangoute,* 180.
58 Ibid., 50.
59 Jacques, "The Spontaneous-Autobenefactive Prefix in Japhug Rgyalrong," 170.

has a cognate in Stau-Horpa: Stau *ndzɔ*, Geshiza *ndzo*, which agrees with the Tangut form.

As a matter of fact, the stative **ŋa-* does not induce the Tangut nasal preinitial before other initial types. Japhug *-astu* "straight," for example, corresponds to 祋$_{1569}$ *twụᵏ¹* {*twụ¹*} "straight (esp. morally)." The verb 缀$_{2226}$ *wiᵏ²* {*we²*} "to become," in an etymology not discussed by Jacques in the *Esquisse*, is a passive formation derived from 䊷$_{5113}$ *wi¹* {*wji¹*} "to do," parallel to Japhug *-apa* "to become," derived from *-pa* "to do." The attested form is *wiᵏ²* {*we²*}, not †*n.wiᵏ²* {*wee²*}. In addition, in the case of 祋$_{1638}$ *gi¹* {*gji¹*} "clear (water)" < Pre-Tangut **ŋa-ⁿgri*, cf. Japhug *-amgri*, Zbu *-ɐmgréɣ*, Khroskyabs *ʁgré*, the same **ŋa-* fails to induce the Tangut nasal preinitial even before a voiced (< *prenasalized) stop.

Thus, an intermediate **n-* must be postulated, the outcome of **ŋ-* assimilated to the following acute prenasalized initial. Before a grave initial, such as in the case of 祋$_{1638}$ *gi¹* {*gji¹*}, the result **ŋg-* was absorbed into the ordinary series of Pre-Tangut prenasalized stops, producing the attested form *g-*. Alternatively, as one anonymous reviewer suggests, it could reflect the autobenefactive prefix *nə-*, so that both Tangut 缀$_{2396}$ *ndzuᵏ²*{*dzuu²*} "to sit" and Stau *ndzɔ*, etc. would correspond to Japhug *-n-ɤmdzɯ* "to sit by/for oneself" instead of unprefixed *-amdzɯ*.

4.3.4 Geminate Nasal Assimilation

The examples discussed in §4.3.1–3 can all be traced, in one way or another, to a Pre-Tangut preinitial **n-*. A large number of examples, however, correspond to a wide range of preinitials in modern Rgyalrongic comparanda. Their only commonality is that they are followed by a nasal initial. In these cases, a rather atypical assimilation, whereby any preinitial is assimilated toward a geminate nasal, i.e., *mm-* < **lm-*, **rm-*, **sm-*, **km-* ... must be postulated.

In two examples, the geminate nasal arises from an earlier preinitial **l-*:

- 獱$_{5677}$ *mmeᵉ¹* {*mjiij¹*} "tail" is a cognate of Japhug *tɤ-jme*, Zbu *tɐ-lméʔ*, from Proto-Rgyalrong **-ləmɛ*.
- 豳$_{0330}$ *mmeᵉ¹* {*mjiij¹*} "dream" is a cognate of Japhug *tɯ-jmŋo*, Zbu *ta-lmáʔ*, from Proto-Rgyalrong **-ləmaŋ*.

However, earlier **l-* before nasals also give reflexes as tense syllables. Known examples are 绷$_{2325}$ *mạ²* {*mjɨ²*} "to forget," cf. Japhug *-jmɯt*; 躔$_{4600}$ *ɴwụᵏ¹* {*ŋwụ¹*} "oath," cf. Japhug *kɯjŋu*.

In two examples, the geminate nasal arises from an earlier preinitial *r-:

- 緲$_{2639}$ *mme²* {*mjiij²*} "name" is a cognate of Japhug *tɤ-rmi*.
- 廁$_{1892}$ *mmi¹* {*mjii¹*} "house" is probably a cognate of Japhug *-rma* "to pass a night in someone's home."

However, earlier *r- before nasals also give reflexes as plain syllables. Some known examples are 尼$_{1671}$ *ne¹* {*njij¹*} "red," cf. Japhug *-ɣɯrni*, 怉$_{2563}$ *me²* {*mej²*} "hair," Japhug *tɤ-rme*.

In three examples, the geminate nasal arises from an earlier preinitial *s-.

