

РОССИЙСКАЯ АКАДЕМИЯ НАУК
Институт восточных рукописей

ТАНГУТЫ в Центральной Азии

*Сборник статей
в честь 80-летия
профессора Е.И.Кычанова*



МОСКВА
Издательская фирма
«Восточная литература»
2012

УДК 94(5)
ББК 63.3(5)
Т18

*Издание выполнено при поддержке
Фонда Цзян Цзин-го
(Chiang Ching-kuo Foundation for
International Scholarly Exchange), Тайвань*

Составитель и ответственный редактор
И.Ф. Попова

Тангуты в Центральной Азии : сб. ст. в честь 80-летия проф. Е.И. Кычанова / сост. и отв. ред. И.Ф. Попова ; Ин-т восточных рукописей РАН. — М. : Вост. лит., 2012. — 501 с. : ил. — ISBN 978-5-02-036505-6 (в пер.)

Сборник, в который вошли статьи отечественных и зарубежных ученых, посвящен 80-летию известного российского востоковеда, доктора исторических наук, профессора Е.И. Кычанова. Проблематика сборника задана основными доминантами многолетнего исследовательского творчества юбиляра, который, являясь в первую очередь тангутоведом и опираясь на широчайшую источниковедческую базу, блестяще разработал многие актуальные проблемы истории государственности, права, этногенеза, письменного наследия народов Китая и Центральной Азии. Большинство авторов статей постарались показать, как вопросы, поставленные в свое время в работах Е.И. Кычанова, получили дальнейшее развитие в науке.

ISBN 978-5-02-036505-6

© Институт восточных рукописей РАН, 2012
© Редакционно-издательское оформление.
Издательская фирма
«Восточная литература», 2012

Complexity from Compression: a Sketch of Pre-Tangut

Nearly half a century ago, E.I. Kychanov and M.V. Sofronov co-authored *Issledovanija po fonetike tangutskogo jazyka* (1963), the first monograph with a systematic reconstruction of Tangut phonology. Several other reconstructions have appeared since then. All have distinct values for most, if not all, of the 105 rhymes of the 文海寶韻 *Precious Rhymes of the Sea of Characters*, a monolingual Tangut dictionary. These reconstructed values generally contain few final consonants and no final obstruents. G. Clauson was skeptical about such a rhyme system: “Sofronov’s (1963) list contains sixty-five open vowels <...> It does seem impossible that a Tangut phonetician,¹ however acute his hearing, could have distinguished sixty-five different open vowel sounds, even if some of these were in fact diphthongs”.² His objections could also apply to later reconstructions. Nonetheless, there is no Chinese, Tibetan, or Sanskrit transcription evidence for a more elaborate set of final consonants in Tangut, so it is safest to continue reconstructing a large number of final vowels.

How did such a large set of vocalic distinctions come into being? In this paper, I present a scenario in which pre-Tangut, a language with a relatively simple phonology, developed into Tangut, a language with a much more complicated phonology, through a process that I call ‘compression’. Due to space limitations, I cannot offer full arguments for my speculations, though I will mention parallels in other languages for various features and sound changes.

1. Pre-Tangut

Pre-Tangut is the unattested, hypothetical ancestor of Tangut reconstructed on the basis of (1) phonological alternations in Tangut and (2) comparison with related languages. It is an intermediate stage between Proto-Tibeto-Burman and Tangut.

¹ And, I would add, any Tangut native speaker.

² Clauson 1964, p. 66.

2. Word structure of pre-Tangut

Many if not most words of pre-Tangut were sesquisyllables consisting of an unstressed presyllable followed by a stressed syllable.

*presyllable (C)(V) + syllable (C)(G)(V́)(C)(H)³

This iambic structure is similar to the structure of Old Chinese as reconstructed by Sagart (1999). It is found today in the minor-major syllable sequences of Burmese and the unrelated Mon-Khmer languages. Perhaps it can be projected back to the ancestors of pre-Tangut: Proto-Tibeto-Burman or even as far back as Proto-Sino-Tibetan.

3. Pre-Tangut presyllables

L. Sagart proposed that Old Chinese had two kinds of prefixes: fused prefixes that combined with root initials and iambic prefixes that were lost.⁴ I reconstruct a similar distinction in pre-Tangut between *three* kinds of presyllables:

1. Fused preinitials or presyllables that conditioned medial *-w-*, tense vowels, aspiration, and retroflexion (see 3.1)
2. Iambic presyllables that were lost before intervocalic lenition (see 3.2.1)
3. Iambic presyllables that were lost after intervocalic lenition (see 3.2.1).

The unstressed vowels of all three types of presyllables may have conditioned the warping of the vowel of the stressed syllable before fusion or presyllabic loss (see 3.2.2).

3.1. Preinitial consonants

Preinitial consonants could either be primary or secondary.

Primary preinitials were never followed by unstressed vowels. In other words, they were never onsets of presyllables.

Secondary preinitials were onsets of presyllables that lost their vowels:

*presyllable CV- > *preinitial C-

Preinitial consonants fused with the initial consonants of stressed syllables, resulting in *Cw*-clusters (3.1.1.1), tense consonants that in turn conditioned tense vowels before being lost (3.1.1.2), aspirates (3.1.1.3), and retroflexion (3.1.2.1).

3.1.1. Preinitial obstruents

3.1.1.1. Preinitial labials

<bC> in Tibetan transcriptions of Tangut corresponds to Tangut *Cw* (Tai 2008). This may suggest that *bC-* had become *Cw-* in the native dialect(s) of the Tibetan

³ I write asterisks before my pre-Tangut reconstructions. However, I do not write asterisks before my Tangut reconstructions because (1) all non-Tangut script representations of Tangut are reconstructions by definition and (2) the absence of asterisks helps to distinguish Tangut reconstructions from pre-Tangut reconstructions which I always write with asterisks. See appendixes 1 and 2 for lists of the initials and finals of my Tangut reconstruction. All reconstructions in this paper are mine unless explicitly stated otherwise.

⁴ Sagart 1999, pp. 17–18.

transcribers of Tangut. The <bC> transcriptions could also be taken at face value as evidence for a Tangut dialect preserving an earlier preinitial labial obstruent *P-. If Tibetan represented a real Tangut preinitial, then Tibetan medial <w> might have represented a Tangut ‘primary waw’ as opposed to a Tangut ‘secondary waw’ that developed from *P- in other dialect(s) such as the standard dialect codified in dictionaries.

*-w- > primary waw -w- in all (?) Tangut dialects

*P- > secondary waw -w- (except in the dialect(s) transcribed in Tibetan?)

