Tangut and Horpa languages: Some shared morphosyntactic features

Mathieu Beaudouin

Tangut and Horpa languages: some shared morphosyntactic features

Mathieu Beaudouin

Inalco & Crlao (Paris, France)

Fieldwork from the past decade has yielded new data from a cluster of languages in Western Sichuan (China), resulting in new observations relevant for the understanding of Tangut grammar. In this paper, I intend to present morphosyntactic evidence pointing to the Tangut language's membership within the Horpa taxon, located within the larger Gyalrongic group of the Qiangic branch of Sino-Tibetan. Tangut exclusively shares with Horpa languages cognates that are far too peculiar to be the result of mere chance. By successively considering the verbal, nominal, and postpositional domains, the present paper highlights evidence that links Tangut to Horpa, while proposing new paths to the understanding of grammatical categories of Tangut proper, such as orientational/aspectual preverbs.*

Keywords : Tangut, Horpa and West Gyalrongic languages, orientation/TAME preverbs, agreement history, locative cases

Introduction

Tangut is a medieval non-Sinitic Sino-Tibetan language which was spoken in the Western Xià empire (1038-1227 AD). Though its syntax has been quite well understood since studies as early as Morisse (1904), the most important contribution to the understanding of its grammar today has been Kepping (1985). Though the main focus of her study is said to be on morphology (she named her work *морфология* 'morphology'), she made important observations with implications for other aspects of Tangut grammar which are still of value today. However, numerous features of Tangut have remained quite difficult, if not impossible, to account for through the sole use of documents written in the language.[1]

*I want to express my gratitude towards all those who helped improve this paper: Sami Honkasalo, Lai Yunfan, Jesse Gates, Nie Hongyin, Guillaume Jacques, Li Shang, Gong Xun, Zhang Shuya, Mark Miyake, Elizabeth Zeitoun, Xiao Xiao, Emmett Strickland, and of course the two anonymous *Language and Linguistics* reviewers. Naturally, all remaining errors are entirely mine. Except for an ensemble of abbreviations listed in Section 5, the glosses used in the present paper follow the *Leipzig Glossing Rules* (see https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf).

[1] For example, the two different possible values for each series of the orientational system (see Subsection 1.2.6).

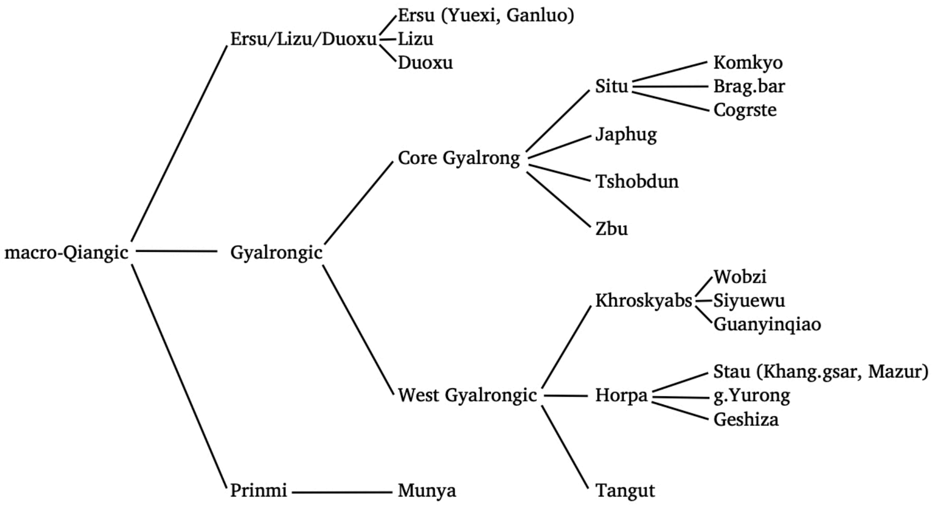

Figure 1: Overview and classification of the languages of the present study

Over the two last decades, scholars have made considerable progress in the description of non-written languages spoken in Western Sichuan, which have long been suspected of being close relatives of Tangut. Laufer (1916) already included Tangut in a group also comprising the Mosuo and Lolo languages, and since then, other scholars have used modern languages to improve our understanding of Tangut. For example, Nishida (1973, 1976) highlighted some shared features between Tangut and Duoxu. However, the first systematic comparison of Tangut with Japhug, a language of the Gyalrongic clade, dates only to Jacques (2014).

The Gyalrongic subgroup within Qiangic was first proposed by Sun (2000a, 2000b), who listed five common characteristics shared by the Horpa, core Gyalrong, and Khroskyabs languages: glottal inversion, tonal inversion, ablaut, aspiration polarity, and parallelism between the verbal past and progressive stems. Recently, Sun (2019) provided new evidence supporting the existence of the Horpa subgroup within Gyalrongic, proposing new insights on the history of tonal polarity.[2] Lai (2017) and Jacques et al. (2017) gave further evidence, both of lexical and morphological nature (desyllabification of preinitials) that led to the subgrouping of Khroskyabs and Horpa into a shared clade, itself genetically linked to core Gyalrong. At the end of Jacques et al. (2017: 611), they mention grammatical cases potentially shared by Horpa languages and Tangut.

Figure 1 gives an overview of the languages mentioned in the present study, along with their classification. I argue here that Tangut and Horpa languages should placed together in the same clade, either Horpa itself, or a superior clade distinct from Khroskyabs. The classifications proposed by these two hypotheses

[2]Several different names have been given for this subgroup. I employ the 'Horpa' label first proposed by Sun (2007b).

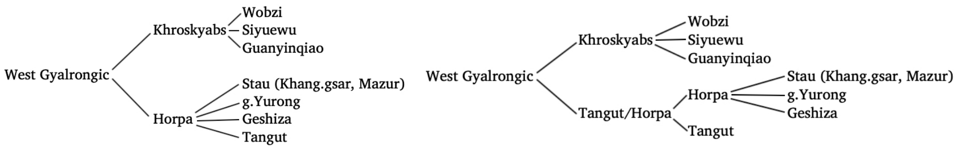

Figure 2: The genetic position of Tangut: two possibilities

are presented in Figure 2.

The lexical proximity between Tangut and the Horpa languages is quite apparent if one compares, for example, the lexicon given in Sun (2019) with Tangut.[3] However, the documentation of the Horpa languages being very recent, with the only complete grammars written by Honkasalo (2019), Tunzhi (2019), and now Gates (2021), it has only now become possible to use Horpa languages to understand Tangut's grammar and genetic relationships more fully.