- 徲$_{5700}$ *nni²* {*njii²*} "nose" is a cognate of Japhug *tɯ-ɕna*.
- 姩$_{2518}$ *nne¹* {*njiij¹*} "heart" is a cognate of Japhug *tɯ-sni*.
- Jacques did not discuss the Burmo-Qiangic etymology of 娳$_{2440}$ *nnə²* {*njɨɨ²*} "day."[60] However, it is superposable to Khroskyabs *ânsnə* "a day," Geshiza *bə-sni* "today," etc. Japhug *sɲi* and Tshobdun *sɲi*, both "day," are also clearly cognate, though slightly irregular.

However, earlier *s- before nasals also give reflexes as tense syllables. Known examples are 繈$_{5731}$ *nạᵏ¹* {*nạ¹*} "nasal mucus," cf. Japhug *tɯ-ɕnaβ*; 毅$_{0549}$ *nʃọᵏ¹* {*niọ¹*} "sister of a woman," cf. Japhug *tɤ-snom*; 豬$_{5990}$ *nʃọᵏ²* {*niọ²*} "ear (of grain)," cf. Japhug *kɯɛnom*.

In four examples, the geminate nasal arises from an earlier preinitial *ɣ- or *ʁ-, probably passing through an intermediate stage as *ŋ-.

- 徲$_{5702}$ *mma¹* {*mjaa¹*} "ulcer, wound" is a cognate of Japhug *tɯ-ɣmaz* "blessure," Proto-Rgyalrong *-*kəmas*.
- 褲$_{4027}$ *nnə¹* {*njɨɨ¹*} "two" is a cognate of Japhug *ʁnɯz*, Proto-Upper-Rgyalrong **qənes*.
- 蘘$_{4408}$ *mməᵏ¹* {*məə¹*} "fire" is a cognate of Japhug *smi*. The Tangut form itself corresponds with a West Rgyalrongic dialectal root with *ʁ-: Khroskyabs *ʁmə́*, Stau *ɣmə*, Geshiza *wmə*, which induced the uvularity compression *$ʁmə^{ᵏ}$ < *ʁ-mə.[61]
- Jacques did not discuss the Burmo-Qiangic etymology of 紇$_{2192}$ *mme¹* {*mjiij¹*} "corpse." However, it is cognate to Khroskyabs *jmô*, thus reflecting a proto-form akin to Proto-Rgyalrong **kəmaŋ*. This root is clearly ancient,

[60] Jacques, *Dictionnaire Japhug-chinois-français, Version 1.1*, 161.
[61] Gong, "Uvulars and Uvularization in Tangut Phonology," 198–99.

cf. Proto-Lolo-Burmese (Matisoff) **maŋ*, etc. Although Matisoff[62] considers the Proto-Sino-Tibetan form to have a preinitial **s*-, it can be argued that the Khroskyabs preinitial *j*- < **kə*- is earlier. Matisoff's **s*- forms could instead be regarded as a later composition with the pan-Sino-Tibetan verb "to die," Chinese 死 *sijX*, etc., cf. Lǒngchuān Ngochang *ʂɿ˥˥mʐuaŋ˥˥* "corpse".[63]

No exceptions to this correspondence are known. The presence of uvularity compression in 𗈪$_{4408}$ *mmə^{ʁ1}* {*məə^{1}*} "fire" but the lack of it in 𗓁$_{5702}$ *mma^{1}* {*mjaa^{1}*} "ulcer, wound" and 𗍫$_{4027}$ *nnə^{1}* {*njɨɨ^{1}*} "two" should nevertheless be noted.

Finally, in two examples the extant modern Rgyalrongic comparanda do not permit the identification of the Pre-Tangut preinitial in question.

- 𗎫$_{2436}$ *mma^{1}* {*mjaa^{1}*} "fruit" is a cognate of Japhug *ɯ-mat* "its fruit." Among modern Rgyalrongic languages, this etymon is only attested in Upper Rgyalrong: Tshobdun *té-mɛ* (*ta^{55}mᴇ^{33}*, DB-caob-0318), Zbu *və-mêt*, always without a preinitial.
- 𗊖$_{2128}$ *mmə^{ʁ1}* {*məə^{1}*} "to blow" is a cognate of Japhug *kɤ-ɣɤmɯt*. This verb is a deverbal from a noun reflected as Japhug *tɤmɯt* "exhaled breath." However, the deverbal formant is the **p*- one in Upper Rgyalrong: Japhug *kɤ-ɣɤmɯt*, Tshobdun *kɐwémo* (*ka^{33}wa^{44}mo^{33}*, DB-caob-1316), Zbu *ka-vamô^{ɣ}t*. In Geshiza, a West Rgyalrongic language like Tangut, one finds *wmə* < **ɣmə* instead. It is difficult to tell if the Tangut verb reflects the **p*- deverbal in Upper Rgyalrong of the **k*- deverbal in Geshiza.