Tangut *zero ~medial -w- alternations⁵ originated as zero ~ *P-alternations: e.g.,

𗉑 *Idzi* < **dzi* ‘calm’ (adjective)⁶

𗉑 *Idzwi* < **P-dzi* ‘to calm’ (verb).

Nonalternating native Tangut medial -w- may be either primary or secondary. There is no guarantee that all *P-less cognates of *P-words survived in Tangut, so a medial -w-word without a medial -w-less counterpart may not necessarily have a primary waw: e.g., 𗉑 *2dzwio* ‘person’ could be from **Cuu-dzwoH* with primary waw or from **Puu-dzoH* whose presyllable conditioned a secondary waw.⁷

There are no Tangut words with labial initials followed by -w- (*pw-*, *phw-*, *bw-*, *mw-*, *vw-*). If pre-Tangut had **PP*-sequences, they were simplified to *P-* in Tangut: e.g., **P-m-* > **mw-* > *m-*, etc.

3.1.1.2. Preinitial coronals

Gong Hwang-cherng observed alternations between Tangut lax and tense vowels.⁸ Gong (1999) then proposed that tense vowels (written here with subscript dots) originated from preinitial **s-* on the basis of external comparisons: e.g.,

𗉑 *Itəu* ‘thousand’: Written Tibetan *stong* ‘id.’

Since lax-tense vowel alternations in Tangut have multiple functions,⁹ perhaps tense vowels originated from more than one voiceless coronal obstruent that I will symbolize as **S-*. This **S-* could either be part of the root or a prefix. I reconstruct it as a prefix if a tense vowel word has a lax vowel cognate within Tangut or has an **s-*-less external cognate: e.g.,

𗉑 *Ikhwa* < **khwa* ‘distant’

𗉑 *Ikhwa* < **S-khwa* ‘to keep at a distance.’

⁵ Gong 1988, p. 798–800.

⁶ English glosses of Tangut words are based on the glosses in Gong 1988 and Li 2008.

⁷ A presyllable with a high vowel is necessary to account for the warping of -o to -io. See 3.2.2.2.

⁸ Gong 1988, pp. 805–811. At the time, Gong was using Sofronov’s reconstruction with “minor revisions” (1988, p. 784). Sofronov’s reconstruction did not have any retroflex vowels, so some of the lax-tense cognate sets in Gong (1988) would now be reinterpreted as nonretroflex-retroflex cognate sets in reconstructions with retroflex vowels like the reconstruction in Gong (2003) or the reconstruction in this paper.

⁹ Gong 1988, pp. 810–811.

How could a consonant condition tension in a following but nonadjacent vowel? Modern Korean tense consonants (*pp-*, *tt-*, *ss-*, *cc-*, *kk-*) originated from Late Middle Korean clusters with *p-* and/or *s-*. According to S. Martin,¹⁰ “The laryngeal tension [of modern Korean tense consonants] continues on into the vowel, which can be described as ‘laryngealized’”. The development of tense consonants and vowels in Korean could be formulated as

$$p/sCV > CCV > \underset{\cdot}{C}V > \underset{\cdot}{C}\underset{\cdot}{V} / \underset{\cdot}{C}\underset{\cdot}{V}$$

with the subscript dots used by Tangutologists to represent tenseness. Note that in modern Korean, only the tenseness of consonants is phonemic, whereas the tenseness of vowels is subphonemic. However, in Tangut, the tenseness of consonants was lost, so the tenseness of vowels became phonemic:

$$*SCV > *CCV > *\underset{\cdot}{C}V > *\underset{\cdot}{C}\underset{\cdot}{V} > *C\underset{\cdot}{V} / C\underset{\cdot}{V}.$$

3.1.1.3. Preinitial gutturals

Gong Hwang-cherng¹¹ found alternations between Tangut nonaspirated and aspirated initials. I derive these alternations from earlier *zero ~ *K-alternations. *K- was a voiceless velar, uvular, or glottal obstruent that devoiced voiced/ consonants: e.g.,

$$*Kb- > ph-, *Kd- > th-, *Kg- > kh-, *Kd\underset{\cdot}{z}- > t\underset{\cdot}{j}h-, *Kl- > lh-$$

Voiced consonants are preserved in nonprefixed members of voiced-aspirated cognate sets: e.g.,

$$\begin{array}{l} \text{𐰇} \text{ } Igi < *gi \text{ ‘to fall, to lose’} \\ \text{𐰇} \text{ } Ighi < *K-gi \text{ ‘to let fall, to cause to lose.’} \end{array}$$

Note that not all such sets involved a *K-prefix. Some doublets reflect different strata of borrowing from Chinese: one before devoicing and another after devoicing: e.g.,

$$\begin{array}{l} \text{𐰇} \text{ } Idza \text{ ‘mixed’} < \text{Late Middle Chinese } \text{𐰇} \text{ } *dzap \text{ ‘id.’ (early loan)} \\ \text{𐰇} \text{ } Itsha \text{ ‘mixed’} < \text{Tangut Period Northwestern Chinese } \text{𐰇} \text{ } *tsha < \text{Late Middle} \\ \text{Chinese ‘id.’ (late loan; aspirated } tsh\text{- directly from Chinese rather than from} \\ \text{pre-Tangut } *K\text{-}dz\text{-).} \end{array}$$

*K- aspirated most voiceless obstruents: e.g.,

$$\begin{array}{l} \text{𐰇} \text{ } Ika < *ka \text{ ‘center’} \\ \text{𐰇} \text{ } Ikha < *K-ka \text{ ‘in’ (postposition)} \end{array}$$

One might expect *k- to have merged with *S- and become k- followed by a tense vowel (see 3.1.1.2). If such a merger occurred, then *Ikha* ‘in’ must have had a non-*k- guttural preinitial (e.g., *x-ka; see below). If such a merger did not occur,

¹⁰ Martin 1992, p. 27.

¹¹ Gong 1988, pp. 785–796.

then perhaps aspiration preceded tension, so **k-k-* became *kh-* before **sk-* became a new **kk-* that was ultimately reduced to *k-* before a tense vowel:

Early pre-Tangut	<i>*k-k-</i>	<i>*Sk-</i>
Aspiration	<i>*kh-</i>	<i>*Sk-</i>
Gemination	<i>*kh-</i>	<i>*kk-</i>
Tangut	<i>kh-</i>	<i>k-</i> + tense vowel

The relative chronology of the rules in this paper has yet to be worked out.