Indeed, all the shared characteristics existing between Tangut and West Gyalrongic, especially Horpa, not only help to establish a link between Tangut and Horpa (the first goal of the present paper), they also justify the (cautious) use of these living languages as a methodological tool for hypotheses regarding Tangut *per se*, in order to refine a synchronic account of Tangut grammar (the second goal of the present paper).

This work was conducted independently from Lai et al. (2020), and its scope is different. First, as stated above, the end goal of the approach used here is twofold, being as much about Tangut's grammar as its phylogenetic position. Second, the data used focuses on Tangut and Horpa within West Gyalrongic, while also exploring outside Gyalrongic. It contains fewer examples from non-Horpa West Gyalrongic languages, while incorporating some data from Lizu/Ersu and Munya (Qiangic). The combination of these factors can sometimes lead to different conclusions than what was found in Lai et al. (2020).

The corpus employed is mainly composed of my transcriptions of the [illegible] *Djịj*^{1} *bo*^{1} (類林 *Lèilín*, "Forest of Categories") in the edition of Shi et al. (1993), and the [illegible] *Yạ*^{2}*njị*^{1} *lʰjịj* (十二國 *Shíèr guó*, "The Twelve Kingdoms") in the edition of Solonin (1995). These two texts were translated from Chinese to Tangut, in a way making them closer to the language spoken by the Tanguts (Lin & An 1992).[4] For each example extracted from these documents, the translation of the

[3]This is a topic to be discussed in detail in a future paper; a preview of this lexical proximity —between Geshiza Horpa and Tangut—can be seen below in Table 8 of the present article.

[4]Kepping was the first scholar to use this kind of texts for grammatical description. The reason for such a choice is well explained in Kepping (1985: 17): "В нашей работе впервые в тангутоведении в качестве материала использованы переводы китайских неканонических сочинений. В отличие от переводов канонических сочинений произведения неканонические, светские, обычно переводились весьма вольно, а то и перелагались без соблюдения текстуальной точности и особенностей оригинала." *"In our work, for the first time in Tangut studies, translations of Chinese non-canonical works were used as source material. Unlike translations of canonical works, non-canonical and secular works were usually translated quite freely, or even shifted without observing the textual accuracy and peculiarities of the*

editor precedes my translation. I also conducted searches in the [illegible] *Sjiw¹ śio̱¹ njij² ·wə̱¹ la¹ mjij² X* (新集慈孝記下卷 *Xīn jí cí xiào jì xià juàn*, "New Collection on Parental Love and Filial Piety"), edited by Jacques (2007), a document which pertains to the same 'oral-like' category. Finally, some analyses used the parts transcribed by Downes (2018) of the autochthonous Tangut Code (also referred to here as legal texts), and the Tangut version of the *Avataṃsaka Sūtra*, with the Japanese translation by Arakawa (2011). The Tangut phonetic reconstruction I follow is the Gong Hwang-Cherng (cf. Gong 2006) system as it appears in Li (2008).

The present study is structured as follows. After undertaking a systematic comparison of the verbal morphology (Section 1), I discuss numerals and nominalization (Section 2), before analyzing locative phrases (Section 3).

1 Verb

1.1 The Tangut verbal template

Table 1 presents the structure of the Tangut verb, which is templatic in nature (Bickel & Nichols 2007). Each morpheme in the template is assigned to only one slot, but not all slots need to be filled for each instance of a given verb. Each slot of this template will be subject to a cross-analysis with similar morphemes found in West Gyalrongic languages.

Slot 1 is occupied by two series of orientational preverbs (Type-A/B orientational preverbs), traditionally known (Kepping 1985) to encode the perfective aspect for the first series, and optative mood for the second (a description I will refine in Subsection 1.2.6).[5] The orientational value is available only when employed with some verbs, typically motion verbs. Type-B preverbs are derived from Type-A ones by fusion with a former irrealis/interrogative morpheme **-i-* (see Subsection 1.2.6).

original." Note that this characteristic should make any linguist always depart from the Tangut text in his translations, and consult the Chinese original only as a secondary reference.

[5]With regard to preverbal positions, the terms 'preverb' and 'prefix' should be seen as synonyms in the present study. The usage in the practice of Tangut description is to refer to 'prefixes', where the term 'preverb' should be favored as these so-called prefixes only occur in preverbal position.

1: TAM (dir)		2: NEG	3: modal	4: noun inc.	5: Σ	6: person	7: TEL/FUT	8: IFR	9: PROG
PFV / IMP	[illegible] *·a* (up)								
	[illegible] *nja*^{1} (down)								
	[illegible] *·wjɨ*^{2} (centrf)								
	[illegible] *kjɨ*^{1} (centrp)	[illegible] *mjɨ*^{1}				1SG [illegible] *ŋa*^{2}			
	[illegible] *rjɨr*^{2} (unspec.)								
	[illegible] *dja*^{2} (/)	[illegible] *mjij*^{2}			Σ[[illegible]]	2SG [illegible] *nja*^{2}			
	[illegible] *djɨ*^{2} (/)		POT [illegible] *tśʰjɨ*^{1}	*σ*			[illegible] *·jij*^{1}	[illegible] *sji*^{2}	[illegible] *djij*^{2}
INTRG / OPT	[illegible] *·jij*^{1} (up)	[illegible] *mjɨ*^{1}	CONC [illegible] *ljɨ*^{1}		Σ[B]	1DU [illegible] *kjɨ*^{1}			
	[illegible] *njij*^{2} (down)		[illegible] *tsjɨ*^{1}						
	[illegible] *·wjij*^{2} (centrf)	[illegible] *tji*^{1}				2DU [illegible] *tsjɨ*^{1}			
	[illegible] *kjij*^{1} (centrp)								
	[illegible] *rjijr*^{2} (unspec.)					12PL [illegible] *nji*^{2}			
	[illegible] *djij*^{2} (/)								
INTRG	[illegible] *·a*								

Table 1: Maximal template of the Tangut verb

Slot 2 is filled by negation: general negation [Tangut] *mjɨ*^{1}, past/perfective negation [Tangut] *mjij*^{2} (NEG.PFV), modal negation [Tangut] *mjɨ*^{1} (NEG.MOD) and prohibitive [Tangut] *tjɨ*^{1}. Modals can occupy slot 3: the potential [Tangut] *tśʰjɨ*^{1} (POT), the concessive [Tangut] *ljɨ*^{1} (CONC), and a morpheme [Tangut] *tsjɨ*^{1} whose role remains unclear. Each of these modal preverbs usually collocate with [Tangut] *mjɨ*^{1}- in negative configurations.[6] Slot 4, which is the final prefixal position, can be filled with a monosyllabic incorporated noun, which I will not discuss here. An example illustrating the succession of prefixal morphemes is given in (1):

(1) [Tangut] [Tangut] [Tangut] [Tangut] [Tangut]

tjị^{1} *rjɨ*^{r2}*-mjɨ*^{1}*-tśʰjɨ*^{1}*-tʰji*^{1}
food PFV:DIR-NEG.MOD-POT-eat

"…я не ел пищу." (Solonin 1995: 39)

"…I could not eat food." (12K, 132.22.05)

Next comes the stem (slot 5), followed by the person agreement suffix (slot 6), which precedes a suffix marking telic aspect (slot 7), traditionally described as a future marker.[7] The next morpheme (slot 8) is a suffix formerly referred to in the literature as perfective, but that will be analyzed here as inferential (IFR, see Subsection 1.3.3). Finally, this suffix can be followed by the progressive [Tangut] *-djij*^{2} (slot 9). Two examples illustrating suffix ordering are given in (2)—slots 6-7-8, and (3)—slots 8-9.