As a conclusion, the geminate nasal in Tangut unambiguously indicates the existence of a preinitial in Proto-Rgyalrongic. However, the same Proto-Rgyalrongic preinitial before a nasal initial can lead either to a geminate nasal or a different result. There is no obvious solution to this problem. For **l*- and **s*-, nonetheless, one might tentatively suggest a preference for the geminate reflex in open syllables (and quasi-open syllables with *-*ŋ*), and a preference for tense reflex in close syllables. Compare, for example, 𗓀$_{5700}$ *nni^{2}* {*njii^{2}*} "nose," which has a geminate nasal and derives from earlier *-*a*, cf. Japhug *tɯ-ɕna*, with 𗖵$_{5731}$ *nạ^{ʁ1}* {*nạ^{1}*} "nasal mucus," which has a tense vowel and derives from earlier *-*ap*, cf. Japhug *tɯ-ɕnaβ*.

---

62 James A. Matisoff, *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction* (Berkeley: University of California Press, 2003), 265.

63 Huáng Bùfán, Xǔ Shòuchūn, Chén Jiāyīng, Wáng Huìyín eds., *Zàngmiǎnyǔzú yǔyán cíhuì* [A Tibeto-Burman Lexicon] (Beijing: Zhongyang minzu daxue chubanshe, 1992), 54.

4.3.5 Unexplained Comparison

One remaining comparison needs to be discussed. [Tangut]$_{2801}$ *nlhu*^{ʁ2} {*lhuu*^{2}} “marrow” is considered by Jacques in the *Esquisse* (p. 53) to be cognate to Japhug *tɯ-pju*. If the words are indeed cognate, the most likely proto-form would be approximatively **-məl̥u*. This correspondence between Tangut *nlh-* and Modern Rgyalrongic *pj-* remains unattested elsewhere.

4.3.6 Cases Not Discussed in the *Esquisse*

One etymology not proposed by Jacques in the *Esquisse* merits some consideration.

- The stem-alternating verb [Tangut]$_{0750}$ *ŋgi*^{1} {*gjii*^{1}} / [Tangut]$_{1249}$ *ŋgo*^{1} {*gjoo*^{1}} “to chew, to bite” is considered by Gong Hwang-cherng[64] to be borrowed from Chinese 齧 *ŋet* “to bite, to gnaw.” However, I argue that shows that it reflects in fact a pan-Rgyalrongic etymon **-nəka*, cf. Japhug *kɤ-nɤŋka*, the semantics of which have been bleached to “to eat” in Modern Stau-Horpa, cf. Stau and Geshizha *ŋgə*.[65]

The discovery that column-2 rhymes could reflect the Rgyalrongic autobenefactive derivation also allows us to understand the origin of the verb [Tangut]$_{4489}$ *mphi*^{1} {*phjii*^{1}}. It is clearly related to the stem-alternating verb [Tangut]$_{0749}$ *phi*^{1} {*phji*^{1}} / [Tangut]$_{4568}$ *pho*^{2} {*phjo*^{2}}, which is used as an unmarked causativiser “to make, to order.” The most common meaning for the form with nasal preinitial [Tangut]$_{4489}$ *mphi*^{1} {*phjii*^{1}} is a more specific one, “to send someone as representative,” cf. (1). It is interesting to note that a common alternative verb in the same context, [Tangut]$_{5871}$ *nziw*^{ʁ2} {*zeew*^{2}}, also has a nasal preinitial.