One also might expect **Ks-* to become an aspirated *sh-* like modern Burmese *ə*. However, there is no evidence for such an initial in Tangut. **K-* may have conditioned tense vowels after *s-*: e.g., 𐞗 *Isə* ‘three’ may be from **sə* < **sə* < **sso* < **xso* < **Kso* (cf. the *g-* of Written Tibetan *gsum* ‘three’).

In Korean, **hVC-* as well as **kC-* developed into Late Middle Korean aspirates.¹² I assume Tangut also underwent similar sound changes and therefore cannot rule out the possibility of velar, uvular, and/or glottal fricative sources of aspiration: e.g., **xC-* > *Ch-*. Modern Mawo Qiang, a distant relative of Tangut, has *xC-* and *χC-* clusters.¹³

3.1.2. Preinitial sonorants

3.1.2.1. Preinitial **r-*

Pre-Tangut preinitial **r-* was one source of retroflexion in Tangut vowels: e.g.,

𐞗 *lɿəʔ* < **ru-ləʔ* ‘four.’

For the other source of retroflexion, see 4.4.2.2.

Retroflex vowels are very common in Tangut. Perhaps some were conditioned by preinitial **l-* and even preinitial dental stops that merged with preinitial **r-*: e.g., **TV-* > **T-* > **r-*.

Nonretroflex-retroflex cognate sets can be reconstructed with **Ø-* ~ **r-*: e.g.,

𐞗 *lza* < **za* ‘red face’

𐞗 *lzaʔ* < **r-za* ‘red-faced ancestor.’

I reconstruct **r-* as a prefix even in retroflex vowel words like 𐞗 *lɿəʔ* ‘four’ which lack nonretroflex vowel cognates within Tangut if they have **r-*-less exterior cognates: e.g., Written Tibetan *bzhi* < **b-lyi* ‘four’ and Old Chinese 四 **s-li-s* ‘four.’

3.1.2.2. Preinitial nasals?

I do not know of any voiceless ~ voiced obstruent alternations that suggest **zero* ~ **preinitial nasal* alternations in pre-Tangut: e.g., **p-* ~ **b-* < **p-* ~ **Np-*, etc.

¹² Vovin 2010, p. 11; Lee and Ramsey 2011, p. 89.

¹³ Sun Hongkai 1981, p. 27.

However, perhaps some Tangut voiced obstruent initials are from pre-Tangut *preinitial nasal + obstruent initial sequences: e.g., *b-* < **Nb-*, etc.

3.2. Presyllabic vowels

The vowels of pre-Tangut presyllables have left two kinds of traces in Tangut.

3.2.1. Intervocalic lenition

Pre-Tangut presyllables that were lost at a very late date conditioned the lenition of main syllable initials in intervocalic position:

	Early presyllable loss	Late presyllable loss	Fusion
Early pre-Tangut	* <i>CV-CV'</i>	* <i>CV-CV'</i>	* <i>CV-CV'</i>
Loss of presyllabic vowel; presyllable becomes preinitial	* <i>CV-CV'</i>	* <i>CV-CV'</i>	* <i>C-CV'</i>
Early loss of presyllable	* <i>CV'</i>	* <i>CV-CV'</i>	* <i>C-CV'</i>
Lenition	* <i>CV'</i>	* <i>CV-C_{lenited}V'</i>	* <i>C-CV'</i>
Late loss of presyllable	* <i>CV'</i>	* <i>C_{lenited}V'</i>	* <i>C-CV'</i>

Forms subject to sound changes are in **bold**.

All obstruents at the same point of articulation merged into a single lenited initial. The reflexes of Tangut lenition are similar to those of intervocalic lenition in Vietnamese and Korean.

*Labials > *v-* (phonetically [β]?; cf. Middle Vietnamese [β] < **-p-*, **-b-* and Middle Korean [β] < **-p-*)

*Dentals > *l-* (cf. Middle Korean [r] < **-t-*)

*Alveolars > *z-* (cf. Middle Korean [z] < **-s-*, **-ts-*)

*Alveopalatals > *ʒ-* (cf. Middle Vietnamese [ʒ] < **-c-*, **-ʃ-*)

*Velars > *ɣ-* (cf. Middle Vietnamese [ɣ] < **-k-*, **-g-* and Middle Korean [ɣ] < **-k-*)

Lenition obscures etymological relationships: e.g., the Tangut cognate of Written Tibetan *gcig* 'one' and Old Chinese 隻 **tek* 'single' is 𪛗 *llew* < **k_A-tek* or **k_A-tik*. (I assume the pre-Tangut prefix had an initial **k-* corresponding to Written Tibetan *g-*, though other initials are possible. See 3.2.2 for the reasoning behind reconstructing **k* as the vowel of the presyllable. See 4.4.1.1 for the **k* > *-w* shift.)

3.2.2. Stressed vowel warping

In 2008, I proposed that the Old Chinese type A/B distinction was conditioned by presyllabic vowels.¹⁴ The following adaptation of that theory and A. Schuessler's

¹⁴ Miyake 2008.

(2007, 2009) theory of vowel warping in Chinese can account for much of the large rhyme inventory of Tangut.

I reconstruct at least two different vowels in Tangut presyllables:

- a lower vowel symbolized¹⁵ as **ɿ* (cf. the Middle Korean ‘minimal vowel’ [ɿ])
- a higher vowel symbolized as **u* (cf. the Middle Korean ‘minimal vowel’ — [u])

These vowels may have resulted from the merger of a larger number of even earlier unstressed vowels.

Pre-Tangut main syllable vowels also belonged to lower and higher classes:

Higher	<i>*i</i>	<i>*u</i>
Lower	<i>*e</i>	<i>*ə</i> ¹⁶
	<i>*a</i>	<i>*o</i>

Pre-Tangut had partial vowel harmony (under Chinese influence?). If the height class of an unstressed presyllabic vowel matched the height class of a stressed vowel, the latter did not change either before or after presyllable loss: e.g.,

- *Cu-Cí* > **Cu-Cí* > *Ci* (higher + higher)
- *Cɿ-Cá* > **Cɿ-Cá* > *Ca* (lower + lower).

However, if the height class of an unstressed presyllabic vowel did not match the height class of a stressed vowel, the latter warped (partly lowered or raised) before the presyllable was lost: e.g.,

- *Cɿ-Cí* (lower + higher) > **Cɿ-Cəí* > *Cəi* (lower + **partly lowered**)
- *Cu-Cá* (higher + lower) > **Cu-Ciá* > *Cia* (higher + **partly raised**).