(2) [Tangut] [Tangut] [Tangut] [Tangut] [Tangut]

mə^{1} *·wu*^{2}*-nja*^{2}*-·jij*^{1}*-sji*^{2}
sky help-2-TEL-IFR

天將助矣 (Shi et al. 1993: 290)

"Heaven will help you." (Leilin, 06.15B.7)

[6] There are two facts that indicate that modal morphemes belong to the same slot. Firstly, as stated above in this article, each of these modal preverbs usually collocate with [Tangut] *mjɨ*^{1}- in negative configurations. Secondly, two modal morphemes occuring consectutively is unattested.

[7] Arakawa (2014) mentions examples where a negation follows a verb and is attached to an auxiliary verb, then puts this negation inside the verbal template, after the main verb. I believe these examples can be seen as clues indicating that the auxiliary is independent (i.e., not as a part of the template). Indeed, the auxiliary verbs, even if they are semantically dependent to the verb they modify, behave as verbs from a templatic point of view. In [Tangut] [Tangut] [Tangut] [Tangut] *tsʰjɨɨ*^{1} *njwi*^{2}*-ŋa*^{2}*=·ji*^{2} (recite can-1=say) 'I can recite' from the *Leilin* (04.28A.4, 4-7) the auxiliary [Tangut] bears an agreement suffix, as would any independent verb. The consistent lack of agreement for the main verb, on the other hand, only indicates an infinitive form.

(3) [illegible]...

·we² dzju² ɣar²=gu² *bə²lụ¹ dja²-tś ʰju¹ ŋowr²*
city lord belly=MEDE insect PFV-EXV whole
mjij²-ljɨ¹-śjɨj¹-sji²-djij²...
NEG.PFV-CONC-become-IFR-PROG...

府君胃中有蟲欲成 (Shi et al. 1993: 289)

"There are worms in my commander's belly. Even if they have not completely grown yet..." (Leilin, 06.11B.7)

Some morphemes, always occurring after the verb (see Subsection 1.3.4), are analyzed as enclitics, i.e., not part of the verbal template. Note, however, that very little research has been done on word boundaries in Tangut and that this question may be subject to reevaluation in future work. As the affix ordering is quite similar in the two main groups used for comparison (Khroskyabs and Horpa), the analysis will follow the frame of the verbal template. Subsection 1.2 will focus on prefixal slots, and Subsection 1.3 on suffixal positions (in which I include suffixes as well as enclitics).

1.2 Prefixal slots

1.2.1 Orientational preverbs: an overview

The series of so-called 'orientational preverbs' occupies the first position of the template.[8] Table 2 lists these preverbs, in parallel to those of Horpa (Geshiza, g.Yurong, and Mazur Stau varieties), Khroskyabs (Guanyinqiao and Wobzi varieties), core Gyalrong (Tshobdun and Japhug varieties), Ersu (Ganluo variety), and Munya.[9]

These preverbs tend to distribute themselves according to subsystems which originate from three major types (Sun 2003), namely, the solar (East, West, North, South), the riverine (upstreams, downstreams), and the vertical (up, down). At first sight, it is worth noting that the preverb system of Munya, a language many researchers have argued to be closely related to Tangut in the past, is the most distant to Tangut.

[8] The system formed by Tangut's orientational preverbs has been described in detail by Kepping (1985: 176-203, 208-216). The present work is the first to reconsider some of her conclusions.

[9] Geshiza Horpa data is from Honkasalo (2019), Mazur Stau Horpa data from Gates (2021), Wobzi Khroskyabs, Guanyinqiao Khroskyabs, and g.Yurong Horpa data from Lai (2017), Tshobdun Gyalrong data from Sun (2007a), Japhug Gyalrong data from Jacques (2021), Kyomkyo Situ data from Prins (2016), Brag-bar Situ data from Zhang (2020), Ganluo Ersu data from Chirkova & Wang (2017), and Munya data from Bai (2019). The place of the Ersu/Lizu/Duoxu cluster (abbreviated Ers/Lz/Dx in Table 2) within Qiangic is to date still questioned, hence the precedence of 'macro' when I include data from one of these languages.

		dir.0		dir.1: up	*INTRG*	dir.2: down	dir.3	dir.4	dir.5	dir.6	dir.7
Horpa	Tangut	rʲɨr^2-		·a-	← *·a-*	nʲa^1-	kʲɨ^1- (inw.)			·wʲɨ^2- (outw.)	dʲa^2- (PFV)
	Geshiza	(rə-)	=	rə-	*æ-*	næ-	gæ- (upstr.)			wə- (downstr.)	dæ- (PFV)
	g.Yurong			ə-	← *ə-*	nə-	gə- (N)			ɣə- (S)	də- (PFV)
	Mazur Stau	(rə-)	=	rə-	*æ-*	nə-	kə- (NW)			ɣə- (SE)	tə- (PFV)
Khroskyabs	Guany.	rə-		a-	← *a-*	na-	ka- (upstr.)	nə- (downstr.)	la- (h. alt.)	və- (l. alt.)	
	Wobzi	rə-		æ-	← *æ-*	næ-	kə- (upstr.)	nə- (downstr.)	læ- (h. alt.)	və- (l. alt.)	
Gyalrong	Tshobdun			tə-		nɐ-	kə- (E)	nə- (W)	lɐ- (upstr.)	tʰɐ- (downstr.)	←
	Japhug			tɤ-		pɯ-	kɤ- (E)	nɯ- (W)	lɤ- (upstr.)	tʰɯ- (downstr.)	←
(Situ)	Kyomkyo	ro- (rə-)		to-		na-	kə- (upstr.)	nə- (downstr.)			
	Brag-bar	rɐ-	=	rɐ-		na-	wo- (upstr.)	nə- (downstr.)			
Ers/Lz/Dx	Ganluo Ers.			dɛ-		nɛ-	kʰɛ- (inw.) kʰwa˞- (N, upstr.)	ni- (obl.down)	dʑi- (obl.up)	ŋɛ- (outw.) ŋwa˞- (S, downstr.)	tʰɛ- (PFV)
Prinmi	Munya	ro-		tə-		no-	ɣɤ- (upstr., inw., E) kʰɯ- (unspec.)			ɛ- (downstr, outw., W) ngɯ- (cisl.)	tʰo- (transl.)