(1)	[Tangut]	[Tangut]	[Tangut]	[Tangut]	[Tangut]	[Tangut]	[Tangut]
	dzwo^{2}	*mphi*^{1}	*tshi*^{1}	*ne*^{2}	*do*^{ʁ2}	*ntshe*^{1}	*phi*^{1}
	person	send	Qí	king	POST	speak	make

“He sent someone to the King of Qí to tell (the story).”[66]

The other frequent meaning of [Tangut]$_{4489}$ *mphi*^{1} is “to employ someone as a ser-

[64] Gong, “Xīxiàyǔ zhōng de Hànyǔ jièci.”

[65] Gong, *Le rgyalrong zbu, une langue tibéto-birmane de Chine du Sud-ouest*, 302–3.

[66] Inv N° 616:7, cf. Jacques, *Textes tangoutes I. «Nouveau recueil sur l amour parental et la piété filiale»*, 42.

vant," cf. (2). A common word for the job of "servant," indeed, is the *nomen patientis* of this verb, [illegible] *mphi*^{1}-*le*^{ʁ}*w*^{2} {*phjii*^{1}-*lew*^{2}}.

(2)

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
thə^{2}	*ta*^{1}	*mo*^{2}	*dzẽ*^{1}	*dźo*^{1}	*rər*^{2}-*mphi*^{1}.*qʕe*^{ʁ2}				*ni*^{2}	*mphi*^{1}	*mə*^{1}-*wo̜*^{2}-*na*^{2}	
this	TOP	I	time	long	PFV-order			you	order	NEG-can-2SG		

"I had these people as servants (*lit.* ordered these people) for a long time; you'll not be able to work with them (*lit.* order them)."[67]

Both uses of the preinitialed verb [illegible]$_{4489}$ *mphi*^{1} {*phjii*^{1}} can be understood as the effect of an autobenefactive prefix. Compared to the unprefixed [illegible]$_{0749}$ *phi*^{1} {*phji*^{1}} / [illegible]$_{4568}$ *pho*^{2} {*phjo*^{2}}, which has a general causative meaning of "to make, to order," both common meanings [illegible]$_{4489}$ *mphi*^{1} {*phjii*^{1}}, whether "to send someone as a representative"—to represent oneself—or "to engage someone as a servant"—i.e., in one's own service—strongly imply that the subject of the verb is a beneficiary, and thus can be regarded as autobenefactive derivations from the unprefixed base verb.

Another potential example of the autobenefactive prefix is [illegible]$_{5435}$ *n.wi*^{ʁ1} {*wee*^{1}} "to be born." This word could be an autobenefactive derivation from [illegible]$_{2226}$ *wi*^{ʁ2} {*we*^{2}} "to become" (for its etymology cf. §4.3.3), parallel to Geshiza *nʑæ* "to be born," probably autobenefactive from *ʑæ* "to come."

4.3.7 The Origins of the Tangut Nasal Preinitial

In conclusion, the Tangut nasal preinitial seems to have two principal origins:

- Before nasals, it reflects the result of the assimilation of any Pre-Tangut preinitial to a geminate nasal: NN < CN;
- Before other consonants, it reflects a Pre-Tangut dental preinitial *n*-.

The nasal preinitial hypothesis settles some etymological problems and opens up fruitful possibilities for further etymological research, especially with regard to Tangut reflexes of the pan-Rgyalrongic autobenefactive derivation. On the other hand, the evolution of Proto-Rgyalrongic preinitials in Tangut[68],

---

67 Inv N° 616:5, cf. ibid., 29.

68 Gong Hwang-cherng, "Xīxiàyǔ de jǐnyuányīn jí qí qǐyuán" [The Tense Vowels in Tangut and Their Origins], *Bulletin of the Institute of History and Philology* 70, no. 2 (1999): 531–58. Miyake, "Complexity from Compression: A Sketch of Pre-Tangut"; Jacques, *Esquisse de phonologie et de morphologie historique du tangoute*, 21–35; Gong, "Uvulars and Uvularization in Tangut Phonology,"198–99.

which already has a rather chaotic picture, is further complicated by the non-obligatory geminate nasal assimilation. Further research is needed to elucidate the specific conditions of the nasal assimilation, preferably based on further etymological proposals.

5. CONCLUSION

This essay proposes the nasal preinitial hypothesis, recapitulated in Table 5, according to which syllables having a column-2 rhyme have a nasal preinitial instead of a long vowel as Gong Hwang-cherng proposed, whereas column-1 syllables do not.

Table 5: The nasal preinitial hypothesis.