Partly lowered vowels developed into diphthongs beginning with *ə*: *əu*, *əi*.

Partly raised vowels developed into diphthongs beginning with *i* (after *v*-, *l*-, and alveopalatals) or **i* (after all other initials): *ia*, *iə*, *ie*, *io* ~ *ia*, *iə*, *ie*, *io*. (There are exceptions to this pattern of complimentary distribution.) The *i* that resulted from partial raising is not to be confused with the *i* that developed before high vowels after *v*-, *l*-, and alveopalatals (see 4.3.1).

If a presyllable has lenited a following initial but has not warped a following stressed vowel, I reconstruct the presyllabic vowel with the height class of the stressed vowel: e.g.,

- 𗰨 *llew* < **kɿ-tek* ‘one’¹⁷ (lower + lower) (**kuu*- with a higher vowel would have warped **e* to **ie*.)

¹⁵ I use the term ‘symbolized’ to indicate that **ɿ* and **u* may not have been the precise phonetic values of the Tangut presyllabic vowels. They could have been central **i* and **e*, etc. What matters is their heights relative to each other.

¹⁶ It is also possible that **ə* belonged to the higher vowel class of **i* and **u*, but then its behavior would be anomalous, as it would be the only higher class vowel that bent upward and never bent downward.

¹⁷ The pre-Tangut form could also have been **kɿ-tik*. The lower vowel of the presyllable would have conditioned the warping of **i*: **ik* > **əik* > *ew*.

藏 *lʒiɪw* < **Cu-fuk* ‘juniper tree’¹⁸ (higher + higher) (**CA-* with a lower vowel would have warped **u* to **əu* which would then have monophthongized to *e* before *-w*. For *-iɪw* < **-uk*, see 4.4.1.1).

Medial *-i-* alternations¹⁹ may reflect earlier prefixes: e.g.,

藏 *lʒhəu* < **CA-tshu* ‘shovel’ (prefix conditioned vowel warping)

藏 *lʒhiu* < **tshu* ‘shovel’ (no prefix; **u* became *iu* after **tsh-*; see 4.3.2).

However, “no semantic difference can be observed” between alternating forms.²⁰ Furthermore, these alternations occur mostly in words with *u*. These cognate sets may reflect interdialectal and/or dialect-internal variation in the pronunciation of /u/ rather than morphology.

3.2.3. Stressed vowel brightening

Perhaps there were more than two kinds of presyllabic vowels. ‘Brightening’ (raising of **a* to *i*) in Tangut²¹ may have been conditioned by high front vowels in presyllables: e.g.,

**Ci-Cá* > *Ci* (= *Cji* in Gong’s reconstruction used by Matisoff).²²

The height of a palatal presyllabic vowel may have determined the degree of brightening: e.g.,

**Ce-Cá* > *Cie* (= *Cjij* in Gong’s reconstruction used by Matisoff)

with a partly high diphthong rather than *Ci* with a high monophthong.

There are also sporadic cases in which pre-Tangut **a* was raised to *ə*: e.g.,

藏 *lɪwə* < **PV-ɪa* ‘five’: Written Tibetan *lŋa*, Old Chinese 五 *ɲʰaʔ* ‘id.’

I hesitate to reconstruct yet another presyllabic vowel to account for only a few instances.

4. Pre-Tangut stressed syllables

4.1. Pre-Tangut stressed syllable initials

I tentatively project the Tangut initial inventory (see Appendix 1) back into pre-Tangut.

¹⁸ Jacques (2004, p. 160; 2006) compared this Tangut word to Japhug rGyalrong *ɕɣɣ* ‘juniper tree’ and Written Tibetan *shug-pa* ‘juniper tree’.

¹⁹ Gong 1988, pp. 796–798.

²⁰ Gong 1988, p. 798.

²¹ Matisoff 2004.

²² The negative particle 藏 *lmi*, cognate to Old Chinese 無 **ma* ‘not have’, may pose a problem for this derivation, as it would have to come from a sesquisyllabic **Ci-ma*. Would such a high-frequency particle really be so phonologically complex? On the other hand, it is hard to believe that **ma* would brighten to *lmi* without any conditioning factor. Not all Tangut **a* brightened, so one cannot attribute the raising to a regular vowel shift.

A few Tangut initials may be secondary in origin: e.g., an initial may always be the result of lenition like Vietnamese *g-* [ɣ] which is only from **CV-K-*.

I presume that pre-Tangut had more stressed syllable initials than presyllabic initials: e.g., **k-*, **kh-*, **g-* were possible stressed syllable velar stop initials, but **k-* may have been the only possible presyllabic velar stop initial.

All vowels after pre-Tangut syllable-initial **r-* became retroflex: **rV > rV'*. Note that *medial *r-* did not condition retroflex vowels. See 4.2.4.

A couple of external correspondences suggest that uvulars may have conditioned Tangut Grade II vowels *u* and *i*:

𗑦 *lyu* < **Gu?* ‘head’: Baxter and Sagart’s (2012) Old Chinese 后 **G^ʰ(r)o?* ‘sovereign’ (< ‘head of a state’), Written Tibetan *mgo* ‘head’
 𗑦 *lkhɪ* < **Ci-qa?* ‘bitter’: Mawo and Taoping Qiang *qha*,²³ Zhongu Tibetan *qhɛnde* ‘to be bitter’,²⁴ Written Tibetan *kha* ‘bitter’, Baxter and Sagart’s (2012) Old Chinese **kh^ʰa?* (not **qh^ʰa?*).

(See 4.2.4 for more on Grade II.) However, the reconstruction of uvulars in Old Chinese is still unsettled. A. Schuessler (2007; 2009) does not reconstruct them in Old Chinese. Moreover, note that Baxter and Sagart reconstruct a velar in **kh^ʰa?* ‘bitter’ instead of a uvular corresponding to a uvular in Qiang and Zhongu. N. Hill²⁵ regarded Zhongu uvulars as being “due to the influence of a Qiangic substrate.” Perhaps the uvular in Old Chinese ‘head’ is primary whereas the uvular in Qiang and Zhongu ‘bitter’ is secondary.²⁶ Did Tangut inherit a secondary uvular in ‘bitter’ from Proto-Qiangic? In any case, there is no strong evidence for a medial **r-* in either ‘head’ or ‘bitter’ that would normally condition Grade II (see 4.2.4), so the vocalism of those words needs another explanation.

4.2. Pre-Tangut medial glides

4.2.1. Pre-Tangut medial **-w-*

This medial is preserved in Tangut. It is primary waw, whereas secondary waw reflects an earlier **P-* (see 3.1.1.1).