Table 2: Orientational preverbs in macro-Qiangic (cognacy marked in red/blue, semantic oppositions marked with column separators)

Building on Arakawa (2012)'s insight that 𗗙 *dji²*- encodes unspecified orientation, I also analyze 𗈪 *dja²*- as orientationally unspecified.[10] As discussed below (Subsection 1.2.4), the direction for that morpheme is indeed not as straightforward compared to the other orientational prefixes.

1.2.2 The history of dir.0 **rV-*, dir.1 **tV-*, and dir.2 **nV-*

The cognacy for the direction 'down' is quite transparent: the prefixes all share the same initial (except for Japhug), and the correspondence Tangut (T.) *a* :: Geshiza (G.) *æ* is widely attested after coronal onsets between the two languages (see Table 7). As for the direction 'up', the configuration is more intricate. Geshiza, Gyalrong, and other macro-Qiangic languages display an apparent discrepancy, with a coronal onset not found in other West Gyalrongic languages (i.e., Tangut, g.Yurong, and Khroskyabs languages). I believe Tangut's unspecified preverb 𗅋 *rjɨr²*, Geshiza *rə*-, and Mazur Stau *rə*- to be cognates. Below is the hypothesis I propose to explain the distribution.[11]

1) The archaic form in macro-Qiangic for 'vertically up' was **tV-*. This form still exists in Munya, Tshobdun, Japhug, and Kyomkyo (it is voiced in Ersu). It coexisted in the system with another preverb **rV-*, present in Situ languages ('towards the mountain'), Tangut and Khroskyabs languages (unspecified direction), and Munya ('in a circle', now unproductive).[12]

2) The initial consonant of **tV-* lenited into **rV-*. The correspondence Gyalrong *t-* :: Khroskyabs *r-* :: Horpa *r-* is actually attested.[13] If we look for other cognates in the languages listed in Table 2, another example can be found with the numeral 'one', with Japhug *tɤɣ* :: Wobzi *rɑ̂ɣ* :: Geshiza *rəu* (Tangut *lew¹*).[14] From a diachronic point of view, an intervocalic **t- > r-* change (with a probable flap stage

[10] 𗗙 *dji²*- is an autobenefactive derivation of 𗈪 *dja²*- seen with a reduced subset of verbs. In the present study, 𗈪 *dja²*- and 𗗙 *dji²*- should be understood as two manifestations of the same category.

[11] What follows in this sub-subsection does not try to provide exhaustive evidence, but only a scenario constructed by abduction.

[12] The more conservative orientation system of Cogrtse Situ also presents the opposition found in Kyomkyo between *ro*- 'towards the mountain' and *rə*- 'towards the river'. This second orientation seems to be unrelated to all the other languages of Table 2, which specify an upward-like direction for dir.0 **rV-*. As Tangut's dir.0 remains unspecified (not allowing the exclusion of any of the potential cognates), I indicate in Table 2 this downward-like value between parentheses, even if it is unlikely that it is cognate with Tangut's dir.0. Brag-bar Situ is representative of the stage when its ancestor departed from proto-Situ (which illustrated by Komkyo and Cogrtse), by loosing this second preverb. The original pair's surviving prefix was inherited by West Gyalrongic languages, together with its acquired homonymy with dir.1.

[13] As mentioned by an anonymous reviewer, the change one can infer from this correspondence is problematic, as indeed the syllable onset position cross-linguistically favors fortition rather than lenition. However, very few sentences begin directly with a preverb, whose initial can therefore often be in an intervocalic position. In any event, this does not invalidate the correspondence itself, which I leave unexplained for now until future fieldwork yields more data.

[14] The labio-velar found in Tangut and Geshiza is a shared innovation, from a former *-ɣ. The former coda still exists in other Horpa languages (see Sun 2019).

T. :: G.	Tangut	Geshiza	Meaning
r- $^{-r}$ :: r-	r^{j}ar^{1}	ræ	‘to write’
	r^{j}ar^{2}	ræ	‘turnip’
	r^{j}ijr^{1}	rji	‘horse’
·^{j}- $^{-r}$ (< r^{j}- $^{-r}$) :: rj-	.^{j}ir^{2}	rjə	‘hundred’
	·^{j}ar^{1}	rjɛ	‘eight’
	·^{j}ər^{1} / ·^{j}ər^{2}	rjæ	‘to ask’
	wor^{1}·^{j}a̠r^{2}	wərja	‘chicken’
·^{j}- :: j-	·^{j}ɨ^{2}	jə	‘to say’
	·^{j}ij^{2}	ji/jæ	‘sheep’

Table 3: Lateral and approximant onsets correspondences between Tangut and Geshiza Horpa

**r-*) is documented in core Gyalrong Brag-bar Situ (Zhang 2020: 38, 462-464), even if the precise conditions of the sound change are still under investigation.

3) The phonetic and semantic proximity in Geshiza and Mazur Stau between **rV-* ‘dir.0’ and the lenited *rə-* (‘up’) resulted in a unique unanalysable *rə-* which still presents characteristics of the ‘dir.0’ encoded in Situ. Indeed, in Geshiza *rə-* means ‘away from the river’, i.e., the same meaning as ‘towards the mountain’ in the configuration of a valley. Actually, that exact change happened also in Brag-bar Situ. The only difference is that comparison allows for analyzing two different *rɐ-* in Brag-bar Situ, while there is no comparative data available for Geshiza and Mazur Stau.

4) A reanalysis happened in Tangut, g.Yurong, and the Khroskyabs languages between two morphemes occupying the same slot of the template: the interrogative and the direction ‘up’. In Tangut, only one character [illegible] exists for the two morphemes (i.e., the interrogative and the orientational) encoded by the preverbal syllable *·a*. The reanalysis may come from the fact that **rV-* continued to lenite until the proximity between the interrogative and the orientational became too close in the speakers’ minds. Inside West Gyalrongic, a potential former lateral can be found in the optative counterpart *la-* of ’Jorogs Khroskyabs *o-* ‘upwards’, which is evidence supporting this interpretation. In any case, the T. *·a* :: G. *rə-* correspondence is problematic, as the Geshiza initial should be reflected as a rhotacisation in Tangut (see Table 3), but is reflected neither in Type-A preverb [illegible] *·a* nor in Type-B preverb [illegible] *·jij¹*.