<table>
<tr><td rowspan="2">initial classes</td><td colspan="2" rowspan="2">Column-1 rhyme
short vowel → absence of nasal preinitial</td><td colspan="3">Column-2 rhyme
long vowel → presence of nasal preinitial</td></tr>
<tr><td>Tangut character</td><td>revision</td><td>Note</td></tr>
<tr><td rowspan="2">voiced</td><td>[illegible]$_{3806}${bu^{1}} "cattail"</td><td>bu^{ʁ1} → bu^{ʁ1}</td><td>[illegible]$_{2138}$ {buu^{2}} "tomb"</td><td>buu^{ʁ2} → mbu^{ʁ2}</td><td>borrowed from Chinese 墓 Héxī *mb- < muH</td></tr>
<tr><td colspan="2">no revision</td><td colspan="3">homorganic nasal preinitial: mb-, nd-, ŋg-, ɴɢ-, ndz-, ndź-</td></tr>
<tr><td rowspan="2">voiceless</td><td>[illegible]$_{4568}$ {phjo^{2}} "to make"</td><td>pho^{2} → pho^{2}</td><td>[illegible]$_{1223}${phjoo^{2}} "to combine"</td><td>phoo^{2} → mpho^{2}</td><td>Tibetan transcription as འཕོ <'pho>.</td></tr>
<tr><td colspan="2">no revision</td><td colspan="3">homorganic nasal preinitial: mp(h)-, nt(h)-, ŋk(h)-, ɴq(h)-, nts(h)-, ntś(h)-</td></tr>
<tr><td rowspan="2">nasals</td><td>[illegible]$_{0092}$ {mja^{1}} "mother"</td><td>ma^{1} → ma^{1}</td><td>[illegible]$_{5702}$ {mjaa^{1}} "ulcer, wound"</td><td>maa^{1} → mma^{1}</td><td>< *ŋma, cognate to Japhug tɯ-ɣmaz</td></tr>
<tr><td colspan="2">no revision</td><td colspan="3">geminate nasal preinitial: mm-, nn-, ŋŋ-, ɴɴ-</td></tr>
<tr><td rowspan="2">sibilants</td><td>[illegible]$_{2104}$ {śji^{1}} "before"</td><td>śi^{1} → śi^{1}</td><td>[illegible]$_{0716}$ nśi^{1} {śjii^{1}} "to butcher"</td><td>śjii^{1} → nśi^{1}</td><td>cognate to Khroskyabs ɲɕî</td></tr>
<tr><td colspan="2">no revision</td><td colspan="3">dental nasal preinitial: ns-, nś-, nz-, nź-</td></tr>
<tr><td rowspan="2">glides</td><td>[illegible]$_{4957}$ {wə^{1}} "fur jacket"</td><td>wə^{ʁ1} → wə^{ʁ1}</td><td>[illegible]$_{0320}$ {·wəə^{1}} "soft"</td><td>·wəə^{ʁ1} → n·wə^{ʁ1}</td><td>cognate to Geshizha nvə</td></tr>
<tr><td colspan="2">no revision</td><td colspan="3">dental nasal preinitial: n·w-, n·j-</td></tr>
<tr><td>gutturals</td><td>[illegible]$_{5689}$ {ɣa^{1}} "door"</td><td>ʁa^{ʁ1} → ʁa^{ʁ1}</td><td colspan="3">nx-/nχ-, nɣ-/nʁ- and n·- do not exist (cf. §4.2.3)</td></tr>
</table>

The "long vowel" problem has remained an open question in Tangut scholarship ever since Gong Hwang-cherng's article "A Hypothesis of Three Grades and Vowel Length Distinction in Tangut". It is a testimony to the power of Sino-Tangutica in the reconstruction of Tangut phonology that a definitive solution to this problem can only come from a thorough examination of Chinese-to-Tangut materials, especially from comparing different chronological layers of borrowings.

The revision of [Tangut character]$_{3323}$ "goose" from NNʕ*a*^{ʁ2} {*ŋiaa*^{2}} to NGʕ*a*^{ʁ2} {*giaa*^{2}} (§4.2.2) draws attention to other potential misreadings of *fǎnqiè* spellings from the *Combined Edition of Wénhǎi and Homophones*. This work is crucial in the reconstruction of Tangut phonology, preserving the only testimony of the pronunciation of a large number of *shǎng*-toned characters. Its unique importance, unfortunately, is rivaled only by its paleographic difficulty. It is hoped that further insights into the rules and mechanism of Tangut *fǎnqiè* might bring even more emendations to the pronunciation of individual Tangut characters.