4.2.2. Pre-Tangut medial **-j-*

A palatal glide may be the source of some *-i-* and *-i-* in Tangut: e.g.,

𗑦 **sjeH > 2sie* ‘knowledge’: Written Tibetan *shes-pa*, Proto-Tibeto-Burman **syey-s* ‘id.’²⁷

It is also possible to derive *2sie* from a yodless **Cuu-seH* with partial raising of **e*.

²³ Sun Hongkai 1981, p. 216.

²⁴ Sun Jackson 2003, p. 772.

²⁵ Hill 2010, p. 120.

²⁶ I think it may be possible to reconstruct a uvular in Old Chinese ‘bitter’ on entirely internal grounds, enabling me to reconstruct a uvular at the Proto-Sino-Tibetan level for that word.

²⁷ Matisoff 2003, p. 614.

4.2.3. Pre-Tangut medial *-rj-

The pre-Tangut cluster **rj-* became Tangut *ʃi-'*: e.g.,

𐵑 **rjat* > *ʃia'* 'eight': Written Tibetan *brgyad*, Old Chinese **p^hret* 'id.'

4.2.4. Pre-Tangut medial *-r-

According to G. Jacques (2009), Gong (1993) derived his Grade II *-i-* from an earlier **-r-*. Gong's Grade II *iV*-diphthongs correspond to my Grade II lowered vowels:

Pre-Tangut	<i>*ru</i>	<i>*ri</i>	<i>*ra</i>	<i>*rə</i>	<i>*re</i>	<i>*ro</i>
Gong's Grade II ²⁸	(none)	<i>ie</i>	<i>ia</i>	<i>iə</i>	<i>iej</i>	<i>io</i>
Grade II in this paper	<i>ʊ</i>	<i>ɪ</i>	<i>æ</i>	<i>ɐ</i>	<i>ɛ</i>	<i>ɔ</i>

This vowel shift pattern is similar to what Schuessler (2007, 2009) reconstructed in Chinese:

Old Chinese **râ* > Later Han Chinese *a* (a low front vowel close to [æ] and distinct from back [ɑ])

Old Chinese **rê*, **rê* > Later Han Chinese **ɛ*

Old Chinese **rô* > Later Han Chinese **ɔ*

In Chinese, this shift only occurred in type A syllables (indicated with circumflexes over vowels in Schuessler's notation). Perhaps the Tangut shift only occurred in syllables with low vowels or partly lowered vowels:

Pre-Tangut after vowel lowering	<i>*rəu</i> < <i>*ru</i>	<i>*rəi</i> < <i>*ri</i>	<i>*ra</i>	<i>*rə</i>	<i>*re</i>	<i>*ro</i>
Grade II	<i>ʊ</i>	<i>ɪ</i>	<i>æ</i>	<i>ɐ</i>	<i>ɛ</i>	<i>ɔ</i>

**-r-* may have vanished before high vowels:

Pre-Tangut after vowel raising	<i>*ru</i> > <i>*riu</i>	<i>*ri</i>	<i>*ria</i>	<i>*riə</i>	<i>*rie</i>	<i>*rio</i>
Grade III	<i>iu</i>	<i>ii</i>	<i>ia</i>	<i>iə</i>	<i>ie</i>	<i>io</i>
Grade IV	<i>iu</i>	<i>i</i>	<i>ia</i>	<i>iə</i>	<i>ie</i>	<i>io</i>

See 4.3.1 and 4.3.2 for the *-i-* and *-i-* that developed before **u* and **i*.

²⁸ Gong's pre-Tangut forms might not necessarily correspond to mine.

The correspondence of 𪛗 *lʃhiw* ‘six’ to Written Tibetan *drug* ‘id.’ suggests that some Tangut alveopalatal affricates may be from **Tr*-clusters. Perhaps ‘six’ was once **k-truk* with a preinitial **k-* that conditioned aspiration (see 3.1.1.3). (See 4.4.1.1 for the development of *-iiw* from **-uk*.)

4.2.5. Pre-Tangut medial **-l-*?

There are several instances of Tangut *lh-* corresponding to Japhug rGyalrong *k-* presyllables followed by *l*, *l̥*, or *j* < **lj-* in Jacques (2006): e.g.,

𪛗 *llhew* < **-k* ‘to graze’: Japhug rGyalrong *kɣ lɣy* ‘id.’

These correspondences suggest that some Tangut *lh-* may be from **kl-*.

Pre-Tangut **-l-* in other environments might have merged with another medial or disappeared without a trace.

4.3. Pre-Tangut stressed vowels

I project the six basic vowel types of Tangut (*u*, *i*, *a*, *ə*, *e*, *o*; see Appendix 2) back into Proto-Tangut with only a few changes:

- **-a* is restored in ‘brightened’ syllables (see 3.2.3).
- *-iiw* in ‘six’ and ‘juniper tree’ (see 3.2.2, 4.2.4) and perhaps other words is derived from **-uk* (see 4.4.1.1). *-iw* may also sometimes be from **-uk*.
- *-o* is partly from **-aŋ*,²⁹ cf. Japhug rGyalrong *-o* < **-aŋ*,³⁰ and Tangut period Northwestern Chinese *-o* < **-aŋ*).

It is not clear whether the long vowels of Tangut are primary or secondary (see 4.4.4.1 and 4.4.4.2). So pre-Tangut may have had either six or twelve vowels (six short and six long). Nasalization, tensing, retroflexion, and diphthongization occurred later.

Old Chinese as reconstructed by W. Baxter and L. Sagart (2012) also had the same basic six vowels as Tangut, though one should not expect simple one-to-one correspondences between the two vowel systems: e.g., Baxter and Sagart’s Old Chinese 馬 **m^hraʔ* ‘horse’ may correspond to pre-Tangut **Cu-re* (> Tangut 𪛗 *lrie’*) ‘id.’, not **mraH*.

4.3.1. Grade III *-i-*

The high vowels **i* and **u* became *ii* and *iu* after Grade III initials (*v-*, *l-*, and alveopalatals).

**-iuk* became *-iiw* (see 4.4.1.1).

4.3.2. Grade IV *-i-*

The high vowel **u* became *iu* after Grade IV initials (initials other than *v-*, *l-*, and alveopalatals) whereas **i* remained unchanged.

²⁹ Gong 1995.

³⁰ Jacques 2004, p. 232.

