

Yunfan Lai*, Xun Gong, Jesse P. Gates and Guillaume Jacques

Tangut as a West Gyalrongic language

<https://doi.org/10.1515/flih-2020-0006>

Received April 10, 2019; revision invited November 12, 2019; revision received January 27, 2020; accepted March 24, 2020

Abstract: This paper proposes that Tangut should be classified as a West Gyalrongic language in the Sino-Tibetan/Trans-Himalayan family. We examine lexical commonalities, case marking, partial reduplication, and verbal morphology in Tangut and in modern West Gyalrongic languages, and point out nontrivial shared innovations between Tangut and modern West Gyalrongic languages. The analysis suggests a closer genetic relationship between Tangut and Modern West Gyalrongic than between Tangut and Modern East Gyalrongic. This paper is the first study that tackles the exact linguistic affiliation of the Tangut language based on the comparative method.

Keywords: Tangut, West Gyalrongic, Sino-Tibetan, subgrouping, linguistic phylogeny

1 Introduction

Tangut is an extinct language that was spoken in the Tangut Empire (1038–1227), which covered mainly today's Ningxia and Gansu Provinces in China. Since its rediscovery in the late nineteenth century, many publications dedicated to the decoding of Tangut's complex orthography and to its phonology and morphology have seen the light of day. However, its linguistic affiliation is still poorly understood. While the Sino-Tibetan affiliation of Tangut was understood from the very beginning of its decipherment, scholars are divided on its precise affiliation within Sino-Tibetan. Laufer (1916), by comparing Tangut with various Lolo-Burmese languages, believed Tangut to be closely related to Lolo-Burmese.

***Corresponding author: Yunfan Lai**, Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History, Jena, Germany, E-mail: khroskyabs@gmail.com

Xun Gong, Department of East Asian Languages and Cultures, School of Oriental and African Studies, University of London, London, UK, E-mail: minus273cn@gmail.com

Jesse P. Gates, Centre de Recherches Linguistiques sur l'Asie Orientale, École des Hautes Études en Sciences Sociales, Paris, France, E-mail: stauskad@gmail.com

Guillaume Jacques, Centre de Recherches Linguistiques sur l'Asie Orientale, École des Hautes Études en Sciences Sociales, Paris, France, E-mail: rgyalrongskad@gmail.com

However, many of the relevant examples he found also have cognates in Gyalrongic languages, which rule out his correspondences for the purposes of subgrouping. Because of the lack of data and the insufficient understanding of Sino-Tibetan historical linguistics in Laufer's day, the comparative work of Laufer is no longer useful in answering the question of Tangut's genetic position. The Lolo-Burmese affiliation of Tangut is partly held by Nishida (1976), who compares Tangut with Tosu, a Qiangic language spoken in Sichuan. Nishida proposes that Tangut was a language with a Lolo-Burmese stratum which eventually became influenced by Qiangic languages due to contact.

Wolfenden (1931), using Tibetan transcriptions of Tangut characters, compared twenty-four lexical examples in Tangut mainly with Tibetan and Gyalrong, and stated that it was possible that Tangut was related to the ancestor of the Gyalrong language. Wang (1933) deemed Tangut to be closely related to Muya and Qiang, and this opinion is continued by Sun (1991) and Matisoff (2004). Sun (1991) compares Tangut with more than twenty "Tibeto-Burman" languages, and finds that Tangut is allied to those known as Qiangic languages. However, his comparison was based on Chinese transcriptions of Tangut, which hide a great amount of phonological information. Matisoff (2004) states that the phenomenon of "brightening" (i.e. a sound change from *a* to a high vowel, such as *i* and *e*) is a shared innovation of Qiangic, and because Tangut underwent this sound change, it should be classified as a Qiangic language together with "Tibeto-Burman" languages mainly spoken in Sichuan, such as Gyalrong, Stau, Pumi, nDrapa, Muya, etc. However "brightening" is arguably recent in Gyalrongic languages, since loanwords from Tibetan also undergo this change (Gong 2018: 23–27), and therefore is at least in part a parallel innovation across non-Tibetan languages of the area.

Two articles have been recently published on Sino-Tibetan phylogeny using Bayesian methods, Zhang et al. (2019) and Sagart et al. (2019), but only the latter included Tangut in its language sample. According to Sagart et al. (2019)'s results, Tangut belongs to the Gyalrongic clade, but its exact place within this clade is less clear: it is linked with West Gyalrongic with a posterior probability of only 0.51.

The purpose of this paper is to investigate with more data the precise relationship between Tangut and Gyalrongic languages, validating it as a West Gyalrongic language on the basis of data from four domains: lexicon (Section 4), case marking (Section 5), partial reduplication (Section 6) and verbal morphology (Section 7). Before the conclusion, Section 8 speculates about the migration of the Tanguts.

2 Tangut and Pre-Tangut reconstructions used in this paper

Before going into the details, it is necessary to present the Tangut and Pre-Tangut reconstructions used throughout this paper.

A slightly modified version of Gong Hwang-Cherng's Tangut reconstruction of Tangut is selected in this paper (Gong 2002). While the reconstruction can mostly be interpreted with the International Phonetic Alphabet (IPA), there are several major non-IPA characteristics. In terms of vowels, an underdot stands for a tense vowel, e.g. *ɑ̣* stands for a tense realisation of [a]; a post-vocalic *-r* refers to the retroflexion of the preceding vowel, e.g. *ar* stands for [a̠]; and a long vowel is represented by the doubling of the vowel in question, e.g. *aa* stands for [a:]. In terms of consonants, a period in the initial position stands for a glottal stop initial, e.g. *ɑ̣* is equivalent to [ʔa]; *lh-* stands for *l-*; and the fricatives and affricates with *ś* and *ź* in Gong's original reconstruction are written *ç* and *ž* respectively, so as to facilitate the comparison with Gyalrongic languages. The two tones are marked with superscripted numerals, "1" for the plain tone, and "2" for the rising tone, after the segmental form: *ɑ̣*¹ (plain tone) and *ɑ̣*² (rising tone).

All the Tangut reconstructions are accompanied with Tangut characters followed by a four-digit number corresponding to those in Li Fanwen's dictionary (Li 1997). For example, the character 𐞪³¹⁷⁷ *kur*² 'ice, freeze' can be found in Li's dictionary under the number 3177.

Jacques (2014) reconstructed the phonology of Pre-Tangut. By taking attested Tangut forms and reversing sound changes that he had identified, Jacques was able to generate Pre-Tangut forms that could be more easily compared to forms in other Sino-Tibetan languages. Gong (2020) reconstructs the distinction of uvularization for Tangut and Pre-Tangut, but a reconstruction of Pre-Tangut that accounts for uvularization remains to be worked out in full. In the present article, the Pre-Tangut system of Jacques (2014) is retained as is. Cases where the question of uvularization is relevant receive a separate discussion.

Reconstructed forms of Pre-Tangut are systematically presented after Tangut reconstructions in the comparisons made in this paper. Standard IPA symbols are generally used in the Pre-Tangut system, except for two uppercase letters that stand for multiple possibilities of reconstruction: **C-* stands for **k-* or **t-* and **S-* stands for **s-*, **l-*, **ç-*.

3 Gyalrongic languages (before this paper)

The Gyalrongic branch was first proposed by Sun (2000a, 2000b), based on morphosyntactic parallelisms among its three subgroups: East Gyalrongic languages (then termed simply ‘Gyalrong’) including Situ, Japhug, Tshobdun and Zbu; Khroskyabs (then termed ‘Lavrung’); and Stau (then termed ‘Horpa’).¹ Gyalrongic languages are spoken in seven counties of southwestern Rngaba (Aba) Tibetan and Qiang Autonomous Prefecture (Li, Heishui, Ma’erkang, Xiaojin, Jinchuan, Aba, and Rangtang Counties) and four counties of central and eastern Dkarmdzes (Ganzi) Tibetan Autonomous Prefecture (Danba, Daofu, Luhuo, and Xinlong Counties) of Sichuan Province, China. These languages share similar tense marking strategies, ablauting patterns and transitivity markings which are unknown in other Qiangic languages. Recent years have seen various discussions on the relation between Khroskyabs and Stau, showing their morphological and lexical commonalities not shared by East Gyalrongic languages, such as the generalization of inverse marking into all 3 → 3 scenarios in Khroskyabs and Stau (Lai 2017; Jacques et al. 2017). Scholars are increasingly convinced that these two varieties should form a single node under the Gyalrongic branch, termed *West Gyalrongic* (Jacques et al. 2017: 609–611; Lai 2017: 10; Gong 2018: 21; Gates 2012: 14 as ‘Western Gyalrongic’). The *Stammbaum* of Gyalrongic is illustrated in Figure 1.

Note that ‘East Gyalrongic’ is not to be confused with ‘Eastern Gyalrong’ or ‘East Gyalrong’, both referring to Situ only and not including Japhug, Tshobdun, and Zbu. The terms ‘Eastern Gyalrong’ or ‘East Gyalrong’ come from Lin (1993), who called Situ the ‘East Gyalrong dialect’ (嘉戎语东部方言). Besides Situ, our term East Gyalrongic is inclusive of Japhug, Tshobdun, and Zbu, three languages that are often written about as ‘Northern Gyalrong’ in the literature.

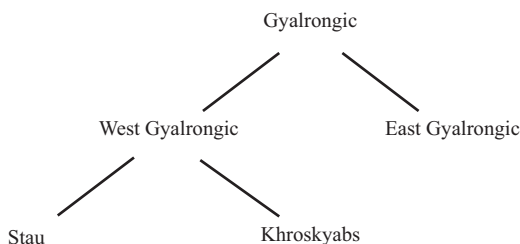


Figure 1: *Stammbaum* of Gyalrongic.

¹ In fact, the term was ‘Horpa-Shangzhai’, which is actually a complex of several languages that include Stodste and Geshiza. Stau (also called Rtau, Horpa, and Daofuhua) is just one of these language varieties, but for the purposes of this paper Stau suffices as a representative of the ‘Horpa-Shangzhai’ language complex.

This paper is based on first-hand data of at least six Gyalrongic varieties, including two East Gyalrongic languages (Japhug and Zbu), and four West Gyalrongic varieties, Wobzi and Siyuewu Khroskyabs, and Khang.gsar and Mazur Stau.

4 Lexicon

In this section, based on the work of Jacques (2014), we first compare common lexical items between Tangut and other Gyalrongic languages in Section 4.1, then we list cognates exclusively found between Tangut and modern West Gyalrongic languages in Section 4.2.

4.1 Gyalrongic lexical commonalities

Jacques (2014) includes 321 cognates shared between East Gyalrongic and Tangut; out of these 321 cognates, 309 lexical items have West Gyalrongic (Khroskyabs or Stau) cognates or have cognates in other Sino-Tibetan languages. Only 12 lexical items lack convincing cognates from other languages of the family. However, one should not immediately consider these 12 lexical items as shared innovations between East Gyalrongic and Tangut for the following reasons.