Additional research is called for primarily in two directions: a systematic treatment of Tangut *fǎnqiè* behavior, especially with regard to the less well-behaving nasals and Sanskrit transcription characters (§4.2), and an investigation of the specific conditions of preinitial assimilation before nasals (§4.3.4).

Among the rhymes assigned by Gong Hwang-cheng with a "long vowel," i.e., to column 2, the rhymes R.21 (1.21–2.18) and R.59 (1.57) are shown to be unrelated to the phenomenon discussed in this article. Their nature, as well as that of R.60 (2.50), will be addressed in forthcoming articles. The same remark applies to "long vowel" rhymes outside the major cycle. I consider "long vowel" rhymes in the second (R.80–R.98) and third minor cycles (R.99–R.103) to be unrelated to the "vowel length" distinction discussed in this article. In my opinion, only after a thorough revision of the major cycle could the reconstruction of the minor cycles be updated through Gong Hwang-cherng's method of phonological alternation.

REFERENCES

Arakawa, Shintarō 荒川慎太郎. "Seikago tsūin jiten 西夏語通韻字典" [A Rhyme Dictionary of Tangut]. *Gengogaku Kenkyū* 16 (1997): 1–153.

—. "Kazō taionshiryō kara mita Seikago no seichō 夏藏対音資料からみた西夏語の声調" [A Study on Tangut Tones from Tibetan Transcription Materials]. *Gengogaku Kenkyū* 17–18 (1999): 27–44.

—. "Seikagoon fukugen no tame no kakushu shiryō 西夏語音復元のための各種資料" [Sources for the Reconstruction of Tangut Phonology]. *Rekishi to chiri* 629 (2009): 27–35.

Baxter, William H. *A Handbook of Old Chinese Phonology.* Berlin: Mouton de Gruyter, 1992.

Gong, Hwang-cherng 龔煌城. "Voiced Obstruents in the Tangut Language." *Bulletin of the Institute of History and Philology* 52, no. 1 (1981): 1–16.

—. "Shíèr shìjì mò Hànyǔ de Xīběi fāngyīn (shēngmǔ bùfèn) 十二世紀末漢語的西北方音（聲母部分）" [A Northwestern Dialect of Chinese at the End of the 12th Century, Part 1: Initials]. *Bulletin of the Institute of History and Philology* 52, no. 1 (1981): 37–78.

—. "Xīxiàyǔ zhōng de Hànyǔ jièci 西夏語中的漢語借詞" [Chinese Loanwords in the Tangut Language]. *Bulletin of the Institute of History and Philology* 52, no. 4 (1981): 681–780.

—. "Phonological Alternations in Tangut." *Bulletin of the Institute of History and Philology* 59, no. 3 (1988): 783–834.

—. "The Phonological Reconstruction of Tangut Through Examination of Phonological Alternations." *Bulletin of the Institute of History and Philology* 60, no. 1 (1989): 1–45.

—. "Lèilín Xīxiàwén yìběn Hàn-Xià duìyīnzì yánjiù 類林西夏文譯本漢夏對音字研究" [A Study on the Tangut Transcription of Chinese in the Tangut Translation of the Lei-lin]. In *Kǎogǔ yǔ lìshǐ wénhuà (Qìngzhù Gāo Qùxún Xiānshēng Bāshí Dàshòu Lùnwénjí)* 考古與歷史文化（慶祝高去尋先生八十大壽論文集 [Anthropology and historical cultures: Festschrift on the Occasion of Kao Chü-hsün's 80th Birthday], edited by Wen-hsün Sung 宋文薰, vol 2: 185–223. Taipei: Cheng Chung, 1991.

—. "Xīxiàyǔ de yīnyùn zhuǎnhuàn yǔ gòucífǎ 西夏語的音韻轉換與構詞法" [Phonological Alternations and Derivational Morphology in Tangut]. *Bulletin of the Institute of History and Philology* 64, no. 4 (1993): 935–68.