1.2.3 Dir.3 to dir.6: a semantic and systemic analysis

1.2.3.1 Innovations by loss and reanalysis

The directions 3, 4, 5, and 6 underwent semantic shift due to considerations proper to each language and speakers’ geographical position. Thus ‘direction 3’, which encodes the centripetal in Tangut (‘inwards’), refers to the North in

		(Tangut/Horpa) dir.3/dir.6 /		
macro-Qiangic (Ersu) dir.3/dir.6 dir.4/dir.5	↗ *loss* ↘ *reanalysis*		↗ *reanalysis*	(Tshobdun/Japhug) dir.3/dir.4 dir.5/dir.7
		(Gyalrong/Khroskyabs) dir.3/dir.4 dir.5/dir.6	⟶ *loss*	(Situ Gyalrong) dir.3/dir.4 /

Figure 3: Reanalysis and losses of the macro-Qiangic preverbs system (dir.3 to 6)

Khang.gsar Horpa, the East in Tshobdun Gyalrong, and an upstream direction in Guanyinqiao Khroskyabs. We can nonetheless establish a common origin not only because the preverbs for that particular slot display semantic and sound correspondences but also parallel oppositions with their counterpart. For example, while *ka-* means 'upstream' in Guanyinqiao Khroskyabs and *kə-* 'East' in Gyalrong, their counterparts (nə- for both languages) mean 'downstream' in Khroskyabs and 'West' in Tshobdun Gyalrong. Considering these 'grouped switches', cognacy can be induced for directions 3 and 6. The only breaks are an autonomous 'downstream' preverb in core rGyalrong (Tshobdun *tʰɐ-* Japhug *tʰɯ-*), and incomplete parallelism for dir.3 and dir.6 in Khroskyabs languages (e.g., Wobzi *kə-* 'upstream' vs. *və-* 'low altitude'). The reanalysis hypothesis presented in Figure 3 explains all these discrepancies.

One of the most interesting points in the distribution relates to directions 4 and 5 ('downstream' and 'high altitude' in Khroskyabs). The loss of dir. 4 and dir. 5 in both Tangut and Horpa languages could be analyzed as a shared innovation. This loss is probably synchronic to a readjustment which happened in Khroskyabs and Gyalrong, of the former antinomy between dir. 3 and 6 (whose meaning is not so far from that of dir. 4).[15] An alternative analysis is that the

[15]Note the perceptual compatibility between the labels 'oblique down' & 'oblique up' of Ganluo Ersu, and 'downstream' in Khroskyabs languages & 'upstream' in Gyalrong languages. The correspondence Khroskyabs/Gyalrong *l-* :: Ersu *dʑ-* is not straightforward, even if it is systemically suggested by the distribution of the prefixes. There are examples of *dʑ-* :: *l-* correspondences between Duoxu, Ersu, and Lizu, but in the other direction (Ersu has the lateral); e.g., for 'head', Duoxu *wu*^{53}*dʑu*^{32} :: Ersu *ɣ̥lí* :: Lizu ^{LP}*wuli*. However, at the same time, the orientational preverb *dʑi-* of Ganluo Ersu seems to be cognate to Duoxu's orientational preverb *dʑi-*, which would indicate a permeability between the two sounds *l-* and *dʑ-* in the cluster of languages. The possibility of palatalization is not so odd, knowing that orientational preverbs do not begin an utterance, which makes the sound change – if it took place – correspond to the pattern of Duoxu *wu*^{53}*dʑu*^{32}.

emergence of dir. 4 and 5 is an innovation shared exclusively by Ersu, core Gyalrong, and Khroskyabs. Still, this view is problematic as Ersu is not known as a Gyalrongic language whereas Horpa, Khroskyabs, and core Gyalrong are known as Gyalrongic languages.

The same pattern of innovation by loss can be seen in Situ for dir.5 and dir.6. This loss is probably the conclusion of a reanalysis cycle where the two first oppositions dir.3/dir.6 (common to Tangut, Horpa languages, Ersu, and Munya) and dir.4/dir.5 were replaced by the oppositions dir.3/dir.4 and dir.5/dir.6 (seen in Khroskyabs languages and Gyalrong languages). While Tangut and Horpa did not reanalyze the configuration of these four directions due to loss of dir.4 and dir.5 preverbs, Situ lost dir.5 and dir.6 after the completion of the cycle, while Tshobdun and Japhug's dir.6 was filled through reanalysis by a cognate of the Ganluo Ersu perfective *tʰɛ-* (dir. 7), which was originally a translocative.[16]

1.2.3.2 The semantic diversification of the dir.3/dir.6 pair

The pair formed by dir.3 and dir.6 preverbs (𘕕 *·wjɨ²* 'outwards' *vs.* 𘓺 *kjɨ¹* 'inwards' in Tangut) is well attested in languages of the macro-group constituted by languages of the West Gyalrongic and Ersu taxons. However, the semantics attached to these preverbs display variation, and only Ganluo Ersu has a pair of preverbs, which establish a one-to-one semantic correspondence with Tangut. The explanation for such diversity could be provided by that language, which shows two sets of preverbs (*kʰɛ-* 'inwards' *vs.* *ŋɛ-* 'outwards' and *kʰwa˞-* 'North, upstream' *vs.* *ŋwa˞-* 'South, downstream') phonetically close enough to have merged into a unique category. Interestingly, all the meanings resulting from the union of these two preverbs' respective meanings can be found in related West Gyalrongic languages. This fact allows one to map in Figure 4 an illustration of that semantic shift path.

1.2.4 Dir.7: Tangut 𘜶 *dja²*- as a perfective marker

The loss of dir.4 & dir.5 is not the only orientational preverb feature shared by Horpa and Tangut: the undefined orientational prefixes Geshiza *dæ-*, g.Yurong *də-*, and Stau *tə-* can also reasonably be assessed to be cognate to 𘜶 *dja²-* which, as said before, does not clearly encode direction. Indeed, as I will demonstrate now, there are some problems with an orientational interpretation of the preverb 𘜶 *dja²-*.