1. First, some comparisons are judged uncertain in Jacques (2014). For example, 𐞗𐞙⁰³⁶⁹ *t^hjuu*¹ ‘inspect’ is possibly related to Japhug *t^hu* ‘ask’, but there is not enough comparative evidence to be sure.
2. Second, some comparisons do not have clear textual examples in Tangut, therefore the glosses are exclusively based on dictionaries, which can be misleading. Examples of this type include 𐞗𐞙⁵¹⁴³ *kje*¹ ‘insult’ (Japhug *nymqe* ‘insult’) and 𐞗𐞙²⁵⁵² *tchia*¹ ‘draw, remove’ (Japhug *tx̣t* ‘draw, remove’).
3. Third, examples which appear as common innovations between Tangut and East Gyalrongic because they are not found in West Gyalrongic, could have existed in West Gyalrongic and later could have been replaced by loanwords from Tibetan or Chinese. For example, for the Tangut word 𐞗𐞙⁰⁴⁷⁵ *lju*¹ ‘mat’ (Japhug *-βju* ‘mat’), Khroskyabs employs a Tibetan loan word *stæn* ‘mat, carpet’, from Tibetan *stan* of the same meaning; for 𐞗𐞙⁴⁰⁹² *khie*¹ ‘hate’ (Japhug *q^ha* ‘hate’), Khroskyabs uses the negative form of the Tibetan loan word *ndæ* (Wobzi) or *ndôd* (Siyuewu) ‘love’ (from Tibetan འདོད *’dod* ‘love’, *mə-ndæ* (NEG-love)); for 𐞗𐞙²⁴³⁶ *mjaa*¹ ‘fruit’ (Japhug *sumat* ‘fruit’), Khroskyabs uses the Tibetan loan word *cânṭoṅ* (Wobzi) and *cânṭoṅ* (Siyuewu) ‘fruit’, from Tibetan ཤིང་ཐོག *shing.tog* ‘fruit’; and for 𐞗𐞙⁵⁹³² *mə*² ‘kind, species’ (Japhug *(ɛmuuz-)-muuz* ‘two kinds’), Khroskyabs

uses *rây* (in both Wobzi and Siyuewu), which is a borrowing of རྩལ *rigs* ‘species, kind’ from Tibetan.

Only one example, 𐰚⁰¹⁴⁸ *tq¹* ‘coagulate’ (Japhug *stxm* ‘coagulate’), could qualify as candidates of shared innovations between East Gyalrongic and Tangut. However, having only one example is rather insignificant when compared to the overwhelming evidence for classifying Tangut as a West Gyalrongic language, as will be shown in this paper.

We do not present the exhaustive list of Gyalrongic commonalities here. Instead, we list several newly discovered cognates in this section.

1. 𐰚⁰⁵⁸⁵ *çj²* ‘grass’ < Pre-Tangut **çje*, Japhug *xçaj* ‘grass’, Wobzi and Siyuewu Khroskyabs *χçj* ‘grass’.
2. 𐰚⁰²⁶⁸ *jir²* ‘ask’ < Pre-Tangut **r-jæt*, Zbu *rgwât* ‘ask’, Wobzi Khroskyabs *ryæ* ‘ask’, Siyuewu Khroskyabs *ryæd* ‘ask’, Tshobdun *rwa* ‘ask’, Mazur Stau *rje* ‘ask’. The retroflex vowel in Tangut reflects the *r-* in the Gyalrongic cognates (Jacques 2014: 23–29). For the vowel correspondence, cf. 𐰚²³⁶⁷ *tshj¹* ‘goat’, Zbu *tsutâ* ‘goat’, Wobzi Khroskyabs *ts^hæ* ‘goat’, Siyuewu Khroskyabs *ts^hæd* ‘goat’, Mazur Stau *ts^hε* ‘goat’.
3. 𐰚³⁵⁴⁷ *lia²* ‘drunk’, Mazur Stau *je* ‘drunk’, Zbu *ldâ* ‘drunk’. The phonological form of the Tangut word is highly uncommon. Tangut Grade II probably corresponds with the Zbu velarized vowel *a*.²
4. 𐰚⁵³²¹ *nej²* ‘knead’ < Pre-Tangut **r-nej*, Bragbar Situ *mêj* ‘knead’, Zbu *lni* ‘knead’, Wobzi Khroskyabs *lnè* ‘knead’. Pre-Tangut **r-* is lost before Tangut *n-* (Jacques 2014: 25) and did not leave a trace as a Tangut retroflex syllable.
5. 𐰚⁵³⁷⁰ *pja¹* ‘palm’ < Pre-Tangut **s-pjat*, Bragbar Situ *nə-çphiēt* ‘slap on the face’, Wobzi Khroskyabs *djæpjâ* ‘palm’, Siyuewu Khroskyabs *dæpjâ* ‘palm’. The Wobzi final vowel *æ* in most cases comes from a closed syllable. In this case, it is most likely derived from an earlier **æt*. However, the absence of a coda in Siyuewu is irregular and unexplained.
6. 𐰚⁵⁸⁰⁰ *tijj²* ‘block’ < Pre-Tangut **s-tij*, Japhug *sti* ‘block’, Bragbar Situ *çtiē* ‘block’, Zbu *stâ* ‘block’, Wobzi Khroskyabs *sté* ‘block’, Siyuewu Khroskyabs *fstéd* ‘block’. This comparison must be regarded as tentative, as the Tangut rime *jij/aj* usually corresponds to *-Vm* in other Burmo-Qianguic languages.
7. 𐰚⁰⁰¹¹ *lo^o* ‘rich’ < Pre-Tangut **C-tVm*, Zbu *tâm* ‘rich’, Wobzi Khroskyabs *jdâm* ‘rich’, Siyuewu Khroskyabs *ydâm* ‘rich’. For the vowel correspondence cf. 𐰚²⁹¹⁵ *no¹* ‘rib’, Zbu *ta-mâm?* ‘rib’, Wobzi Khroskyabs *mâmçærâ* ‘rib’.

² According to Gong Hwang-Cherng’s reconstruction (Gong 1994), rimes are divided into twelve rime classes (攝 *shè*), and each class contains three grades (等 *děng*) of rimes: Grades I, II, and III. Grade I represents plain vowels, Grade II vowels with the medial *-i-*, Grades III vowels with the medial *-j-*.

8. 𐞧𐞦⁵⁸¹¹ *rjir*² ‘juice, soup’ < Pre-Tangut **rje*, Zbu -*rī?* ‘soup’, Wobzi and Siyuewu Khroskyabs *ré* ‘soup’. In East Gyalrongic other than Zbu, the term for ‘soup’ is based on the term for ‘water’, such as Japhug *ɛrci*, and Bragbar *ɕaŋi*, literally meaning ‘meat water’.

4.2 West Gyalrongic lexical commonalities

In this section, we consider West Gyalrongic and Tangut lexical commonalities from three aspects. First, in Section 4.2.1, we analyze cognates that are found in other Gyalrongic or Sino-Tibetan languages, which however underwent idiosyncratic changes in West Gyalrongic. Then in Section 4.2.2, we discuss shared monosyllabic innovations between Modern West Gyalrongic and Tangut, and finally in Section 4.2.3, we examine shared bisyllabic innovations.

We have ensured that the West-Gyalrongic and Tangut lexical commonalities listed in this section are indeed shared innovations, and we have excluded some clear retentions. For example, Tangut 𐞧𐞦²⁴⁰⁷ *dzeej*² ‘ride’ corresponds with Wobzi and Siyuewu Khroskyabs *rtsə* ‘ride’, and is without a cognate in East Gyalrongic languages. However, Tangut 𐞧𐞦²⁴⁰⁷ *dzeej*² ‘ride’ is comparable with Proto-Lolo-Burmese *dzi*² 0651 (Jacques 2014: 172), therefore, this word is not included.

4.2.1 “Westernized” Gyalrongic cognates

1. 𐞧𐞦³⁹¹² *bji*² ‘step’ < Pre-Tangut **mbja*, Wobzi Khroskyabs *blī* ‘step’, Siyuewu Khroskyabs *blīy* ‘step’. This noun is probably related to Japhug -*mglə*, Zbu -*mglət* ‘step’. The initial of the Tangut *b*- < **mb*- agrees with Khroskyabs *b*, and hence implies a West Gyalrongic phonological innovation **mb*l- < **mgl*-.
2. 𐞧𐞦⁰⁹⁶⁰ *mji*¹ ‘daughter’ < Pre-Tangut **S-mjej*, Wobzi Khroskyabs *vəsmé* ‘wife’, Mazur and Khang.gsar Stau *smi* ‘woman’. The tense vowel -*i* in Tangut indicates a Pre-Tangut preinitial **S*-, which is not found in East Gyalrongic languages: Japhug *tu-me* ‘girl’, Zbu *tə-mī* ‘woman’. Note that East Gyalrongic languages regularly preserve the preinitial **s*-. The Pre-Tangut **S*- finds its parallel, however, in Khroskyabs and Stau.
3. 𐞧𐞦⁰⁷⁶⁴ *rji*¹ ‘horse’ < Pre-Tangut **rjaŋ*, Mazur and Khang.gsar Stau *rji* ‘horse’. The most widespread etymon for ‘horse’ has a form akin to Burmese *mraŋ*³, cf. Japhug *mbro* ‘horse’ < **aŋ*. The rime of Tangut -*jij* (Tangut) and Stau -*i* point to a proto-rime **-aŋ*. The Tangut and Stau forms could reflect an irregular loss of the preinitial **m*-, or, less likely, could reflect a separate version of ‘horse’, which is at any rate a Wanderwort, cf. Chinese 馬 (Baxter-Sagart) **m̥ra?* or Korean *mar*. The Khroskyabs form, Wobzi and Siyuewu Khroskyabs *bró* ‘horse’, must be a

loan word from East Gyalrongic, as we would expect *-e* as the vowel: cf. Wobzi Khroskyabs *brê* ‘tall’, corresponding to Japhug *mbro* ‘tall’, which shows the regular reflex of **-aŋ*.

4. 藏⁰⁰²⁰ *tɕja*¹ ‘road’, Mazur and Khang.gsar Stau *tɕɛ* ‘road’, Wobzi Khroskyabs *tɕ^hæ* ‘road’, the construct state of *tɕ^hi* ‘road’, attested in the incorporational form *ntɕ^hæfse* ‘guide’ (*fsê* ‘lead’).³ The form *ntɕ^hæfse* can be analytically used, as shown in (1). The Gyalrongic cognates, Japhug *tɕu*, Khroskyabs *tɕ^hi*, do not correspond well to the Tangut form, which *prima facie* calls for Pre-Tangut **tsrjaC*, as Khroskyabs *-i* generally corresponds to Japhug *-a* and to Tangut *-i*. In this case, the *-a* rime in 藏⁰⁰²⁰ *tɕja*¹ ‘road’ is similar to the isolation of the construct state *tɕ^hæ* in Wobzi Khroskyabs.

- (1) *æ̣t̪ə=ɣə ŋ̣æ̣=ji tɕ^hæ̣ n-u-fsé*
 3SG=ERG 1SG=GEN road PST-INV-lead₂
 ‘He guided me.’

5. The verb 藏⁵⁴⁴⁹ *tji*¹ ‘to put’ (Pre-Tangut **S-tja*, stem B 藏⁵⁴⁴⁹ *tjɔ*¹) finds an exact cognate in Wobzi Khroskyabs *stɪ* ‘put’ (Lai 2017). The Stau inanimate non-stationary existential copula *sti* could be the de-transitivized version of this verb. The *-ji/-jo* stem alternation is indicative of an ancient **a* (see Jacques 2014: 231–232).

The verb 藏⁵⁴⁴⁹ *tji*¹ ‘to put’ and its Khroskyabs and Stau cognates are related to a root widespread in the Trans-Himalayan family, but in languages other than West Gyalrongic, the corresponding root has a simple stop initial, as in Japhug *ta* ‘put’. The presence of a *st-* cluster in Khroskyabs and Stau corresponds to tense voice in Tangut, indicative of a **S-t-* cluster (Gong 1999). This unexpected *s-* must be a lexicalized prefix. A potential explanation for it would be to analyze it as a trace of the cognate of the translocative associated motion prefix *ɕu-* prefix found in Japhug and other modern East Gyalrongic languages (see Jacques 2013). In Japhug texts, this prefix does occur in direct contact with the verb root in the non-past factual form, as in (2).

- (2) *nutɕu fsaŋ ɕu-ta-nu ra.*
 there fumigation TRANSLOC-put:FACT-PL have.to:FACT
 ‘They have to go and put juniper fumigation offerings there.’