—. "A Hypothesis of Three Grades and Vowel Length Distinction in Tangut." *Journal of Asian and African Studies* 46–47 (1994): 305–14.

—. "Xīxiàyǔ de jǐnyuányīn jí qí qǐyuán 西夏語的緊元音及其起源" [The Tense Vowels in Tangut and Their Origins]. *Bulletin of the Institute of History and Philology* 70, no. 2 (1999): 531–58.

—. *Xīxià yǔwén yánjiù lùnwénjí* 西夏語文研究論文集 [Collected Papers on Tangut Philology]. Taipei: Academia Sinica, 2002.

—. "Tangut". In Randy LaPolla and Graham Thurgood (eds.) *The Sino-Tibetan Languages*, 602–20. London & New York: Routledge, 2003.

Gong, Xun. *Le rgyalrong zbu, une langue tibéto-birmane de Chine du Sud-ouest: une étude descriptive, typologique et comparative*. PhD thesis, INALCO, 2018.

—. "Uvulars and Uvularization in Tangut Phonology." *Language and Linguistics* 21, no. 2 (2020): 175–212.

Hán, Xiǎománg 韩小忙. Tóngyīn Wénhǎibǎoyùn hébiān *zhěnglǐ yǔ yánjiū* 《同音文海宝韵合编》整理与研究 [Combined Edition of the Homophones and *Wénhǎi*: A Critical Edition with Extensive Commentary]. Beijing: Zhongguo shehui kexue chubanshe, 2008.

Honkasalo, Sami. *A Grammar of Eastern Geshiza: A Culturally Anchored Description.* PhD thesis, University of Helsinki, 2019.

Huáng, Bùfán 黄布凡 ed. *Zàngmiǎnyǔzú yǔyán cíhuì* 藏缅语族语言词汇 [A Lexicon of Tibeto-Burman Languages]. Beijing: Zhongyang minzu xueyuan chubanshe, 1992.

Institute of Oriental Manuscripts of the Russian Academy of Sciences, Institute of Ethnology and Anthropology of the Chinese Academy of Social Sciences, eds. *Écáng Hēishǔichéng wénxiàn* 俄藏黑水城文獻 / *Pamjatniki pis'mennosti iz Chara-choto chranjaščiesja v Rossii*, 29 vols. Shanghai: Shanghai guji chubanshe, 1996–2019.

Jacques, Guillaume. *Textes tangoutes I. «Nouveau recueil sur l amour parental et la piété filiale».* Munich: LINCOM, 2007.

—. *Esquisse de phonologie et de morphologie historique du tangoute*. Leiden: Global Oriental, 2014.

—. "The Spontaneous-Autobenefactive Prefix in Japhug Rgyalrong." *Linguistics of the Tibeto-Burman Area* 38, no. 2 (2015): 271–91.

—. *Dictionnaire Japhug-chinois-français, Version 1.1*. 2016. Accessed November 20, 2021. http://himalco.huma-num.fr/dictionaries/japhug/dictionary.pdf.

Jacques, Guillaume, Yunfan Lai, Anton Antonov, Lobsang Nima. "Stau." In *The Sino-Tibetan Languages (2nd edition)*, edited by Graham Thurgood and Randy LaPolla. London & New York: Routledge, 2017.

Lai, Yunfan. *Grammaire du khroskyabs de Wobzi*. PhD thesis, Université Paris 3, 2017.

Lai, Yunfan, Xun Gong, Jesse P. Gates and Guillaume Jacques. "Tangut as a West Gyalrongic Language." *Folia Linguistica Historica* 41, no. 1 (2020): 171–203.

Lǐ Fànwén 李範文. *Sòngdài Xīběi fāngyīn:* Fānhàn Héshí Zhǎngzhōngzhū *duìyīn yánjiū* 宋代西北方音——《番汉合时掌中珠》对音研究 [The Northwestern Dialect of Chinese During the Sòng Period: A Study of the Transcription Practices in the *Pearl in the Palm*]. Beijing: Zhongguo shehui kexue chubanshe, 1994.

—. *Xià-Hàn Zìdiǎn* 夏漢字典 [The Tangut-Chinese Dictionary]. Beijing: Zhongguo shehui kexue chubanshe, 1997.

—. *Xià-Hàn Zìdiǎn* 夏漢字典 [The Tangut-Chinese Dictionary], Second Edition. Beijing: Zhongguo shehui kexue chubanshe, 2008.