1.2.4.1 The semantic inadequacy of the orientational/modal interpretations

[16]The existence of a least one cognate in Ersu makes an innovation proper to Khroskyabs and Gyalrong unlikely. Note the regularity of the correspondences if one compares dir.7 (reanalyzed into dir.6 in Japhug and Tshobdun) with dir.2 'down': Tangut *nja¹-* / *dja²-*; Geshiza *næ-* / *dæ-*; g.Yurong *nə-* / *də-*; Tshobdun *nɐ-* / *tʰɐ-*; Ganluo *nɛ-* / *tʰɛ-*; Munya *no-* / *tʰo-*.

{inwards} ≠ {outwards} ∧ {N, upstrm} ≠ {S, downstrm}

↓

proto-West Gyalrongic

Dir.3 **Dir.6**

{inwards, N, upstrm} ≠ **{outwards, S, downstrm}**

↙ ↓ ↘

Tangut	*Geshiza (Khr.)*	*g.Yurong*
{inwards} ≠ {outwards}	{upstrm} ≠ {downstrm}	{N} ≠ {S}

Figure 4: Emergence (by reunion and redistribution) of dir.3 and dir.6 preverbs in Tangut, Geshiza Horpa and g.Yurong Horpa

The first problem with an orientational interpretation lies in the distribution of the verbs occurring together with *dja²*-. In Table 4, the multiplicity of semantic features attached to the verbs cooccurring with *dja²*-, which do not consistently encode a distancing from the agent, tends to seriously invalidate the orientational interpretation. The second group's verbs are more indicative of mood, similar to the Mandarin Chinese resultative 掉 *diào*, which indicates loss and disapearance (a view also formerly expressed in Kepping 1985 and retained in Lai et al. 2020). Nevertheless, that interpretation overlooks the other verbs Kepping had in mind when proposing an orientational interpretation, plus some others which only indicate a change of state (*·we²* 'to become', *ŋwər²* 'to heal, recover').

The label 'perfective' is, in fact, the best choice for *dja²*- because it covers the semantic range of this preverb (i.e., a true change of state, including disappearance and distancing from the agent, depending on the verb).

1.2.4.2 The systemic inadequacy of the orientational/modal interpretations

The orientational interpretation is also invalidated by facts from the system formed by *dja²*- and its Type-B counterpart *djij²*-. If Type-B orientational preverbs usually appear in uses similar to their Type-A counterpart, equilibrium breaks with *dja²*- and *djij²*-.[17] Table 5 shows that of the 24 verbs associated with *djij²*- in Leilin, most also occur with *dja²*-, but others are associated with another Type-A orientational preverb. *djij²*- is the only Type-B preverb analyzable as a counterpart of such a range of different Type-A preverbs, some of them with completely opposing meanings.

As a matter of fact, this contrast is already observable with the preverb *dja²*-, which is replaced by a preverb with clearly orientational semantics when this is required by the context. In Table 6, adapted from Table 4, I put all the verbs

[17]The only exception to that equilibrium being the verb *ljwį²* 'sink', which correlates in my corpora with the Type-A centripetal and with the Type-B preverb marking vertical downward direction.

Verb	Meaning	
·jɨr^{1}	'to ask'	
ta^{1}	'to flee'	
pʰji^{2}	'to throw'	
kʰjow^{1}	'to give'	orientational meaning?
rjɨr^{2}	'to leave, to go'	
·wjij^{2}	'to send, unleash'	
·wjạ^{2}	'to send, unleash'	
tjị^{1}	'to put'	
sja^{1}	'to kill'	
kʰjwɨ^{1}	'to cut'	modal meaning?
/ *sji*^{2}	'to die$_{2}$'	
ljwij^{1}	'to die' (honorific)	
ljiij^{2}	'to destroy'	
·jijr^{2}	'to execute'	
bej^{1}	'to be defeated'	
zar^{2}	'be ashamed'	
lia^{2}	'to be drunk'	
ljij^{1}	'to change' (tr.)	
lej^{2}	'to change' (intr.)	aspectual meaning?
ŋewr^{1}	'be disordered'	
·we^{2}	'to become'	
dźjwa^{1}	'to finish' (intr.)	
ŋwər^{2}	'to heal, recover'	
·o^{1}	EXV.on	
tśʰju^{1}	'to have'	
njiij^{1}*.ljịj*^{2}	'to feel happy'	

Table 4: 27 most frequent verbs occurring with *dja*^{2}- in Leilin.

[illegible] *ljij¹*	to change	[illegible] *śju¹*	to be damaged, destroyed (*dja²*)
[illegible] *rjir¹*, [illegible] *rjor²*	get (*·a*)	[illegible] *sji¹*	die₁ (*dja²*)
[illegible] *rejr²*	be many	[illegible] *dạ²*	know (*dja²*)
[illegible] *·ju²*	to find (*nja¹*, *rjir²*)	[illegible] *śjij¹*	to become (*dja²*)
[illegible] *sej¹*	be calm, clean	[illegible] *dźjiij¹*	to live, be somewhere (*dja²*)
[illegible] *·wẹ¹*	be mediocre	[illegible] *·wəə¹*	to defeat (*dja²*)
[illegible] *ljụ²*	to put, install (*nja¹*)	[illegible] *·we²*	be defeated (*dja²*)
[illegible] *sjwij¹*	be evident	[illegible] *bej¹*	be vanquished (*dja²*)
[illegible] *sjiw¹*	be new	[illegible] *·o¹*	EXV.on (*·a*, *kji¹*, *dja²*)
[illegible] *njwi²*	be able, capable (*·a*)	[illegible] *tjị¹*	to put, place on (*nja¹*, *dja²*, *rjir²*)
[illegible] *·ji²*	to say (*kji¹*, *rjir²*)	[illegible] *sja¹*	to kill (*dja²*)
[illegible] *ŋwuu¹*	to declare	[illegible] *lhjwi¹*	to take, catch (*dji²*)

Table 5: Verbs associated with [illegible] *djij²*- in Leilin (with their Type-A counterpart in parentheses)

occurring only with [illegible] *dja²*- or its optative counterpart [illegible] *djij²*- at the top, and all the verbs occurring not only with [illegible] *dja²*-, but also with other orientational prefixes at the bottom. Sometimes mismatches are only present with Type-B preverbs.[18] As this situation only occurs in legal texts, I indicate the occurrences of these unexpected Type-B preverbs without taking them into account for now. Finally, some texts only present one occurrence of the verb, with a preverb not being [illegible] *dja²*-, an occurrence which is boldfaced. The verbs below the dashed line are verbs for which an orientational analysis could work but is not attested with genuine orientational preverbs.