³ Incorporation is a particular type of compounding in which a V and an N combine to form a new V (Mithun 1984). In Gyalrongic languages, most incorporational constructions are a combination of a denominal prefix, a modified version of a noun (construct state), and a verb.

Since this form is relatively common, it is possible that the translocative prefix was reanalyzed as part of the root in Pre-Tangut. Other possibilities could be taken into consideration (an incorporated noun for instance, as this phenomenon is well-attested in Khroskyabs and Tangut, see Jacques 2011), but in any case, this verb is a good example of an innovative feature that is unlikely to have independently arisen twice.

4.2.2 West Gyalrongic monosyllabic innovations

1. 𐞧⁵³¹³ *kq*¹ ‘pull up’ < Pre-Tangut **S-kaC*, Wobzi Khroskyabs *qâ* ‘pull’. The Wobzi vowel *a* can always be traced back to an earlier **æ*, except in loanwords. Tangut Grade I (Ø) corresponds with the uvular initial in Gyalrongic languages.
2. 𐞧¹²⁷⁸ *ji*² ‘say’ < Pre-Tangut **ji*, Siyuewu Khroskyabs *jû* ‘say’, Mazur and Khang.gsar Stau *jə* ‘say’. East Gyalrongic has unrelated terms for this concept: *ku-ti* in Japhug, *ka-tsê* in Bragbar Situ, and *kæ-ts^hô* in Zbu.
3. 𐞧⁰⁷⁵⁶ *dzju*² ‘meet’ < Pre-Tangut **ndzo*, Mazur and Khang.gsar Stau *dzə* ‘meet’. Only unrelated terms are found in East Gyalrongic, e.g. Japhug *atuɣ*, Bragbar Situ *o-mə-mtê*.
4. 𐞧¹²⁷⁴ *wo*² ‘able’ < Pre-Tangut **wan*, Wobzi Khroskyabs *yó* ‘able’. For the rime correspondence cf. 𐞧³⁴⁴³ *po*¹ ‘paternal uncle’, Wobzi Khroskyabs *əyô* ‘uncle’. In East Gyalrongic, the word for ‘able’ is generally the cognate of Japhug *c^ha*, such as Bragbar Situ *c^hâ* and Zbu *kə-c^hê*.
5. 𐞧⁰²¹⁴ *lu*² ‘poor’ < Pre-Tangut **C-to*, Wobzi Khroskyabs *jdô* ‘poor’, Siyuewu Khroskyabs *ydô* ‘poor’. The correspondence between Tangut *l-* and Khroskyabs *d-* can be also found in 𐞧⁰⁶³⁰ *la* ‘weave’ :: Wobzi Khroskyabs *dây* ‘weave’, 𐞧⁵⁸⁴⁵ *lwə* ‘buy’ :: Wobzi Khroskyabs *jdê* ‘buy’. In East Gyalrongic, unrelated terms for ‘poor’ are used: Japhug *ŋgu*, Zbu *kə-ŋgó*, and Bragbar Situ *tciê*.
6. 𐞧⁵³⁹⁰ *phie*² ‘untie’ < Pre-Tangut **phra*, Wobzi and Siyuewu Khroskyabs *p^hrê* ‘untie’, Mazur Stau *p^hrə* ‘untie’. East Gyalrongic has different unrelated terms for this concept, such as Japhug *rla* and Bragbar Situ *rdiê*.
7. 𐞧³²⁸⁷ 𐞧⁵⁵⁶¹ *rur*¹ *rar*² ‘narrow’, Wobzi Khroskyabs *ɛrô* ‘narrow’. The first syllable < Pre-Tangut *ro* shows uvularization that can be attributed to a compression of the preinitial *ɣ-* in Khroskyabs (Gong 2020). The second syllable reflects the partial reduplication of the West Gyalrongic type (Section 6). In East Gyalrongic, the words for ‘narrow’ are Japhug *tɛɣr*, Bragbar *tɛor* and Zbu *kə-tciɛ*.
8. 𐞧⁵⁸⁰⁹ *lu*¹ ‘wet’ < Pre-Tangut **lo*, Wobzi Khroskyabs *lú* ‘wet’, Siyuewu Khroskyabs *luy* ‘wet’. According to Gong (2020), the rime of this example should be reconstructed with a uvularized vowel, namely **-u^ɣ*. This uvularization might

be reflected in the coda -ɣ in Siyuewu Khroskyabs. In East Gyalrongic, Japhug *aci* ‘wet’ is a denominal form of the noun *tu-ci* ‘water’, and in Bragbar Situ the word is *ɕtɕit*, which is unanalyzable.

9. 𐰚⁵⁹⁷⁸ *tər*¹ ‘tread, stamp’ < Pre-Tangut **r-tu*, Siyuewu Khroskyabs *ntərtá* ‘tread, stamp’. A different etymon is found in East Gyalrongic with Bragbar Situ *ra-tɕiāk* and Japhug *ɣtɕaɕ*.
10. 𐰚⁴⁸⁸⁹ *dzjwi*¹ ‘neighbour, each other’ < Pre-Tangut **p-ndzu*, Wobzi and Siyuewu Khroskyabs *vdzə* ‘fellow, mate’, Mazur and Khang.gsar Stau *vdzə* ‘friend’. Compare the compound 𐰚⁴⁹⁸⁹ 𐰚⁴⁸⁸⁹ *.wji*¹ *dzjwi*¹ ‘friend’, the first morpheme being 𐰚⁴⁹⁸⁹ *.wji*¹ ‘friend, relative’. East Gyalrongic uses different terms for ‘neighbour’ or ‘friend, fellow’: Japhug *tx-ɣa* ‘neighbour’, *ɣuɣsu* ‘friend’; Bragbar Situ *ta-jwāk* ‘neighbour’, *ta-rge-psô* ‘friend’.

4.2.3 West Gyalrongic bisyllabic innovations

Bisyllabic words are of great importance for the classification of Tangut. It is possible for similarities in monosyllabic comparisons to have occurred by chance; however, the probability of bisyllabic comparisons to be coincidences is much lower, that is, they are more likely to be shared innovations.

This section is divided into three parts. First, bisyllabic words in which both syllables are unique to West Gyalrongic and Tangut, as well as unanalysable bisyllabic words. Second, bisyllabic words in which one morpheme is unique to West Gyalrongic and Tangut. Third, bisyllabic words in which no morpheme is unique to West Gyalrongic and Tangut, but the combination of the two morphemes are unique.

Unanalyzable bisyllabic words/Words with two innovated morphemes

1. 𐰚²⁴⁵⁵ 𐰚²¹²⁹ *gji*² *bji*² ‘wife’ < Pre-Tangut **ŋgje.mbjɛj*, Siyuewu Khroskyabs *gəvə* ‘wife’. This example is unanalyzable. A monosyllabic etymon, widely attested in all Gyalrongic languages, is used for this concept in East Gyalrongic, Japhug *tx-rzaβ*, Bragbar Situ *ta-ɣɛp*, Zbu *tə-ɣɛv*.
2. 𐰚⁴⁰³⁶ 𐰚³³³² *dziwə*¹ *dziwe*¹ ‘pull’ < Pre-Tangut **ndzrwi-ndzrwa*, Wobzi Khroskyabs *dzədzə* ‘pull’, Mazur and Khang.gsar Stau *ndzədzə* ‘pull’. This example is a reduplication of a simplex verb stem which is absent outside West Gyalrongic. Different unrelated words for similar meanings are used in East Gyalrongic, such as Japhug *ɣxkhuukhrut* ‘drag’ and Bragbar Situ *racə* ‘pull, drag’, etc.
3. 𐰚⁵³⁸² 𐰚⁰⁵⁷⁵ *mji*² *rjar*² ‘footprint’ < Pre-Tangut **mji.rjat*, 𐰚⁵⁷³³ 𐰚⁰⁵⁷⁵ *mji*² *rjar*² ‘follow the footprints’ < Pre-Tangut **S-mji.rjat*, Siyuewu Khroskyabs *moríd* ‘(animal’s) footprint’, *smoríd* ‘follow the footprints’. This comparison is of great interest

because not only is there a correspondence of the base noun, but also there are parallel denominal forms in both Tangut and in Khroskyabs. The tense vowel *-i* in 𐞗⁵⁷³³𐞗⁰⁵⁷⁵ *mji²rjar²* ‘follow the footprints’ indicates a proto-preinitial, corresponding to the *s*-denominal prefix in Siyuewu Khroskyabs.

4. 𐞗³⁶⁹⁶𐞗⁵¹²⁴ *na¹.rar²* ‘tomorrow’, Wobzi Khroskyabs *nâri/nâru* ‘tomorrow’, Siyuewu Khroskyabs *nâri* ‘tomorrow’. The Pre-Tangut form could be reconstructed as **naC.raC*. The Pre-Tangut rime **-aC* corresponds to Khroskyabs *-a*, as in Wobzi Khroskyabs *lalâ* ‘maternal aunt’, comparable to Tangut 𐞗⁰⁰⁰⁴ *la²* ‘maternal aunt’ (Pre-Tangut **lak*). The second syllable, **raC*, is a little problematic. There are two free variants for ‘tomorrow’ in Wobzi Khroskyabs, *nâri* and *nâru*, which are clearly related. There is evidence that the form *nâri* is more likely to be the base form, and *nâru* the modified form. A good number of Khroskyabs *-i* rimes correspond to Pre-Tangut **-a*, and it is plausible that it is also the case for the one in *nâri*. The form *nâru* could be the combination of *nâri* and an additional ending **-C* that modifies the vocalism, hence **nâriC* > *nâru*. Therefore, the Tangut word 𐞗³⁶⁹⁶𐞗⁵¹²⁴ *na¹.rar²* ‘tomorrow’ is probably the exact cognate of *nâru* in Wobzi Khroskyabs. East Gyalrongic has completely different terms for ‘tomorrow’: Japhug *fso* and Bragbar Situ *soçnî*, Zbu *səfsîz*.

Words with one innovated morpheme

1. 𐞗³⁷³²𐞗¹³¹² *kə¹ta¹* ‘puppy, small dog’ < Pre-Tangut **ku.taC*, Wobzi and Siyuewu Khroskyabs *kəta* ‘dog’, Mazur Stau *k^hətə* ‘dog’, Khang.gsar Stau *kəta* ‘dog’. The morpheme 𐞗³⁷³² *kə¹* is probably comparable with Tibetan 𐍌³⁷³² *khyi* ‘dog’, as well as the first elements of Japhug *k^huna* ‘dog’ and of Bragbar Situ *k^həniē* ‘dog’. The second morpheme 𐞗¹³¹² *ta¹* is uniquely found in West Gyalrongic.
2. 𐞗⁵⁶⁴³𐞗³⁰⁹² *mji¹dji²* ‘except, apart from’ < Pre-Tangut **mji-ndjej*, Wobzi Khroskyabs *mənə* ‘unless, apart from’, Siyuewu Khroskyabs *məndi* ‘unless, apart from’, Mazur Stau *mde* ‘except, apart from, otherwise’. While the first morpheme, 𐞗⁵⁶⁴³ *mji¹*, could be a negative marker widely shared by Sino-Tibetan languages, the second morpheme, 𐞗³⁰⁹² *dji²*, is most likely an innovation. The examples (3a), (3b), and (3c) show similar uses of this etymon in Tangut, Siyuewu Khroskyabs, and Mazur Stau, respectively.