Tangut had no simple rhyme *-u* (see Appendix 2). This situation may have arisen under the influence of Late Middle Chinese whose **-u* had similarly shifted to **-iu* or **-iu*,³¹ leaving a gap to be filled later by **-o* after raising.

4.4. Pre-Tangut codas

Although Tangut had no final obstruents and few final consonants, pre-Tangut once had a richer set of codas like its relatives Japhug rGyalrong, Classical Tibetan, Old Burmese, and Old Chinese.

4.4.1. Pre-Tangut obstruent codas

4.4.1.1. Pre-Tangut **-k*

**-k* became *-w* after front vowels but disappeared elsewhere. See Gong (1995) for examples.

Although **-iuk* had a back vowel, this rule applied to this rhyme after **u* dissimilated to a front vowel **i* before a velar coda:

**-iuk* > **-iuɣ* > **-iuɣ* > **-iiɣ* > **-iiw*.

See ‘six’ (3.2.2) and ‘juniper tree’ (4.2.4).

It is tempting to regard the long *-aa* of 𐰇𐰏 *2mia-2niaa* ‘Tangut’ (cf. Written Tibetan *mi-nyag* ‘id.’) as an instance of compensatory lengthening after the loss of **-k*. However, other **-k* words like

𐰇 *Ido* < **dok* ‘poison’; borrowed from Middle Chinese 毒 **dowk* ‘id.’

have short vowels. Could the *-aa* of ‘Tangut’ be from **-aakH* with an original long vowel? (The final **-H* is the source of the second tone. See 4.5.)

4.4.1.2. Other pre-Tangut stop codas

The final **-p* and **-t* that one would expect from comparison with Old Chinese, Written Tibetan, and Old Burmese have vanished without a trace: e.g.,

𐰇 **Cɿ-kap* > *lɣa* ‘needle’: Japhug rGyalrong *ta-qaβ*, Written Tibetan *khab* ‘id.’)

𐰇 **ʔjat* > *lʔaʔ* ‘eight’: Written Tibetan *brgyad*, Old Chinese **pʔret* ‘id.’).

There are a few instances of long vowels in probable **-t* words: e.g.,

𐰇 **Cuu-maat* > *lmiaa* ‘fruit’: Japhug rGyalrong *suu-mat* ‘id.’

but these vowels may be primary long vowels rather than remnants of lost stops.

4.4.1.3. Pre-Tangut fricative codas

See 4.5.

³¹ Compare Kan-on 九 *kiu* ‘nine’ (borrowed from northwestern Late Middle Chinese) with Go-on *ku* ‘id.’ (borrowed from southern Early Middle Chinese prior to **u*-breaking).

4.4.2 Pre-Tangut sonorant codas

4.4.2.1. Pre-Tangut nasal codas

Nasals disappeared after all vowels, leaving behind nasalization in some cases with at least two major exceptions:

— There are no native nasalized *u*-syllables. All nasalized *u*-syllables are Chinese borrowings.

— **-aŋ* became *-o* (see 4.3).

4.4.2.2. Pre-Tangut liquid codas

Final **-r* is another source of vowel retroflexion: e.g.,

𪛗 *lkaaʳ* < **kaar* ‘to measure’: Japhug rGyalrong *kɣ-skɣɹ* ‘to weigh.’

Since a final *-Nr* or *-rN* cluster is absent from languages of the region, I assume that the nasalized retroflex vowels of Tangut rhymes 65, 76, 97, and 98 originated from preinitial **r-* + final **-N* sequences: **r-CVN* > *CṼʳ*.

4.5. Pre-Tangut tonogenetic codas

Tangut had two basic tones, a ‘level tone’ and a ‘rising tone’.³² The terms were obviously adopted from the Chinese phonological tradition and may not be meant to be taken at face value as descriptions of tonal contours. They may have meant nothing more than ‘first category’ and ‘second category’. They could even have referred to phonations rather than tones, but I will continue to use the traditional term ‘tone’.

Given that the Tangut level tone was much more common than the Tangut rising tone and that the rising and departing tones of Middle Chinese originated from Old Chinese final glottals, I derive the Tangut rising tone from a lost final glottal **-H*. This **-H* in turn may be from an even earlier **-s* (cf. Old Chinese **-s* and Written Tibetan *-s*) and/or **-ʔ* (cf. Old Chinese **-ʔ*).

Tonal alternations³³ arose from zero ~ **-H* alternations. An **-H* suffix could be added after other codas: e.g., the rising tone word 𪛗 *2lew* < **Cɿ-tek-H* or **Cɿ-tik-H* ‘same’³⁴ is a suffixed cognate of the level tone word 𪛗 *1lew* < **kɿ-tek* or **kɿ-tik* ‘one’.

Old Chinese **-s* could also follow any coda. Written Tibetan *-s* has a more restricted distribution; homorganic *-Cs* sequences are not possible.

If a Tangut rising tone word has no known level tone cognates, its **-H* can be tentatively regarded as part of its root unless external comparison reveals that the **-H* is a suffix.

³² I will not deal with the ‘entering tone’ in the *Precious Rhymes of the Sea of Characters* and other tonal oddities here.

³³ Gong 1988, pp. 821–832.

³⁴ I am not sure whether ‘same’ had the same numerical **kɿ-*prefix as ‘one’. The unwarped nonhigh *e* of *2lew* necessitates the reconstruction of a nonhigh **ɿ* in the presyllable.

Conclusion

The pre-Tangut phonological system that I have reconstructed in this paper brings Tangut typologically closer to Old Chinese while also accounting for Tangut-internal morphological alternations. It is far from a finished product, as it is based only on a small number of examples. Application of my hypotheses to the Tangut lexicon as a whole will undoubtedly result in the reformulation or even rejection of some of my proposals. Nonetheless, I remain confident that Tangut phonological history will eventually be integrated into the larger saga of monosyllabic compression across the Sinosphere.

Appendix 1

Tangut initials

This system is nearly identical to Gong (2003). I write his $w \ t\acute{s} \ t\acute{s}h \ d\acute{z} \acute{s} \acute{z} \cdot$ as $v \ t\acute{f} \ t\acute{f}h \ d\acute{z} \acute{f} \acute{z} \ ?$. Roman numerals refer to the initial classes of the Tangut 同音 *Homophones* dictionary. Unlike Nishida (1964) or Arakawa (1999), neither Gong nor I reconstruct distinct initials for class IV. Alternative phonetic interpretations are in the right-hand column.