Most of the top category's verbs do not collocate with the semantics of direction, in contrast to verbs at the bottom, whose semantics are compatible with the semantics of direction and orientation. Those directions are most of the time logical: [illegible] *tjị¹* 'to put' can often (four occurrences) be seen together with [illegible] *nja¹*, 'downwards', as in example (4):

(4) [illegible]

ŋwər¹ dzjwi¹ sja¹ kiẹj² zeew²lʰjij² mji¹-kjir²=bju¹ ljụ²=ɣa²
heaven emperor kill want bear NEG.MOD-manage=because body=LOC
sji¹ljwij¹ śjow¹śju¹ nji² nja¹-tjị¹=·wji¹=pʰji¹
cangue fetters etc. PFV:DOWN-put=LV:do[A]=CAUS

黃帝不忍誅之，乃枷械其身 (Shi et al. 1993: 293)[19]

'The celestial emperor could not bring himself to kill them, and so put cangue and fetters on them.' (Leilin 06.28A.2)

[18] I indicate occurrences with a slash '/' and leave the cell blank when the verb does not appear in the document or when I did not find prefixed occurrences.

[19] The scribe made a mistake in his translation, confusing 黄帝 (the yellow emperor), in Tangut [illegible] *nər²dzjwi¹*, with 皇帝 (the emperor), in Tangut [illegible] *ŋwər¹dzjwi¹*.

verb	meaning	Leilin	12K	Cxz	Avtmsk	Code $^{(1261–1376)}$
sja^1	‘to kill’	/	/	/		/ (kjij^1, 3)
k^h*jwɨ*^{1}	‘to cut’	/	/			(.wjij^2, 5)
/ *sji*^{2}	‘to die$_2$’	/	/	/		/
ljwij^1	‘to die’ (honorific)	/	/			
ljiij^2	‘to destroy’	/	/			
bej^1	‘to be defeated’	/	/			
lia^2	‘to be drunk’	/	/	/		
zar^2	‘be ashamed’	/	/			
njiij^1*.ljij*^{2}	‘to feel happy’	/	/			
ŋewr^1	‘to be altered, troubled’	/	/			
dzjwa^1	‘to finish’ (intr.)	/	/			
ljij^1	‘to change’ (tr.)	/	/			/
lej^2	‘to change’ (intr.)	/	/			
·we^2	‘to become’	/	/	/	/	/
ŋwər^2	‘to heal’	/	/	/		
·o^1	EXV.on	/				/ (kjij^1, 3)
·jɨr^1	‘to ask’	/	/			/
ta^1	‘to flee’	/	/	/		
p^h*ji*^{2}	‘to throw’	/	/	.wjɨ^2 (1)		/
k^h*jow*^{1}	‘to give’	/	/	rjɨr^2 (1)		/
rjɨr^2	‘to leave, to go’	.wjɨ^2 (3), kjɨ^1 (1)	/	/	/	
·wjij^2	‘to send, unleash’	/	rjɨr^2 (4)		**rjɨr^2 (1)**	
·wja̱^{2}	‘to send, unleash’	rjɨr^2 (1)	/	/		
tji̱^{1}	‘to put’	nja^1 (4), rjɨr^2 (2)	rjɨr^2 (1)	**rjɨr^2 (1)**		/ (kjij^1 4, rjijr^2 1)
tś^h*ju*^{1}	‘to have’	nja^1 (1)	**nja^1 (1)**			
! *·jijr*^{2}	‘to execute’	.wjɨ^2 (1)	**·wjɨ^2 (1)**			

Table 6: Compatibility with other preverbs of Leilin’s 27 most frequent verbs occurring with *dja*^{2}-

As for *tś*^{h}*ju*^{1} ‘to have’ which does not imply any direction *per se*, in both cases, the prefixed verb combines with a locative particle (subessive, ‘under’) explicitly indicating the notion ‘down’ (5).

(5)

tś^h*jɨ*^{1} *zjọ*^{2} *śjiw*^{2}*kjow*^{1} (*zar*^{1}*ŋwuu*^{1}*bə*^{2}) *djɨ*^{2}*-lju*^{2} *gju*^{2}
DEM time[B] gecko (a_bug_in_Chinese) PFV-catch recipient
kwow^1*=k*^{h}*ju*^{1} *nja*^{1}*-tś*^{h}*ju*^{1}
upside.down=SUBE PFV:DOWN-have

時取守宮，出覆盆下 (Shi et al. 1993: 289)

‘At that time, he caught a gecko (a bug in Chinese) and put it below a recipient placed upside down.’ (Leilin 06.12B.2)[20]

[20]In this example, the two characters are, in the document, smaller than the others. The constituent is an aside, hence the parentheses used in the gloss. *śjiw*^{2}*kjow*^{1} is a loanword from Middle Chinese 守宮 *syuwX kjuwng* (in the transcriptional system of Baxter 1992). The story comes from the 漢書，東方朔傳 *Hàn shū, Dōngfāng Shuò zhuàn*. Even if the meaning ‘to put’ is not clear in Tangut, it is in the transmitted text: 置守宮盂下.

In the case of verbs appearing with the preverb *·wji^{2}-*, e.g., (6) and (7), the context also shows that the motion is directed outwards, sometimes in a metaphorical way. In the two occurrences of *jijr^{2}* 'to execute' with *·wji^{2}-*, the subject appears with the ergative. The use of the ergative seems to be rhetorical; we can understand it as resulting from the fact that the action performed by the verb's subject exceeds the scope of what that subject should do or should not do.

(6)

sjij^{1} rjur^{1}=k^{h}a^{1} l^{h}ji^{2}kja̱^{2}=·wji^{1}-djij^{2}=tja^{1} bju^{1} t^{h}jij^{1}
now world=INTERE mourning_song=LV:do[A]-PROG=TOP reason Tian
xiwəj^{1}=ɣa^{2} ·wji^{2}-rjir^{2}
heng=LOC PFV:OUT-go_out

今之輓歌起此 (Shi et al. 1993: 314)

'The reason by which today in the world one is singing that mourning song comes from Tian Heng.' (Leilin 09.08.A.2)

(7)

kow^{1}-tsə^{1} xjwi^{1} dźjɨ·wji^{1} ·jĩ^{1} kow^{1}=jij^{1} ·wji^{2}-jijr^{2}
Gongzi Hui ERG Yin Gong=ANTIERG PFV:OUT-execute

'Принц Хуй убил Инь-гуна.' (Solonin 1995: 47)

'The prince Hui executed Yin Gong.' (12K, 132.58.07)

Finally, there are cases of verbs taking the preverb *rjɨr^{2}-*. These cases are more challenging to interpret, as *rjɨr^{2}-* is known in the literature for not encoding any particular direction. Beaudouin (2018) formerly noted that the prefix distribution in *Leilin* shows that the verbs associated with *rjɨr^{2}-* are always dynamic and qualified it as marking 'unspecified direction'. This analysis is accurate here, as (8) and (9) show. As the subject of the verb seems to be always at a distant position from the place of the verb's action in these sentences, one also could conjecture that the preverb *rjɨr^{2}-*, when appearing together with a verb usually collocating with *dja^{2}-*, encodes the notion of being distant from the subject/agent of the verb.[21] This conjecture, however, needs to be tested on a larger number of sentences.