- (3) a. 𐞗²⁰¹⁹ 𐞗⁵⁶⁴³ 𐞗³⁰⁹² 𐞗⁵⁶⁴³ 𐞗⁰⁸³² 𐞗²⁰⁹⁸
thja¹ mji¹dji² mji¹ .wjo² ŋa²
 DEM except NEG able[B] 1SG
 ‘Apart from this, I can’t do anything.’
 (Jacques 2014: 242)

b. Siyuewu Khroskyabs

ldzè gəsâd t^hé ldzè rə-ród=næ fslá
 teach₁ certain any teach₁ NPST-need₁=but learn₁
t^hé fanjá=t^ha=məndi
 any holiday=LOC=except
t^hé rə-mə-gô
 any NPST-NEG-be.at.leisure₁

‘We should teach some (of our language), but except during the holidays, they don’t have any time off.’

c. Mazur Stau

t^hε mde mi-mjo-ã
 DEM except NEG-able-1SG

‘Apart from this, I can’t do anything.’ (Elicited)

3. 藏⁵⁴¹⁶ 能⁵⁰⁶⁴ *ɣwə²za²* ‘comb (noun)’ < Pre-Tangut **C-ku.zaC*, Mazur Stau *zɬəzɬə* ‘comb (noun)’, Wobzi Khroskyabs *ɬazəɣ* ‘comb (noun)’ and Siyuewu Khroskyabs *ɬozəɣ* ‘comb (noun)’. In East Gyalrongic, an unrelated etymon is used: Japhug *txɔt* ‘comb (noun)’, Zbu *tɕôṭ* ‘comb (noun)’, Bragbar Situ *ta-ɕêṭ* ‘comb (noun)’. The character 藏⁵⁴¹⁶ *ɣwə²* is glossed as ‘ahead’ in Li (1997: 980), and is certainly related to 纒²⁷⁵⁰ *yu¹* ‘head’ (Jacques 2014: 60). In the Mazur Stau and Khroskyabs cognates, the first elements *ɬə-* and *ɬa-* also mean ‘head’ (*ɬə* in Mazur Stau and *ɬu* in Khroskyabs). As for the second element, which is an innovation, the Mazur Stau initial *z-* is probably from **zj-*, as is justified by the Eastern Geshiza (a language closely related to Stau) cognate *ɣuə-zja* (Honkasalo 2019: 803). The Khroskyabs coda *-ɣ* would point to the precise reconstruction of **-k* for **-C* in Pre-Tangut. The coda is regularly lost in Tangut and Stau. The Khroskyabs rime *-əɣ* is problematic, as we do not find other instances where Khroskyabs *-əɣ* corresponds to Pre-Tangut **-aC*, however, the Tangut vowel *a* in other closed syllables can correspond to Khroskyabs *ə*, as in 𐰇⁰⁵²⁷ *.wja¹* ‘goiter’, cognate with Wobzi Khroskyabs *ɣbêv* ‘to swell’.

Innovated combination of common retentions

The only example found is 𐰇⁵⁵⁰⁰ 𐰇²⁷¹² *sjij¹.wji¹* ‘next year’ < Pre-Tangut **sjej.C-pja*, related to Wobzi and Siyuewu Khroskyabs *fsæpî* ‘next year’, Mazur Stau *sevi* ‘next year’, Khang.gsar Stau *sevə* ‘next year’. The first morpheme corresponds to the one in Japhug *fsɤq^he* ‘next year’, and cognates of the second morpheme is also widely found in East Gyalrongic languages, such as in Japhug *ɣuijpa* ‘this year’, Bragbar Situ *rə-piê* ‘a year’ and Zbu *kə-pjé* ‘a year’. However, the combination of these two morphemes are not attested in any of the documented East Gyalrongic languages.

5 Case marking

Tangut shares with modern West-Gyalrongic languages five case markers not attested in East Gyalrongic languages (see also Jacques et al. 2017) (Table 1).

We note that some of the case markers show a correspondence pattern not attested in cognates between Tangut and East Gyalrongic or other Sino-Tibetan languages. In general, the Pre-Tangut vowel *a*, which corresponds to *a* in most other languages, has different reflexes in Tangut depending on the syllable structure: Pre-Tangut open-syllable rime *-a* becomes Tangut *-i/-e*, while closed-syllable rime *-aC* becomes Tangut *-a*. For example, Pre-Tangut **r-ja* ‘hundred’, cf. Japhug *yurza*, Tibetan རྒྱལ་ *brgya*, gives 𐞗²⁷⁹⁸ *jir*² ‘hundred’; Pre-Tangut *r-jat* ‘eight’, cf. Japhug *kurcat*, Tibetan རྒྱལ་ *brgyad*, gives 𐞗⁴⁶⁰² *jar*¹ ‘eight’. This different reflex is paralleled in West Gyalrongic, cf. Siyuewu Khroskyabs *rji* ‘hundred’, Siyuewu Khroskyabs *vjad* ‘eight’, Mazur Stau *rja* ‘hundred’, Mazur Stau *rje* ‘eight’. The case markers 𐞗¹⁶⁸³ *ja*², 𐞗¹¹³⁹ *tchjaa*¹ and 𐞗¹¹³⁹ *k^ha*¹ have the rime *-a* in Tangut, but their cognates in Khroskyabs and Stau are open syllables with the vowels *a*, *æ*, *a*. This correspondence pattern suggests that these words are innovations later than the branching off of West Gyalrongic (including Tangut).

In Tangut, the marker 𐞗¹¹³⁹ *jij*¹ ‘POST’ has two main functions: allative as in (4a), and genitive as in (4b). Its Khroskyabs cognate =*ji* behaves exactly in the same way, as illustrated by the examples in (5).

- (4)

a.

𐞗¹⁶⁸³
*sjo*²
Xiang

𐞗¹¹³⁹
*jij*¹
POST

𐞗³⁷⁰⁰
*niq*²
poison

𐞗²⁵⁹⁰
*wji*²
DIR

𐞗¹³³²
*dej*¹
offer

𐞗¹⁶⁸⁶
*wji*²
?

‘She offered Xiang some poison.’

(Jacques 2007: 59)

b.

𐞗⁵⁹¹⁹
*tshe*²
humble

𐞗⁰²⁴³
*sji*²
wife

𐞗¹¹³⁹
*jij*¹
GEN

𐞗⁴⁹⁴⁵
*wə*¹
husband

‘The husband of the humble wife’

(Jacques 2007: 94)
- Table 1: Case markers in Stau, Khroskyabs and Tangut.
- | Khang-gsar/Mazur | | Wobzi/Siyuewu | | Tangut | |
|---------------------------------------|-------|---------------|-----|--|-------------------|
| -w/=ɣə | ERG | =ɣə | ERG | 𐞗 ⁵⁸⁸⁰ <i>ɣwu</i> ² | INSTR |
| -j/=ji | GEN | =ji/=jə | GEN | 𐞗 ¹¹³⁹ <i>.jij</i> ¹ | GEN, ANTIERGATIVE |
| -ɬa/=ɬa | ALL | =ɬa | LOC | 𐞗 ⁵⁸⁵⁶ <i>ya</i> ² | LOC |
| -tɕ ^h a/=tɕ ^h ə | LOC | | | 𐞗 ⁰⁰⁸⁹ <i>tchjaa</i> ¹ | LOC |
| -k ^h a/=k ^h ə | INSTR | | | 𐞗 ⁵⁹⁹³ <i>k^ha</i> ¹ | in the middle of |

- (5) a. Wobzi Khroskyabs
ŋæ̌=ji rŋamá p^hól=pa kə-væ̌-n=ni
 1SG=ALL face.cream offer=NMLZ IMP-bring₁₋₂=IMP
 ‘Bring me some face cream as a gift!’
 (Lai 2017: 186)
- b. *tʂač̌=ji pîr*
 Bkrashis=GEN pen
 ‘Bkrashis’ pen’
 (Lai 2017: 185)

The locative marker in Tangut, 𐞗⁵⁸⁵⁶ *ya*² ‘POST’, exemplified in (6), is comparable to Khroskyabs =*ba* and Stau =*ba*, as shown in (7).

- (6) 𐞗²⁵¹³ 𐞗²⁷⁰⁸ 𐞗³⁶⁰¹ 𐞗⁵⁸⁵⁶ 𐞗³⁰⁹⁹
*ju*² *rja*¹ *khew*² *ya*² *dzjiij*¹
 always office entrance POST say
 ‘He always stayed at the entrance of the office.’
 (Jacques 2007: 34)

- (7) a. Wobzi Khroskyabs
ænaeřesi <piānpian>=tə=ba juŋsâ lbé çsærpa=tə dzây ró
 but diaper=DEF=LOC again urine new=DEF exist must
 ‘But again, there will be fresh urine on the diaper.’
- b. Khang.gsar Stau
zdi-ba ya-st^hwej
 wall-LOC IMP-PUT.2SG
 ‘Put your hands on the wall!’

Tangut 𐞗⁰⁰⁸⁹ *təhja*¹ ‘LOC’ is comparable to Khang.gsar Stau *tə^ha* ‘LOC’/Mazur Stau =*tə^hæ* ‘LOC’, examples are shown in (8a) for Tangut, and (8b) for Khang.gsar Stau.

- (8) a. 𐞗²⁷⁸⁷ 𐞗¹⁸⁹³ 𐞗⁵⁸⁸⁰ 𐞗³⁴⁰⁶ 𐞗⁰⁰⁸⁹ 𐞗³⁶⁷⁸ 𐞗¹¹⁰¹
*tü*¹ *lhji*¹ *ŋwu*² *pju*² *təhja*¹ *to*² *jij*¹
 net pull LINK hall LOC go.up FUT
 ‘He (the bear) pulled the net and was about to go up to the hall.’
 (Jacques 2007: 81)
- b. Khang.gsar Stau
təu-tə^ha <pidai> kə-ŋc^hu
 waist-LOC belt PST-PUT.1SG
 ‘I put a belt on my waist.’

Tangut 𐞤⁵⁹⁹³ *kha*¹ ‘LOC: in the middle of’ is cognate to Khang.gsar Stau *-k^ha* ‘LOC, INSTR’ and Mazur Stau *=k^hæ* ‘LOC, INSTR’. A Tangut example is illustrated in (9a). The locative use in Stau is shown in (9b), and the extended instrumental use is illustrated in (9c).

- (9) a. 𐞤³⁵⁶⁶ 𐞤⁰⁴⁷⁷ 𐞤⁵⁰³⁷ 𐞤⁵⁷¹⁰ 𐞤⁵⁹⁹³ 𐞤¹²⁴⁵ 𐞤⁰⁷³¹
*lu*² *zji*¹ *bji*¹ *lji*¹ *kha*¹ *jij*¹ *lju*²
 wife.Lu sword arrow LOC self throw
 ‘The wife from the Lu family threw herself in the middle of swords and arrows.’
 (Jacques 2007: 20)
- b. Khang.gsar Stau
zgøder-de *χpərju-k^ha* *zgezgə-sə* *ci-rə*
 flag-DEF wind-LOC wave-PROG exist-SEN
 ‘The flag is waving in the wind.’
- c. Khang.gsar Stau
ɲa *sco^v-k^ha* *ɣremə-nə* *ɣrə-de* *rə-k^hev*
 1SG spoon-INSTR well-LOC water-DEF PST-spoon.up
 ‘I spooned water from the well with a spoon.’

6 Partial reduplication of the West Gyalrongic type

There are two types of partial reduplication attested in Tangut (Gong 1997, 2003; Jacques 2014). In the first type of reduplication, the reduplicant is prefixed to the stem and has a weakened rime (usually a schwa). In the second type of reduplication, the reduplicant is suffixed to the stem and undergoes rime replacement.

The first type is found in many Gyalrongic languages (see Jacques and Chen 2004; Gates 2017). In Tangut, the rime of the prefixed reduplicant is weakened into *-ə* or *-i*, as shown below.