I	$p-$	$ph-$	$b-$	$m-$		
II					$v-$	[w]?
III	$t-$	$th-$	$d-$	$n-$		
V	$k-$	$kh-$	$g-$	$\eta-$		
VI	$ts-$	$tsh-$	$dz-$		$s-$	
VII	$t\acute{f}$	$t\acute{f}h-$	$d\acute{z}$		\acute{f}	retroflex [tʂ tʂh dz,ʂ]?
VIII	\acute{z}	$x-$	$y-$			glottal [ʔ h h̥]?
IX	$l-$	$lh-$	$z-$		$\acute{z} \ - \ r-$	[ʈ ʈ̚ ʂ z, r]?

$v-$, $l-$, and the alveopalatals were usually followed by Grade III rhymes with $-i-$ rather than Grade IV rhymes with $-i-$. There was something antipalatal about those consonants, so I suspect $l-$ may have been velarized [ʈ] and the alveopalatals were really retroflexes. The correspondence of $t\acute{f}h-$ to Written Tibetan $dr-$ in ‘six’ (4.2.4) suggests that the alveopalatals might have been retroflexes.

$v-$ may have been [w] like Polish $ł$ or Belarusian \check{y} from earlier nonpalatalized l . However, Tibetan transcriptions of $v-$ as <b(w)>, <ḥbh> and even <ww>³⁵ suggest that Tangut $v-$ had more friction than $w-$.

³⁵ Nishida 1964, pp. 82–83; Tai 2008, pp. 177–178.

Appendix 2

Tangut rhymes

This system is a revision of Gong (2003). Although the phonetic values are somewhat different, the rhyme groups are nearly identical to his.

Grade III and IV rhyme numbers marked with a and b are in complementary distribution. Rhymes unique to Chinese loanwords have no pre-Tangut sources and hence are in parentheses. Variants of rhymes with medial *-w-* are not listed.

Pre-Tangut basic vowel	Grade I	Grade II	Grade III	Grade IV
* <i>u</i>	1. -əu	4. -ʊ ³⁶	2. -iu	3. -iu
	5. -əəu	6. -ʊʊ ³⁷	7a. -iuu	7b. -iuu
	(104. -əũ)			
	61. -əu		62a. -iu	62b. -iu
	80. -əu'			81. -iu'
* <i>i</i>	8. -əi	9. -ɪ	10. -ii	11. -i
	12. -əəi	13. -ɪɪ	14a. -iii	14b. -ii
	15. -əĩ		16a. -iĩ	16b. -ĩ
	68. -əi	69. -ɪ	70a. -ii	70b. -i
	82. -əi'	83. -ɪ'	84a. -ii'	84b. -i'
	99. -əəi'		101a. -iii'	101b. -ii'
* <i>a</i>	17. -a	18. -æ	19. -ia	20. -ia
	22. -aa	23. -ææ	21. -iaa	24. -iaa
	25. -ã	26. -æ̃	27a. -iã	27b. -iã
	66. -a		67a. -ia	67b. -ia
	85. -a'	86. -æ'	87a. -ia'	87b. -ia'
	88. -aa'		89a. -iaa'	89b. -iaa'
				(105. -ya)
*ə	28. -ə	29. -ɹ	30. -iə	31. -iə
	32. -əə		33a. -iəə	33b. -iəə
	71. -ə		72a. -iə	72b. -iə
	90. -ə'	91. -ɹ'	92a. -iə'	92b. -iə'
			100a. -iəə'	100b. -iəə'

³⁶ Gong Hwang-cherng classified rhyme 4 as Grade I and reconstructed it as homophonous with Grade I rhyme 1. However, there are minimal pairs distinguishing rhymes 1 and 4, so the two rhymes must have been distinct. Since rhymes 2 and 3 were Grades III and IV, rhyme 4 might have been Grade II. Unfortunately, there are no diagnostic Grade II initials (*v-*, *l-*, alveopalatals) in rhyme 4 syllables. However, the order of Tangut rhymes seems to be based on a Chinese model, and the first four Tangut rhymes (Grade I 1, Grade III 2, Grade IV 3, and Grade II 4) apparently correspond to the first three Middle Chinese rhymes (Grade I 東/冬, Grade III/IV 鐘, and Grade II 江). Moreover, there are no alveolar initials unique to Grades I and IV in rhyme 4 syllables. Rhyme 4 can only be Grade II or Grade III (as in Arakawa 1999).

³⁷ Gong classified the extremely rare rhyme 6 as Grade III. There are only two different rhyme 6 syllables, *khʊʊ* and *ʒʊʊ*. *kh-* and *ʒ-* can only coexist in Grade II, so I classify rhyme 6 as Grade II.

*e	34. -e	35. -ɛ	36. -ie	37. -ie
	38. -ee	39. -ɛɛ	40a. -iee	40b. -iee
	41. -ē	42. -ē̃	43a. -iē̃	43b. -iē̃
		76. -ē̃	65a. -iē̃	65b. -iē̃
		63. -ɛ̃	64a. -iɛ̃	64b. -iɛ̃
	77. -e ^r	78. -ɛ ^r	79a. -ie ^r	79b. -ie ^r
*ik/ek/uk	44. -ew	45. -ɛw	46a. -iew	46b. -iew
	< *-ik/-ek?	< *-ek only?	< *-ek only	< *-ek only
			47a. -iiw	47b. -iw
			< *-ik, *-uk	< *-ik, (*-uk?)
	93. -e ^r w			94. -i(e) ^r w
< *-ik/-ek?			< *-ik/-ek?	
*o	51. -o	52. -ɔ	53a. -io	53b. -io
			50. -wio	
	54. -oo	55a. -ɔɔ	55b. -ioo	55c. -ioo
	56. -ō	57. -ō̃	58a. -iō̃	58b. -iō̃
		59. -ō̃ō̃	60a. -iō̃ō̃	60b. -iō̃ō̃
	73. -o	74. -ɔ	75a. -io	75b. -io
	95. -o ^r	96a. -ɔ ^r	96b. -io ^r	96c. -io ^r
	102. -oo ^r			103. -ioo ^r
	97. -ō ^r			98. -iō ^r

The *o-rhymes had some unusual characteristics (i.e., a separate rhyme 50 -wio distinct from 53a -io which could also be preceded by -w-; a three-way split of rhymes 55 and 96) that deserve investigation.

50 -wio could only have the level tone, whereas 53a -wio with -w- could only have the rising tone.

Perhaps /oo/ was [ɔɔ] after the high vowels /i i/, so 55a [ɔɔ] could rhyme with 55b [iɔɔ] and 55c [ioɔ] and 96a [ɔɔ^r] could rhyme with 96b [iɔɔ^r] and 96c [ioɔ^r].