—. *Jiǎnmíng Xià-Hàn Zìdiǎn* 簡明夏漢字典 [The Concise Tangut-Chinese Dictionary]. Beijing: Zhongguo shehui kexue chubanshe, 2012.

Luo, Charngpeir 羅常培. *Táng Wǔdài Xīběi Fāngyīn* 唐五代西北方音 [Northwestern Dialect of Táng and Five Dynasties Period]. Pei-p'ing: Academia Sinica, 1933.

Maeda, Taeko. *The Mora and the Syllable in KiMvita (Mombasa Swahili) and Japanese.* PhD thesis, SOAS, 2001.

Matisoff, James A. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley: University of California Press, 2003.

Maspero, Henri. "Le dialecte de Tch'ang-ngan sous les T'ang." *Bulletin de l'Ecole française d'Extrême-Orient* 20, no. 2 (1920): 1–124.

Miyake, Marc. "Complexity from Compression: A Sketch of Pre-Tangut". In *Tanguty v Central'noj Azii: sbornik statej v čest' 80-letija prof. E. I. Kyčanova* [Tanguts in Central Asia: A collection of Articles Marking the 80th Anniversary of Prof. E. I. Kychanov], edited by Irina Popova, 244–261. Moscow: Oriental Literature, 2012.

Nagano, Yasuhiko, Mariëlle Prins eds. *rGyalrongic Languages Database*. Accessed November 5, 2021. https://htq.minpaku.ac.jp/databases/rGyalrong.

Oldenburg, Sergej. "Otryvki kašgarskix i sanskritskix rukopisej iz sobranija N. F. Petrovskago" [Some Kashgarian and Sanskrit Manuscripts from N. F. Petrovsky's Collection]. *Zapiski vostočnago otdelenija imp. russ. arx. obščestva* 11 (1899): 207–64.

Sofronov, Mikhail V. *Grammatika tangutskogo jazyka* [Grammar of the Tangut Language], Volume 2. Moscow: Nauka, 1968.

Sun, Jackson T.-S 孫天心. "Gōng Huángchēng Yuànshì tán Xīxiàyǔ yánjiù 龔煌城院士談西夏語研究" [Academician Gong Hwang-cherng on Tangut research]. *Shengyun luncong* 13 (2004): 1–12.

Sūn, Yǐngxīn 孙颖新. "Xīxià yìběn *Sūnzǐ Zhuàn* kǎobǔ 西夏译本《孙子传》考补" [Further Remarks on the Tangut Translation of the *Biography of Sun Tzŭ*]. *Xixia yanjiu* 6 (2010): 70–74.

Sūn, Yǐngxīn 孙颖新. "Xīxià xiěběn *Jìnzhù bā zhāijiè wén* cǎoshū guīlǜ chūtàn 西夏写本《近住八斋戒文》草书规律初探" [A Preliminary Investigation of Regular Features of the Cursive Writing in the Tangut Manuscript Version of the *Eight Upavāsa Precepts*]. *Ningxia shehui kexue* 188 (2015): 124–34.

Tai, Chung-pui 戴忠沛. Xīxiàwén fójīng cánpiàn de Zàngwén duìyīn yánjiù 西夏文佛經殘片的藏文對音研究 [*A Study of Tibetan Phonological Transcription in Tangut Buddhism Fragments*]. PhD thesis, Chinese Academy of Social Sciences, 2008.

Takubo, Shūyo 田久保周誉, ed. *Bonbun Kujaku myōōkyō* 梵文孔雀明王經 / *Ārya-Mahā-Māyūrī Vidyā-Rājñī*. Tokyo: Sankibō busshorin, 1972.

Wáng, Jìngrú 王靜如. "Fómǔ Dàkǒngquè Míngwángjīng Xià-Fàn-Zàng-Hàn hébì jiàoshì 佛母大孔雀明王經夏梵藏漢合璧校釋" [A Comparative Study of the Tangut, Sanskrit, Tibetan and Chinese Editions of the *Mahāmāyūrīvidyārājñī*]. In *Xīxià yánjiù* 西夏研究 [Tangut Studies], vol.1, edited by Wáng Jìngrú, Pei-p'ing: Academia Sinica, 1932: 181–250.