1. 𐞤³³⁵⁵ 𐞤⁰³⁸³ *lə*².*le*² ‘fight each other’
2. 𐞤⁰²²³ 𐞤⁴⁵³³ *kiə*¹.*kie*¹ ‘call’
3. 𐞤²²³⁶ 𐞤⁰⁵⁹⁴ *lhji*¹.*lhji*² ‘slowly’
4. 𐞤¹⁸³³ 𐞤³⁴⁰⁴ *ɕji*¹.*ɕji*¹ ‘hesitate, doubt’

The second type is less clearly described and incompletely documented. Gong (1997) and Jacques (2014) discovered two patterns: first, the vowel of the reduplicant is replaced by a rounded vowel (as in the examples ‘exchange, trade’ and

Table 2: Partial reduplication in Wobzi Khroskyabs.

Meaning	Stem	Reduplicated	Comment
Crow	<i>qadí</i>	<i>qadí~da</i>	Native word
Field	<i>mpələ́</i>	<i>mpələ́~lu</i>	Native word
Village	<i>jâl</i>	<i>jâl~ja</i>	Tibetan ཡུལ་ <i>yul</i> ‘village’
Beverage	<i>jĩnljo</i>	<i>jĩnljo~lja</i>	Chinese 飲料 <i>yǐnliào</i> ‘beverage’

‘stupid’ below); and second, the vowel of the reduplicant is replaced by an unrounded one (as in the example ‘surround’ below).

- 6. 𐄌⁰¹⁴¹𐄌⁰¹²² *wjiɪ¹.wjoo¹* ‘exchange, trade’
- 7. 𐄌⁵²⁰⁸𐄌⁵⁰⁸⁶ *giee¹.gioo¹* ‘stupid’
- 8. 𐄌²⁷⁵⁷𐄌²⁷⁴³ *ror².rejr²* ‘surround’

However, there is a third pattern that has never been described previously. In this pattern, the vowel of the reduplicant is replaced by *-a*, as exemplified below.

- 9. 𐄌⁰⁷⁶⁴𐄌⁰⁸⁰³ *rjiɪr¹.rjar²* ‘horse’
- 10. 𐄌¹⁰³⁵𐄌¹²²⁹ *lhjiɪ².lhaa²* ‘soft’
- 11. 𐄌²¹⁷³𐄌⁰³⁹⁶ *thju².thja¹* ‘this and that’
- 12. 𐄌³⁴³⁹𐄌³²⁸⁰ *pjiɪ¹.pja²* ‘wizard’
- 13. 𐄌⁴⁴⁷¹𐄌⁴⁴⁷⁵ *xiu¹.xa** ‘blow, puff’
- 14. 𐄌⁰⁴³⁹𐄌¹³³³ *ɣie¹.ɣja¹* ‘cook’
- 15. 𐄌³²⁸⁷𐄌⁵⁵⁶¹ *rur¹.rar²* ‘narrow’

This new pattern is rarely found in East Gyalrongic languages, but this reduplication pattern is pervasive in West Gyalrongic languages, and is often productive. Lai (2013) finds this productive reduplicating process in Wobzi Khroskyabs, in which the reduplicant has the vowel *-a* or in a few cases *-u*. The process is fully productive even with recent Chinese borrowings, as shown in Table 2.

In Stau, this process is not productive, but lexicalized forms are attested in many examples, such as Khang.gsar Stau *tɛədza* ‘skin’ (cf. Wobzi Khroskyabs *dzêdza* ‘skin’), *ɣziɣza* ‘learn’ (cf. Wobzi Khroskyabs *ldzê* ‘teach’), *gəga* ‘wear’ (cf. Wobzi Khroskyabs *gí* ‘wear’), *məŋa* ‘fry’ (cf. Wobzi Khroskyabs *mû* ‘fry’).

The pattern with *-a* in Tangut clearly corresponds to that with *-a* in Wobzi Khroskyabs and *-a* in Khang.gsar Stau, and the pattern *-o* in Tangut is probably related to the pattern with *-u* in Wobzi Khroskyabs.

Table 3: The ordering of inflectional prefixes in Tangut.

1	2	3	
Orientation/interrogative	Negative	Modal	Verb

Table 4: Generic ordering of inflectional prefixes in West Gyalrongic.

1	2	3	4	
Orientation	Interrogative/modal	Negative	(Modal)	Verb

7 Verbal morphology

This section focuses on commonalities shared by Tangut and West Gyalrongic within the domain of verbal morphology. We focus on the similar ordering of inflectional prefixes between Tangut and West Gyalrongic in Section 7.1. In Section 7.2, we discuss argument indexation in these languages. In Section 7.3, we compare orientation prefixes between Tangut and West Gyalrongic.

7.1 The ordering of inflectional prefixes

Tangut shares a similar structure of inflectional prefixes with West Gyalrongic languages, which is summarized from Jacques (2011: 437) in Table 3.

West Gyalrongic languages, like Tangut, generally place interrogative and negative prefixes after orientation prefixes, while East Gyalrongic languages have the reverse order, with interrogative and negative prefixes placed before orientation prefixes.⁴ Simplified generic inflectional templates of the two branches are presented in Tables 4 and 5 (for verbal templates in East Gyalrongic languages, see Gong 2018; Jacques 2008; Zhang 2016, for verbal templates in West Gyalrongic languages, see Gates and Kim 2018; Jacques et al. 2017; Lai 2017).

Examples showing the prefix order of Tangut, West Gyalrongic and East Gyalrongic are illustrated in Examples (10), (11) and (12), respectively. The inflectional prefixes are between square brackets.

⁴ Orientation prefixes have two functions. First, they denote the direction of the verb; second, they bear the tense-aspect-modality-evidentiality information of the verb.

c. Cogtse Situ (East Gyalrongic)

wətə=nè [mə-ko-tə]-wasu-n zə [mə-ko]-ŋôš zə

DEM=PL NEG-PFV-2-correct.II-2SG PART Q-IMPV-*be*.II PART

‘Haven’t you got wrong?’

(Lin 2016: 280)

A curious parallelism between Tangut and Khroskyabs is found in the modal slot (Slot 3 in Table 3 and Slot 4 in Table 4). A modal adverb is inserted between inflectional prefixes and the verb stem. In Tangut, there are four such adverbs, 𐰚¹³⁷⁴ *tšhjì*¹ ‘?’, 𐰚⁵⁸¹⁵ *tsjì*¹ ‘?’, 𐰚⁴⁴⁴⁴ *ljì*¹ ‘CONCSV’ and 𐰚³⁵²⁷ *mja*¹ ‘IRR’ (Jacques 2011: 427–436). Although the meanings are not entirely understood, their position in the template can be inferred through textual examples. See Example (13) illustrated with 𐰚⁴⁴⁴⁴ *ljì*¹ ‘CONCSV’, showing that the adverb is placed after the orientation prefix and the negative prefix.

- (13) 𐰚⁰⁷⁹⁵-𐰚⁴⁴⁴⁴-𐰚⁵¹¹³ 𐰚⁵⁶⁴³-𐰚⁴⁴⁴⁴-𐰚⁵¹⁷⁶ 𐰚⁵⁹⁸¹-𐰚⁴⁴⁴⁴-𐰚⁴⁸⁵⁹

*rjir*²-*ljì*¹-*wji*¹ *mji*¹-*ljì*²-*dzioow*² *a*⁰-*ljì*¹-*to*²

PFV-CONCSV-do NEG-CONCSV-may PFV-CONCSV-end

𐰚⁵⁶⁴³-𐰚⁴⁴⁴⁴-𐰚⁰¹³⁹

*mji*¹-*ljì*¹-*nej*²

NEG-CONCSV-profitable

‘Although it is done, it is not satisfying; although it is completed, it is not profitable.’

(Kychanov 1974: 122, 209)

In Khroskyabs, a similar phenomenon is observed. The adverb *də-* ‘even’ can be placed at the end of the inflectional chain, see the example in (14).

(14) Siyuewu Khroskyabs

kə-mə-də-sŋó-ŋ æ-mŋədmŋənd nə-rjé-æŋ

PST-NEG-even-sleep-1SG DEM-appearance PST-sit₂-1SG

‘I did not even sleep and was sitting like that.’

This particular positioning of adverbs is not attested in East Gyalrongic languages, and can be considered a feature of West Gyalrongic.

7.2 Inverse marking and stem alternations

Argument indexation and stem alternation show both Tangut’s Gyalrongic affiliation and its close relationship with West Gyalrongic. As shown in Gong (2016), Tangut shares with West Gyalrongic languages the loss of direct 3 → 3’ forms in transitive verbs, by generalizing the inverse configuration.

Transitive verbs have two arguments: A (a prototypical semantic agent and syntactic subject) and P (a prototypical semantic patient and syntactic object). In Gyalrongic languages, the direct configuration indicates cases where the A ranks higher than the P in the empathy hierarchy (1, 2 $\rightarrow$ 3); the inverse configuration indicates cases where A ranks lower than P in the empathy hierarchy (3, 2 $\rightarrow$ 1) (Gong 2014; Jacques 2010a; Lai 2015; Sun and Shidanluo 2002; Zhang 2019).

In a typical East Gyalrongic language (Gong 2014; Jacques 2010a), cases when both A and P are third person have a distinction between proximate (3) and obviate (3'). The distinction is semantically meaningful, as it shows the topicality and (as shown in 15) animacy of the arguments.

- (15) Zbu
- a. Direct (3 $\rightarrow$ 3')
- | | | | | |
|--------------|-----------|----------------|----------------------------|-----------|
| <i>tʂɛɕi</i> | <i>kə</i> | <i>skutséʔ</i> | <i>nɛ-v-tɕʰóv</i> | <i>ki</i> |
| Bkrashis | ERG | stone | AOR-DIR-smash ₂ | NVIS |
- ‘Bkrashis smashed a stone.’
- b. Inverse (3' $\rightarrow$ 3)
- | | | | | |
|--------------|----------------|-----------|--------------------------|-----------|
| <i>tʂɛɕi</i> | <i>skutséʔ</i> | <i>kə</i> | <i>tə-wə-xśâv</i> | <i>ki</i> |
| Bkrashis | stone | ERG | AOR-INV-hit ₂ | NVIS |
- ‘A stone hit Bkrashis.’
- (Gong 2014: 50)

However, in West Gyalrongic languages, the distinction between 3 and 3' is neutralized, showing marks of the inverse configuration, as exemplified in (16) by Siyuewu Khroskyabs.

- (16) Siyuewu Khroskyabs
- a. *stâɕ=tə=ɣə* *χɕi* *a-dzíd=tsʰi*
- | | | |
|---------------|-------|-------------------------------|
| tiger=DEF=ERG | grass | PST.INV-eat ₂ =IFR |
|---------------|-------|-------------------------------|
- ‘Oops, the tiger ate the grass.’
- b. *sæqé=ɣə* *ɣlé* *nu-scâr*
- | | | |
|-----------|--------|----------------------------|
| stump=ERG | rabbit | PFV.INV-scare ₂ |
|-----------|--------|----------------------------|
- ‘The stump scared the rabbit.’

Although Tangut does not directly preserve an overt inverse prefix, the neutralization of the distinction between 3 and 3' can be observed in its inflectional paradigms. Certain transitive verbs in Tangut exhibit stem alternation (Gong 2001). A stem-alternating verb can have two stems, labeled stem A and stem B. For example, the verb ‘to eat’ has stem A 𐰇𐰺⁴⁵¹⁷ *dzji*¹ and stem B 𐰇𐰺⁴⁵⁴⁷ *dzjo*¹. This distinction, shown by Jacques (2009) as originating in the suffix -w, is cognate with the Stem 1/3 distinction in Zbu Gyalrong. In Zbu Gyalrong, ‘to eat’ has stem 1 ⁿ*dzé* and stem 3 ⁿ*dzó*. As is shown in Table 6, the forms with 3rd person

Table 6: Person marking in stem-alternating verbs in Zbu Gyalrong and Tangut.