References

- Arakawa 1999 — Arakawa Shintarō 荒川慎太郎. “Ka-zō taion shiryō kara mita Seikago no seichō” [A Study on Tangut Tones from Tibetan Transcriptions] 夏藏对音資料からみた西夏語の声調. In *Gengogaku kenkyū* 言語学研究 [Linguistic Research], 17–18 (1999), pp. 27–44.
- Baxter and Sagart 2012 — Baxter W., Sagart L. “Baxter-Sagart Old Chinese reconstruction (Version 1.00).” In <http://crlao.ehess.fr/docannexe.php?id=1202> (accessed January 27, 2012).
- Clauson 1964 — Clauson G. “The Future of Tangut (Hsi Hsia) Studies”. In *Asia Major*, 11.1 (1964), pp. 54–77.
- Gong 1988 — Gong Hwang-chen 龔煌城. “Phonological Alternations in Tangut.” In *Zhongyang yanjiu lishi yuyuan yanjiusuo jikan* [Bulletin of the Institute of History and Philology] 中央研究院歷史語言研究所集刊, 59.3 (1988), pp. 783–834.

- Gong 1993 — Gong Hwang-cherng 龔煌城. “Xi-Xia yu yu qiangyuzhi yuyan tongyuanci de lishi cengci” [Cognate Historical Strata of Tangut and the Qiangic Languages] 西夏語與羌語支語言同源的歷史層次. Unpublished manuscript, 1993.
- Gong 1995 — Gong Hwang-cherng 龔煌城. “The System of Finals in Proto-Sino-Tibetan.” In *The Ancestry of the Chinese Language*. Ed. by William S.Y. Wang. Berkeley: Project on Linguistic Analysis, University of California, 1995, pp. 41–92.
- Gong 1999 — Gong Hwang-cherng 龔煌城. “Xi-Xia yu de jin muyin ji qi qi yuan” [Tangut Tense Vowels and Their Origin] 西夏語的緊母音及其起源. In *Zhongyang yanjiu lishi yuyuan yanjiusuo jikan* [Bulletin of the Institute of History and Philology] 中央研究院歷史語言研究所集刊, 70.2 (1999), pp. 531–558.
- Gong 2003 — Gong Hwang-cherng 龔煌城. “Tangut.” In *The Sino-Tibetan Languages*. Ed. by Graham Thurgood and Randy J. LaPolla. London: Routledge, 2003, pp. 602–620.
- Hill 2010 — Hill Nathan W. “An Overview of Old Tibetan Synchronic Phonology.” In *Transactions of the Philological Society*, 108.2 (2010), pp. 110–125.
- Jacques 2004 — Jacques G. “Phonologie et morphologie du japhug (rGyalrong).” PhD diss. Université Paris Diderot, 2004.
- Jacques 2006 — Jacques G. “Essai de comparaison des rimes du tangoute et du rgyalrong.” In *Medieval Tibeto-Burman Languages II — 10th Seminar of the International Association for Tibetan Studies*. Ed. by Ch. Beckwith. Leiden: Brill, 2006.
- Jacques 2009 — Jacques G. “The Origin of Vowel Alternations in the Tangut Verb.” In *Language and Linguistics*, 10.1 (2009), pp. 17–27.
- Kuchanov and Sofronov 1963 — Кычанов Е.И., Софронов М.В. *Исследования по фонетике тангутского языка (предварительные результаты)*. М.: Издательство восточной литературы, 1963.
- Lee and Ramsey 2011 — Lee Ki-Moon 李基文, Ramsey S.R. *A History of the Korean Language*. Cambridge: Cambridge University Press, 2011.
- Li 2008 — Li Fanwen 李範文. *Xia-han zidian* [Tangut-Chinese Dictionary] 夏漢字典. 2nd edition. Beijing: Zhongguo shehui kexue chubanshe 北京: 中國社會科學出版社, 2008.
- Martin 1992 — Martin S.E. *A Reference Grammar of Korean*. Rutland: Charles E. Tuttle, 1992.
- Matisoff 2004 — Matisoff J.A. “‘Brightening’ and the place of Xi-Xia (Tangut) in the Qiangic Subgroup of Tibeto-Burman.” In *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-cherng Gong on His 70th Birthday*. Ed. by Ho Dah-an. Taipei: Institute of Linguistics, Academia Sinica, 2004, pp. 327–352.
- Miyake 2008 — Miyake Marc Hideo. “Avoiding ABBA: Old Chinese Syllabic Harmony.” In *Evidence and Counter-Evidence: Essays in Honour of Frederik Kortlandt. Vol. 2. General Linguistics*. Ed. by A. Lubovsky, J. Schaeken and J. Wiedenhof. Amsterdam: Rodopi, 2008, pp. 283–301.
- Nishida 1964 — Nishida Tatsuo 西田龍雄. *Seikago no kenkyū* [A Study of the Hsi-Hsia Language] 西夏語の研究. Tokyo: Zauhō kankōkai 東京: 座右實刊行會, 1964.
- Sagart 1999 — Sagart L. *The Roots of Old Chinese*. Amsterdam: John Benjamins, 1999.
- Schuessler 2007 — Schuessler A. *ABC Etymological Dictionary of Old Chinese*. Honolulu: University of Hawai‘i Press, 2007.
- Schuessler 2009 — Schuessler A. *Minimal Old Chinese and Later Han Chinese: A Companion to Grammata Serica Recensa*. Honolulu: University of Hawai‘i Press, 2009.
- Sun Hongkai 1981 — Sun Hongkai 孙宏开. *Qiangyu jianzhi* 羌语简志 [A Brief Description of the Qiang Language]. Beijing: Minzu chubanshe 北京: 民族出版社, 1981.
- Sun Jackson 2003 — Sun Jackson. “Phonological Profile of Zhongu: A New Tibetan Dialect of Northern Sichuan.” In *Language and Linguistics*, 4.4 (2003), pp. 769–836.

- Tai 2008 — Tai Chung Pui 戴忠沛. “Xi-Xia wen fojing canpian de zangwen duiyin yanjiu” [A Study of Tibetan Phonological Transcription in Tangut Buddhism Fragments] 西夏文佛經殘片的藏文對音研究. PhD diss. Graduate School of Chinese Academy of Social Sciences, 2008.
- Vovin 2010 — Vovin A. *Koreo-Japonica: A Re-Evaluation of a Common Genetic Origin*. Honolulu: University of Hawai‘i Press, 2010.