	Zbu	Tangut
1SG → 3SG	ⁿ dzo-ŋʔ	dzjo ¹ ŋa ²
3 → 1SG	və- ⁿ dzé-ŋʔ	dzji ¹ ŋa ²
1PL → 3	ⁿ dzé-jə	dzji ¹ nji ²
3 → 1PL	və- ⁿ dzé-jə	dzji ¹ nji ²
2SG → 3	tə- ⁿ dzoʔ	dzjo ¹ na ²
3 → 2SG	tə-və- ⁿ dzéʔ	dzji ¹ na ²
2PL → 3	tə- ⁿ dzé-ŋə	dzji ¹ nji ²
3 → 2PL	tə-və- ⁿ dzé-ŋə	dzji ¹ nji ²
3sg → 3'SG	ⁿ dzoʔ	dzji ¹
3'SG → 3SG	və- ⁿ dzéʔ	dzji ¹

singular have cognate person-marking patterns in Zbu Gyalrong and Tangut (Gong 2016).

The only thing that is different between the languages in Table 6 concerns the forms for 3SG → 3SG. Unlike Zbu Gyalrong, Tangut has only one 3sg form (dzji¹), which formally corresponds to the inverse form of Zbu Gyalrong (və-ⁿdzéʔ). Tangut shares with Khroskyabs and Stau the collapse of the proximate/obviate distinction which results in a form that is formally equivalent to inverse forms in Zbu Gyalrong.

The 3 → 3' neutralization is considered an important and unusual shared innovation of West Gyalrongic (Jacques et al. 2017: 608), the case in Tangut is highly unlikely to be an independent development.

7.3 Orientation prefixes

Orientation prefixes denote primarily the direction of the action described by its verbal host. It also contains inflectional information such as tense, aspect, modality and evidentiality. In this section, we compare the system and the morphophonology of orientation prefixes between Tangut and other Gyalrongic languages.

It is not straightforward to point out a specific shared innovation in orientation prefixes between Tangut and modern West Gyalrongic, as all the prefixes could be retentions.⁵ However, it is worthwhile to emphasize the striking similarity between Tangut and modern West Gyalrongic in this perspective.

5 We acknowledge that other languages such as nDrapa and Choyu may also have similar systems, however the affiliations of these languages are still unclear and seem very closely related to Gyalrongic.

Khroskyabs varieties have almost the same set of orientation prefixes as Tangut, shown in the second column of Table 7. The prefixes for ‘up’, ‘down’, ‘cis’, ‘TRANS’ and ‘neutral direction’ in Tangut are clearly related to those in Khroskyabs. The prefix 𐰇⁰⁸⁰⁴ *dji*² ‘gain’ corresponds to *nə*- ‘downstream’ in Khroskyabs and *nu*- ‘west’ in Japhug. According to Jacques (2008: 245), the nasal consonant of this prefix comes from an irregular simplification of a proto **nd-*, from which Tangut inherited the occlusive part. The only orientation prefix that has no confirmed West Gyalrongic cognate is 𐰇⁴³⁴² *dja*² ‘loss’. It could however be comparable to Khroskyabs *læ-*, meaning ‘left bank’ in Wobzi and ‘upstream’ in Siyuewu, but the sound correspondence needs to be explained.

As orientation prefixes are prone to rapid semantic change even among Khroskyabs dialects (Lai forthcoming; Lai et al. 2018), the semantic mismatches between Tangut and other West Gyalrongic languages are not obstacles for these etymologies.

Tangut orientational prefixes present two series: Series A, the default series, for realis situations; and Series B for optative and irrealis ones (Gong 2003; Jacques 2014). A vowel change into *-i* characterizes Series B, as shown in Table 8. In Khang.gsar Stau, the two series of orientational prefixes (default and irrealis/interrogative) exhibit exactly the same vowel change, as shown in Table 9 (Jacques et al. 2017: 601).

Exhibiting common innovations or not, the formal and behavioral parallelisms of orientation prefixes between Tangut and modern West Gyalrongic are not negligible. Orientation prefixes can change and be replaced drastically (as is the case of East Gyalrongic, see Gong 2018: 286–290 for the case of Zbu and Zhang forthcoming for a comparison within East Gyalrongic). The fact that the Tangut orientation prefixes are almost the same with the Khroskyabs and Stau ones would

Table 7: Orientational prefixes in Tangut and in other Gyalrongic languages.

Tangut	Wobzi Khroskyabs	Siyuewu Khroskyabs	G.yu.rong Stau	East Gyalrongic
𐰇 ⁵⁹⁸¹ <i>a</i> ⁰ ‘up’	<i>æ</i> - ‘up’	<i>o</i> - ‘up’	<i>a</i> - ‘up’	
𐰇 ¹⁴⁵² <i>nja</i> ¹ ‘down’	<i>næ</i> - ‘down’	<i>næ</i> - ‘down’	<i>nə</i> - ‘down’	<i>nə</i> - (Cogtse)
𐰇 ¹³²⁶ <i>kji</i> ¹ ‘cis’	<i>kə</i> - ‘upstream’	<i>kə</i> - ‘dark side (mountain)’	<i>gə</i> - ‘higher altitude’	<i>kɣ</i> - ‘eastward’ (Japhug)
𐰇 ²⁵⁹⁰ <i>wji</i> ² ‘TRANS’	<i>və</i> - ‘right bank’	<i>və</i> - ‘downstream’	<i>yə</i> - ‘lower altitude’	
𐰇 ⁴³⁴² <i>dja</i> ² ‘loss’				
𐰇 ⁰⁸⁰⁴ <i>dji</i> ² ‘gain’	<i>nə</i> - ‘downstream’	<i>nə</i> - ‘sunny side (mountain)’		<i>nu</i> - ‘west’ (Japhug)
𐰇 ⁰⁷⁹⁵ <i>rji</i> ² ‘neutral’	<i>rə</i> - ‘neutral’	<i>rə</i> - ‘neutral’	<i>də</i> - ‘neutral’	

Table 8: Tangut orientational prefixes.

Series A	Series B	
𐞰 ⁵⁹⁸¹ <i>a</i> ⁰	𐞰 ³⁹⁸⁹ <i>.jij</i> ¹	up
𐞰 ¹⁴⁵² <i>nja</i> ¹	𐞰 ³⁸⁴⁶ <i>njjj</i> ²	down
𐞰 ¹³²⁶ <i>kji</i> ¹	𐞰 ²²¹⁹ <i>kjjj</i> ¹	CIS
𐞰 ²⁵⁹⁰ <i>.wjji</i> ²	𐞰 ²⁵³⁶ <i>.wjji</i> ²	TRANS
𐞰 ⁴³⁴² <i>dja</i> ²	𐞰 ⁴⁸⁴¹ <i>djjj</i> ²	loss
𐞰 ⁰⁸⁰⁴ <i>dji</i> ²	𐞰 ⁴⁸⁴¹ <i>djjj</i> ²	gain
𐞰 ⁰⁷⁹⁵ <i>rjir</i> ²	𐞰 ³⁷⁰⁶ <i>rjjir</i> ²	neutral

Table 9: Khang.gsar Stau orientational prefixes.

Series A	Series B	
<i>rə-</i>	<i>rí-</i>	up
<i>nə-</i>	<i>ní-</i>	down
<i>kə-</i>	<i>kí-</i>	north
<i>yə-</i>	<i>yí-</i>	south
<i>tə-</i>	<i>tí</i>	neutral

be an unusual coincidence if they were not in the same subgroup; the odds of two distantly related languages having the same combination of orientation prefixes is rather low. Furthermore, given the geographic and temporal distance between Tangut and Khroskyabs or Stau and the number of prefixes, it is improbable that the similarity is due to recent borrowing.

7.4 Modal markers

In this section, we focus on two markers, 𐞰³⁹¹⁶ *sjj*² ‘IFR’ and 𐞰⁰⁷³⁴ *mo*² ‘irrealis particle’, of which parallels can be found in West Gyalrongic languages, but not in East Gyalrongic.

7.4.1 𐞰³⁹¹⁶ *sjj*² ‘inferential marker’

The Tangut particle 𐞰³⁹¹⁶ *sjj*² ‘IFR’ is used with perfective verbs with an inferential implication (second-hand or inferred information), as shown in (17). This marker is comparable to Khroskyabs =*si* (Lai 2017: 495–497) and Stau -*sə*, which are

inferential markers generally only compatible with the past or perfective forms of a verb. Examples are shown in (18).

In (17a), the character heard of the affair of the righteous woman's husband before asking the question. The clause with 就³⁹¹⁶ *sjɿ²* 'IFR' is therefore a piece of second hand information. Likewise, in (17b), the newly married character talked to her sister-in-law describing an event before her marriage; a fact that could not be first-hand. As a result, the inferential marker is used here.

- (17) a. 做²⁵⁴¹ 是⁰²⁴³ 像¹⁹⁶² 有¹¹³⁹ 说¹²⁷⁸ 有¹⁰⁴⁵ 有²⁰³⁸
*dzɰw*² *sjɿ*² *tjɿ*² *ji*¹ *jɿ*² *dɔ*¹ *sew*¹
 man woman model ALL say say concubine
 听⁰⁸⁰⁴ 看⁰¹⁴⁷
*dji*² *ɕji*²
 DIR invite
 就¹⁶⁸⁶ 是⁴⁶⁰¹ 就³⁹¹⁶ 有⁵⁹⁹³ 有¹⁹¹⁸ 有⁴⁴⁸¹ 有⁴⁶⁰¹
*wji*² *nja*² *sjɿ*² *kha*¹ *mji*¹ *ɕji*¹ *nja*²
 ? 2SG IFR INSTR NEG go.II 2SG
 就⁰⁵⁸² 就⁵⁰⁵⁷
*thji*² *yie*¹
 why
 说¹²⁷⁸
*jɿ*²
 say
 'Someone asked the righteous woman, "Since your husband married a concubine, why don't you leave?"'
 (Jacques 2007: 99)
- b. 就¹⁸²² 有⁴⁰²⁸ 有³⁹⁸⁶ 有⁴⁸⁹³ 有¹¹³⁹ 有¹⁵²⁶
*ɲwu*¹ *nji*¹ *nji*¹ *wji*¹ *ji*¹ *tshji*²
 say 2SG mother.in. law ALL take.care
 就⁵⁸⁸⁰ 就⁰⁵²⁴
*ɲwu*² *dzju*¹
 CONJ instruction
 就²⁵⁹⁰ 就⁵⁵⁹¹ 有⁴⁶⁰¹ 就³⁹¹⁶ 有⁵⁰²⁴ 有¹⁹¹⁸
*wji*² *lhji*² *nja*² *sjɿ*² *li*¹ *mji*¹
 DIR receive.II 2SG IFR change NEG
 有¹²⁷⁴ 有¹²⁷⁸
*wo*² *jɿ*²
 able say
 'She said, "You took care of your mother-in-law and received her instructions, this should not change."
 (Jacques 2007: 101)

Example (18a) is the most common way to begin a story in Wobzi Khroskyabs. The inferential particle =*si* must be present because it is a fictitious story. Example (18b) is an elicited sentence in which the speaker has no choice but to add the inferential marker at the end because he does not talk about a fact. In (18c), the verb form *jə-rə* (say-SEN) clearly marks a hear-say situation, and as a consequence the inferential marker must be applied.

(18) a. Wobzi Khroskyabs

ɲnâɲna=tə=gə *â-djəm* *næ-dâ=si*
 long.ago=DEF=loc one-family IPFV.PST-exist₂=IFR
 ‘Long long ago, there was a family.’
 (Lai 2017: 496)

b. Khang.gsar Stau

tšaci-w *dzəma* *de* *nə-fse-sə*
 Bkrashis-ERG Sgrolma DEM PFV-INV-kill-IFR
 ‘Bkrashis killed Sgrolma.’
 (Jacques et al. 2017: 608)

c. Mazur Stau

wo *p^hu* *æ-sɛ* *nə-rk^hə-sə* *ŋə-rə* *jə-rə*
 again pocket one-CLF PFV-put.in-IFR COP-SEN say-SEN
 ‘Again she filled up her pocket, it is said.’

7.4.2 𐰇⁰⁷³⁴ *mo*² ‘irrealis particle’

𐰇⁰⁷³⁴ *mo*² is termed ‘irrealis particle’ after Jacques (2007, 2014). However, it has three different uses in Tangut: as the marker of the apodosis of a conditional sentence (19a), as a question particle (19b), and as a conjunction meaning ‘or, otherwise’ (19c); see Duan (2015) and Gates (2021). These three uses of 𐰇⁰⁷³⁴ *mo*² in Tangut correspond to *mo* in Mazur Stau, as seen in (20a), (20b), and (20c), respectively.

- (19) a. 𐰇³⁹²⁶ 𐰇¹⁵⁶¹ 𐰇⁵⁰⁴⁹ 𐰇²⁵¹⁸ 𐰇⁵⁸⁵⁶ 𐰇³⁵²⁷ 𐰇²⁵¹⁶ 𐰇¹⁶⁸⁶ 𐰇⁰⁷³⁴
*nja*² *nji*².*wja*¹ *njiij*¹ *ya*² *mja*¹ *ye*² *wji*² *mo*²
 2SG brother.in.law heart POST HYPOT harm ? IRR
 ‘(If you cry,) you would break your brother-in-law’s heart.’
 (Jacques 2007: 54)

- b. 𐰇⁴⁰²⁸ 𐰇⁰⁰⁹² 𐰇⁰⁵⁰⁸ 𐰇⁰⁷³⁴
*nji*² *mja*¹ *ŋwu*² *mo*²
 2SG mother be Q
 ‘Is this your mother?’
 (Jacques 2007: 39)

- c. 難⁰²⁶¹ 疑⁰³⁹⁶ 死⁵¹⁶⁵ 批⁵⁹¹⁸ 漱⁰⁷³⁴ 龜⁰⁴⁸⁷ 數⁵⁶⁴³ 統³⁵⁷⁶
*mjo*² *thja*¹ *twu*¹ *sji* *mo*² *sjwu* *mji*¹ *sjwij*
 1SG DEM place die.II or alive NEG clear
 ‘It is not clear whether I will die or stay alive there.’
 (Jacques 2007: 14)

- (20) a. *tɕʰəɣɛ* *ŋæ=tɕʰæ* *mtsʰere* *və-rə* *ni* *tʰi*
 then 1SG=on scold do-SEN 2SG DEM
pi *kɛ-qʰəmæ* *vi* *reɣɛ*
 like INTENS-bad do.2 and
smeze=ji *mo=ni qə-re* *də-rə* *mo jə-rə*
 girl=GEN eye=PL go.blind-NMLZ have-SEN IRR say-SEN
 ‘Then she scolded me, “To do like this is bad and the girl’s eyes will be blinded.”’
- b. *tɕʰəɣɛ* *æti* *tʰi* *æ-ɬe=ɣə* *tə-jə-də*
 then elder.brother DEM ONE-CLF=ERG PFV-say-NMLZ
ɣəzə *kɛ-mtsʰer-gə*
 sparrow INTENS-strange-NMLZ
ɣzæ-rə *æ-tɕʰə* *jə-rə* *mo tə-jə*
 sound-SEN Q-DEM say-SEN Q PFV-say
 ‘Then the elder brother, that one, said, “The sparrow is making a strange sound, what is he saying?”’
- c. *wo* *wo* *tɕʰə=ve* *ni* *ngələ=nə* *qe=tɕʰæ*
 yes yes DEM=SBJV 2SG wok=in throw=time
ŋæ *dɛ* *ɕæræræ* *kə-qe*
 1SG DEM quickly PFV-throw
mde *ŋæ* *pʰjã* *mo* *tə-jə-sə* *ŋə-rə*
 otherwise 1SG flee.1 or PFV-say-IFR COP-SEN
 ‘“Oh, okay. If so, when you throw me in the pot, you’d better throw me in fast or else I’ll escape,” the rabbit said.’

8 Note on the migration of the Tanguts

The homeland of the Tanguts is generally thought to be in north-eastern Tibet (Ruth 1994: 155–156, Beckwith 2009: 172). Under the name of 黨項 *Dǎngxiàng*, they first appeared in the seventh century in Chinese historical sources. According to the Book of Sui (隋書 *Suí Shū*, 636 AD), the territory of the Tanguts extended eastwards to today’s Gansu Province, and westwards to today’s Ruoqiang County

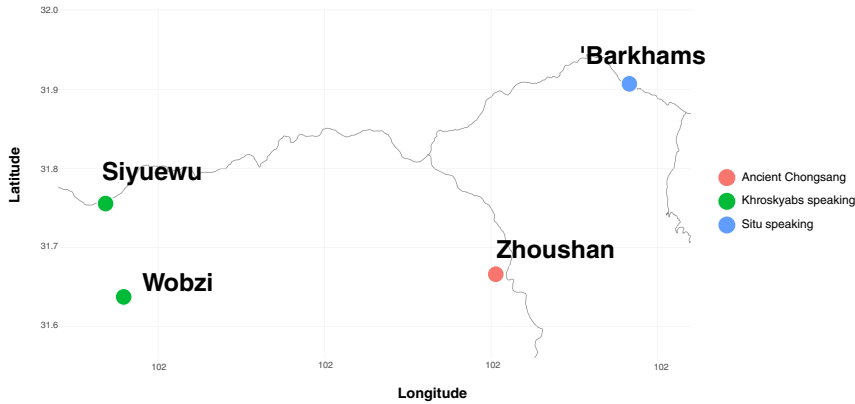


Figure 2: The location of Zhoushan (Ancient Chongsang, the southern border of the Tangut population).

in Xinjiang.⁶ The Old Book of Tang (舊唐書 *Jiù Táng Shū*, 945 AD) states that the east border was located at today's Songpan County, in Rngaba Prefecture, Sichuan, and more importantly, that the south border connected the lands of Chongsang (春桑 *Chōngsāng*) and Misang (迷桑 *Mísāng*).⁷ Chongsang, according to Li (1993: 119, 128), corresponds to today's Zhoushan (周山, *Zhōushān*) in Jinchuan County (金川縣 *Jīnchuān Xiàn*, Written Tibetan: ལྷ་ཆེན་ *Chuchen*), which is just at the frontier that separates Khroskyabs and Situ today, see Figure 2. If Li (1993) is correct, there was probably still a Tangut population residing in today's Gyalrongic speaking region shortly before they moved northeastwards to found the Tangut Empire (1038 AD).

In general, the Tangut population expanded from the southwest to the northeast. The time of separation of East and West Gyalrongic is estimated at about 3000 years before present (Sagart et al. 2019: 10320), which is more than 1500 years before the first documentation of the Tanguts in Chinese history. This time span is sufficient for them to become completely separate from their homeland neighbours. It is conceivable that the homeland of the Tanguts was actually further

⁶ “東接臨洮、西平，西拒葉護，南北數千里，處山谷間” (trans: “To the east, it connects Lintao and Xiping [today's Gansu], to the west, it occupies Yabghu [today's Ruqiang, Xinjiang]. From the south to the north, it situates among several thousand *li* of mountains and valleys.”) (隋書·黨項傳, *The Book of Sui: The Tanguts*)

⁷ “其界東至松州，西接葉護，南雜春桑、迷桑等羌，北連吐谷渾，處山谷間，互三千里” (trans: “Its border reaches Songzhou to the east, Yabghu to the west, the Qiang territories of Chongsang and Misang to the south, and Tuyuhun to the north. Located in the mountains and valleys, it covers over three thousand *li* of land.”) (舊唐書·黨項傳, *The Old Book of Tang: The Tanguts*)

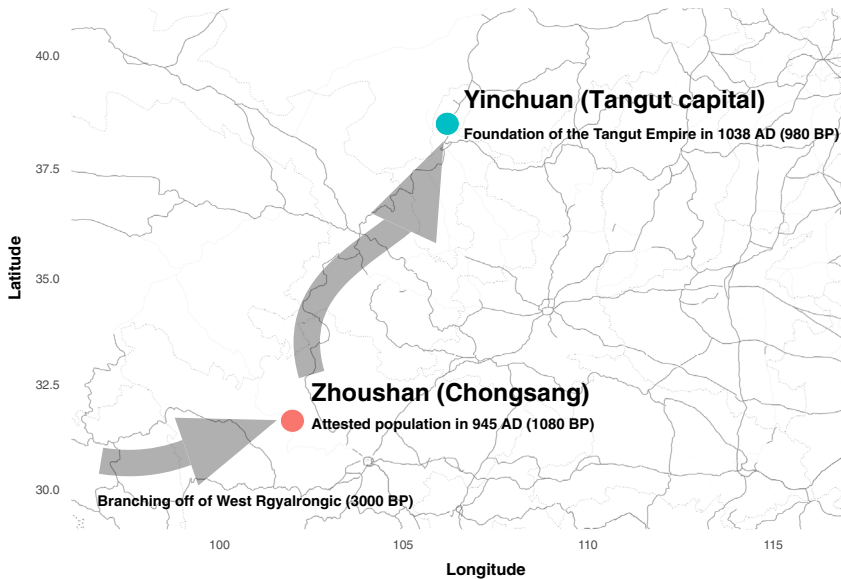


Figure 3: Potential migration route of the Tanguts.

southwest than we previously thought it to be, deep inside the Gyalrongic speaking region, and that they gradually migrated to the northeast and settled down in the Ordos region. Figure 3 illustrates the potential migration route of the Tanguts.

9 Conclusion

The present paper, using traditional methodology for subgrouping, supports the results obtained in Sagart et al. (2019). With phonological, lexical and morphological comparisons, we have shown that Tangut should be classified as a West Gyalrongic language, which means that Tangut is the only documented medieval Gyalrongic language.

In terms of phonology, Tangut shares with modern West Gyalrongic languages the bilabial reflex, *bj-* (*bl-* in Khroskyabs and Stau) of the cluster *mgl-* in East Gyalrongic. It also exhibits innovative additions and losses of preinitials, such as *s-*, *r-* and *m-*, which are absent in East Gyalrongic languages.

Amongst the lexical comparisons, common bisyllabic words unattested in East Gyalrongic languages are strong evidence for subgrouping Tangut in West Gyalrongic: a total of eight examples are found. By contrast, there are no known common bisyllabic words exclusively shared by East Gyalrongic and Tangut.

From a morphological perspective there are three pieces of evidence in favour of placing Tangut within West Gyalrongic. First, Tangut exhibits the partial reduplication of the West Gyalrongic type, with the reduplicant suffixed to the stem. Second, Tangut neutralized the direct-inverse distinction in $3 \leftrightarrow 3'$ scenarios, which is an important innovative feature of West Gyalrongic as well. Third, Tangut exhibits a system of orientation prefixes close to West Gyalrongic languages, which is even almost identical to the system in Khroskyabs.

Furthermore, at the syntactic level, Tangut shares with West Gyalrongic innovative case markers, as well as modal markers.

In-depth studies on Tangut phonology, morphology and syntax will help us better understand the evolution of other Gyalrongic languages. Future studies should aim at establishing detailed sound correspondences and laws among Gyalrongic languages, as well as dating the split of each subgroup.

Abbreviations

ALL	allative
AOR	aorist
CLF	classifier
CONCSV	concessive
CONJ	conjunction
COP	copula
DAT	dative
DEF	definite
DENOM	denominal
DIR	directional/orientation
DU	dual
ERG	ergative
FACT	factual
FUT	future
GEN	genitive
HYPOT	hypothetical
IFR	inferential
IMP	imperative
INSTR	instrumental
INTENS	intensifier
INTERJ	interjection
INV	inverse
IPFV	imperfective
IRR	irrealis
LINK	linking word
LOC	locative
NEG	negative

NMLZ	nominaliser
NPST	non-past
NVIS	non-visual
PART	particle
PFV	perfective
PL	plural
POST	postposition
PROG	progressive
PROH	prohibitive
PST	past
Q	interrogative
SBJV	subjective
SEN	sensory
SG	singular
TRANSLOC	translocative
?	unknown grammatical function.

Acknowledgements: This research is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 715618, Computer-Assisted Language Comparison, <https://digling.org/calc/>). We would like to thank the two anonymous reviewers for their very constructive comments.

References

- Beckwith, Christopher I. 2009. *Empires of the silk road: A history of Central Eurasia from the bronze age to the present*. Princeton and Oxford: Princeton University Press.
- Duan, Yuquan. 2015. 西夏语中的选择连词 mo2 Xīxiàyǔ zhōng de xuǎnzé liáncí mo2 [The alternative conjunction mo2 in Tangut]. 语言研究 *Yǔyán yánjiū* [Research in Language] 35(1). 123–126.
- Gates, Jesse P. 2012. *Situ in situ: Towards a dialectology of Jiārǒng (rGyalrong)*. Master's thesis, Trinity Western University.
- Gates, Jesse P. 2017. Verbal triplication morphology in Stau (Mazi dialect). *Transactions of the Philological Society* 115(1). 14–26.
- Gates, Jesse P. 2021. *A grammar of Mazur Stau*. Doctoral Dissertation, École des Hautes Études en Sciences Sociales.
- Gates, Jesse P. & Won Ho Kim. 2018. Vowel harmony in Stau. *Linguistics of the Tibeto-Burman Area* 41(2). 263–293.
- Gong, Hwang-cherng. 1994. A hypothesis of three grades and vowel length distinction in Tangut. *Journal of Asian and African Studies* 46–47. 305–314.
- Gong, Hwang-cherng. 1997. 西夏語若干韻母轉換的起源-重疊複合詞 Xīxiàyǔ ruògān yùnmǔ zhuǎnhuàn de qǐyuán – chóngdié fùhécí [the origin of some Tangut rime alternations: reduplicated compound words]. 中國境內語言暨語言學 *Zhōngguó Jìngnèi yǔyán jì yǔyánxué* [Language and linguistics in China] 4. 265–289.

- Gong, Hwang-chenng. 1999. 西夏語的緊元音及其起源 Xīxiàyǔ de jǐnyuányīn jíqí qǐyuán [Tense vowels in Tangut and their origin]. *Bulletin of the Institute of History and Philology* 70(2). 531–558.
- Gong, Hwang-chenng. 2001. 西夏語動詞的人稱呼應與音韻轉換 Xīxiàyǔ dòngcí de rénchēng hūyīng yǔ yīnyùn zhuǎnhuàn [Personal agreement and phonological alternations in the Tangut Verb]. *Language and Linguistics* 2(1). 21–67.
- Gong, Hwang-chenng. 2002. *Collected papers on Tangut philology. Language and Linguistics Monograph Series Number C2-1*. Institute of Linguistics (Preparatory Office).
- Gong, Hwang-chenng. 2003. Tangut. In LaPolla, Randy & Graham Thurgood (eds.), *The Sino-Tibetan languages*, 602–620. London: Routledge.
- Gong, Xun. 2014. Personal agreement system of Zbu rGyalrong (Ngyaltsu variety). *Transactions of the Philological Society* 112(1). 44–60.
- Gong, Xun. 2017. Verb stems in Tangut and their orthography. *SCRIPTA* 9. 29–48.
- Gong, Xun. 2018. *Le rgyalrong zbu, une langue tibéto-birmane de Chine du Sud-ouest: une étude descriptive, typologique et comparative*. Doctoral Dissertation, Institut national des langues et civilisations orientales.
- Gong, Xun. 2020. Uvulars and uvularization in Tangut phonology. *Language & Linguistics* 21(2). 175–212.
- Honkasalo, Sami. 2019. *A grammar of Eastern Geshiza*. Doctoral Dissertation, University of Helsinki.
- Jacques, Guillaume. 2007. *Textes tangoutes I, Nouveau recueil sur l'amour parental et la piété filiale*. Languages of the World/Text Collections 25. München: Lincom Europa.
- Jacques, Guillaume. 2008. 嘉絨語研究 *Jiāróngyǔ yánjiū* [Study on the Rgyalrong language]. Beijing: Minzu chubanshe.
- Jacques, Guillaume. 2009. The origin of vowel alternations in the Tangut verb. *Language and Linguistics* 10(1). 17–28.
- Jacques, Guillaume. 2010a. The inverse in Japhug Rgyalrong. *Language and Linguistics* 11(1). 127–157.
- Jacques, Guillaume. 2010b. The three sisters. http://lacito.vjf.cnrs.fr/pangloss/resource/crdo-JYA_ARURNWS_5OUND&crdo-JYA_ARURNWS.
- Jacques, Guillaume. 2011. The structure of the Tangut verb. *Journal of Chinese Linguistics* 39(2). 419–441.
- Jacques, Guillaume. 2013. Harmonization and disharmonization of affix ordering and basic word order. *Linguistic Typology* 17(2). 187–217.
- Jacques, Guillaume. 2014. *Esquisse de phonologie et de morphologie historique du tangoute*. Leiden: Brill.
- Jacques, Guillaume, Anton Antonov, Yunfan Lai & Lobsang Nima. 2017. Stau (Ergong, Horpa). In Thurgood, Graham & Randy LaPolla (eds.), *The Sino-Tibetan languages*, 2nd edn., 597–613. London: Routledge.
- Jacques, Guillaume & Zhen Chen. 2004. 茶堡話的重疊形式 Chāpǔhuà de chóngdié xíngshì [Reduplication in Japhug]. 民族語文 *Mínzú yǔwén* [Minority Languages of China] 4. 7–11.
- Kychanov, Ievgenij Ivanovich. 1974. *Вновь собранные драгоценные парные изречения [Recollected precious pair sayings. Translation from the Tangut]. «Памятники письменности Востока» [Series “Monuments of written oriental languages”]*. Москва: Наука [Moscow: Nauka].
- Lai, Yunfan. 2013. 俄热話的輔音重疊 Èrèhuà de fǔyīn chóngdié [Consonant reduplication in Wobzi]. 民族語文 *Mínzú yǔwén* [Minority Languages of China] 6. 12–18.

- Lai, Yunfan. 2015. The person agreement system of Wobzi Lavrung (Rgyalrongic, Tibeto-Burman). *Transactions of the Philological Society* 113(3). 271–285.
- Lai, Yunfan. 2017. *Grammaire du khroskyabs de Wobzi*. Doctoral Dissertation, Université Paris III.
- Lai, Yunfan. forthcoming. Come is the new go: The evolution of stem alternation of basic motion verbs in Khroskyabs. *Language and Linguistics*.
- Lai, Yunfan, Mei-Shin Wu & List Johann-Mattis. 2018. Rethinking the orientational prefixes in Rgyalrongic languages: The case of Siyuewu Khroskyabs. *Paper presented at the 51th International Conference on Sino-Tibetan languages and linguistics*, 9 2018.
- Laufer, Berthold. 1916. The Si-hia language, a study in Indo-Chinese philology. *T'oung Pao* 17(1). 1–126.
- Li, Fanwen. 1997. 夏漢字典 *Xià Hàn zìdiǎn* [A Tangut-Chinese dictionary]. Beijing: Zhongguo shehui kexue chubanshe.
- Li, Jingxun. 1993. 四川通史 *Sìchuān Tōngshǐ* [A general history of Sichuan]. Chengdu: Sichuan University Press.
- Lin, Xiangrong. 1993. 嘉戎語研究 *Jiāróngyǔ yánjiū* [A study on the Rgyalrong language]. Chengdu: Sichuan minzu chubanshe.
- Lin, Youjing. 2016. 嘉戎語卓克基話語法標注文本 *Jiāróngyǔ Zhuókèjīhuà yǔfǎ biāozhù wénběn* [Cogtse Rgyalrong texts: Fully analyzed spontaneous narratives with an updated sketch grammar of the language]. Beijing: Shehui Kexue chubanshe.
- Matisoff, James A. 2004. “Brightening” and the place of Xixia in the Qiangic branch of Tibeto-Burman. In Lin Ying-chin, Fang-min Hsu, Chun-chih Lee, Jackson T.-S. Sun, Hsiu-fang Yang & Dah-an Ho (eds.), *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-Cherng Gong on His Seventieth Birthday*. Taipei: Institute of Linguistics, Academia Sinica.
- Mithun, Marianne. 1984. The evolution of Noun Incorporation. *Language* 60(4). 847–894.
- Nishida, Tatsuo. 1976. Hsihsia, Tosu and Lolo-Burmese languages. *Studia Phonologica* 10. 1–15.
- Ruth, Dannel. 1994. The Hsi-Hsia. In Herbert Franke & Denis Twitchett (eds.), *The Cambridge History of China: Alien regimes and border states, 907-1368*, vol. 6, 154–214. New York and Oakleigh: Cambridge University Press.
- Sagart, Laurent, Guillaume Jacques, Yunfan Lai, Robin J. Ryder, Valentin Thouzeau, Simon J Greenhill & Johann-Mattis List. 2019. Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *Proceedings of the National Academy of Sciences* 116(21). 10317–10322.
- Sun, Hongkai. 1991. 从词汇比较看西夏语与藏缅语族羌语支的关系 *Cóng cíhuì bǐjiào kàn Xīxià yǔ yǔ Zàngmiǎn yǔzú Qiāng yǔzhī de guānxì* [The relation between Tangut and Qiangic (Tibeto-Burman) through lexical comparison]. 民族语文 *Mínzú yǔwén* [Minority Languages of China] 2(1). 1–11.
- Sun, Jackson T.-S. 2000a. Parallelisms in the verb morphology of Sidaba rGyalrong and Lavrung in rGyalrongic. *Language and Linguistics* 1(1). 161–190.
- Sun, Jackson T.-S. 2000b. Stem alternations in Puxi verb inflection: Toward validating the rGyalrongic subgroup in Qiangic. *Language and Linguistics* 1(2). 211–232.
- Sun, Jackson T.-S. & Shidanluo. 2002. 草登嘉戎語與「認同等第」相關的語法現象 *Cǎodēng Jiāróngyǔ yǔ rèntóng děngdì xiāngguān de yǔfǎ xiànxàng* (Empathy hierarchy in Caodeng rGyalrong grammar). *Language and Linguistics* 3(1). 79–99.
- Wang, Jingru. 1933. 論四川羌語及弭藥語與西夏語 *Lùn Sìchuān Qiāng yǔ jí Mǐyàoyǔ yǔ Xīxià yǔ* [The Shishiah and the Ch’iang and Minia Languages in Szüch’uan Province]. In *Xīxià Yánjiù*, vol. 2, 278–288. Pei-p’ing: Academia Sinica.

- Wolfenden, Stuart N. 1931. On the Tibetan transcriptions of Si-Hia words. *Journal of the Royal Asiatic Society* 63(1). 47–52.
- Zhang, Shuya. 2016. *La phonologie et la morphologie du dialecte de Brag-dbar du rgyalrong situ*. Master's thesis, Université Paris III.
- Zhang, Shuya. 2019. From proximate/obviative to number marking: Reanalysis of hierarchical indexation in Rgyalrong languages. *Journal of Chinese Linguistics* 47(1). 125–150.
- Zhang, Shuya. 2020. *Le rgyalrong situ de brag-bar et sa contribution à la typologie de l'expression des relations spatiales: l'orientation et le mouvement associé*. Doctoral Dissertation, Institut national des langues et civilisations orientales.
- Zhang, Menghan, Yan Shi, Wuyun Pan & Li Jin. 2019. Phylogenetic evidence for Sino-Tibetan origin in northern China in the Late Neolithic. *Nature* 569. 112–115.