(8)

njij^{2} ŋwə^{1} lụ^{1}gur^{1} rjɨr^{2}-·wji^{1} kiew^{2}=rjar^{1} twụ^{1} rjɨr^{2}-tjị^{1} lụ^{1}gur^{1}
King five rock_ox PFV:DIR-do[A] march=range place PFV:DIR-put rock_ox
mjiij^{1}=k^{h}ju^{1} kię^{1} rjɨr^{2}-tjị^{1}=·wji^{1}
tail=SUBE gold PFV:DIR-put=LV:do[A]

[21] That characteristic is present in (8). In (9), Bu Shi has left his farm after giving it to his brother to raise goats in the mountain. He is then far from his brother at the time of the transfer.

秦王作五石牛，置於界首， 遺金於石牛後 (Shi et al. 1993: 316)

'The king of Qin made five oxen with rocks and put them in a place within the marches; then he put gold behind the tails of the oxen.' (Leilin 09.13.B.3)

(9) [illegible]

[tjo̦² śji¹ wa² rjir¹ zji² lʰjo¹ sjwi¹lu̦² we²]=tśʰjaa¹ jij¹ war²
brother before INTRG get[B] all lose poor become=SUPE REFL good
ji̦² djii¹ ŋewr²-dźjow¹ rjir²-kʰjow¹
again separate lot-times PFV:DIR-give[A]

弟尽破其产，式辄复分与弟者数矣 (Hàn shū 56)

'Son frère avait quant à lui perdu tout ce qu'il avait obtenu et était devenu pauvre. (Bu Shi) lui donna des biens à plusieurs reprises.' (Jacques 2007: 52-53)

'When his brother lost everything he had earned before and became poor, he (Bu Shi) gave parts of his goods away to him several times.' (CXZ 4-17)

1.2.4.3 Comparative evidence for the perfective analysis

In Geshiza Horpa, the only function of the preverb *dæ-* is to encode perfective aspect without any orientational implication. This preverb corresponds to Tangut [illegible] *dja²-* in many aspects:

1. First, phonologically speaking, the correspondence G. *-æ* :: T. *-a*, already found for the plain vowel (INTRG, 'one' : G. *æ* :: T. *a*) is attested in other cognates with coronal initials, as seen in Table 7.[22]

meaning	'road'	PFV	'to kill'	'to release'	'to write'	PFV:DOWN
Tangut	tśʲa¹	dʲa²-	sʲa¹	lʲa²	rʲar¹	nʲa¹-
Geshiza	tɕæ	dæ-	sʰæ	læ	ræ	næ-

Table 7: Examples of Geshiza *-æ* :: Tangut *-a* correspondences with coronal onsets

2. Second, as for the verbs collocating with Tangut [illegible] *dja²-*, Geshiza Horpa *dæ-* can collocate with verbs compatible with other preverbs when those verbs have an orientational meaning. Compare (10) and (11).

(10) *gadə gadəɣi braŋgu dæ-ɕhoŋ*
morning early.morning TOPON PFV-go.PST.1
'I went (downriver) to Danba County Town early in the morning.' (Geshiza Horpa, Honkasalo 2019: 546)

[22] For 'one', the particular usages also coincide (see Section 2.1).

(11) *gadə* *gadəɣi* *braŋgu* *wə-ɕhoŋ*
morning early.morning TOPON PFV.DIR-go.PST.1
'I went (downriver) to Danba County Town early morning.' (Geshiza Horpa, Honkasalo 2019: 546)

3. Third, the verbs occurring together with the preverb are similar in both languages, and they have the same distribution on the whole. Table 8 lists the 27 Geshiza verbs most frequently seen with the perfective preverb *dæ-* in Honkasalo (2019). Apart from the presence in the last row of verbs denoting semantic assimilation to the agent ('to eat', 'to drink') or pure action ('to do', 'to build'), the three categories depicted in Table 4 above can be found. The first group is potentially orientational in nature, the second group is related to mood ('loss'), and the third group expresses a simple change in state due to the preverb's perfective nature. As for Tangut, only a generic perfective analysis can encompass such a wide range of usages.

   Ten verbs in my comparative data take the generic perfective in both Tangut and Geshiza. A cognate not preceded by [illegible] *dja*^{2}- (Tangut verbs of the second column) does not imply that the association is impossible. It is possible that we simply do not have enough texts to ascertain it for the time being. Finally, some verbs for which I did not find any cognate in Tangut still have perfect synonyms in Table 4 (e.g., 'to finish', 'to become').

   Beaudouin (2018) has also formerly proposed that [illegible] *dja*^{2}- is a mood marker. However, it is not surprising that any irreversible change (the perfective aspect), associated with specific verbs already denoting loss, could be seen as a loss. Indeed, due to the semantics of the verb, should one use that former explanation to treat examples in (12) from Geshiza (Honkasalo 2019: 545), *dae-* could also be seen as a prefix denoting mood, which is not the case:

(12) a. *rjəu=ke=nɔ* *dæ-bædzo-s*^{h}*i*
wife=DAT=TOP.C PFV-divorce.3-IFR
'He got divorced from his wife.'

b. *<taʈʂɛ>-væ=dʑe* *æ-ɣi* *dæ-sæ-s*^{h}*i*
TOPON-NAT=TOP one-CLF.person PFV-die.PST.3-IFR
'A person from Dazhai village had died.'

Among the orientational prefixes of Ganluo Ersu, apart from the similarity in meaning for the inwards/outwards pair (T. [illegible] *kjɨ*^{1}- 'inwards' :: Ganluo Ersu *khɛ-* 'inwards'; T. [illegible] *·wjɨ*^{2} 'outwards' :: Ganluo Ersu *ŋɛ-* 'outwards'), one can observe the existence of a prefix specialized in the encoding of perfective aspect *t*^{h}*ɛ-*, which collocates with verbs similar to those just de-

Geshiza verb	Meaning	Tangut cognate found with dja^2-	Tangut cognate not found with dja^2-	
k^{h}o	'to give'	*k^{h}jow^1*		
ɕə	'to go'	*śjɨ^1*		
læ	'to release, to send'		*lja^2*	direction?
jə	'to say'		*·jɨ^2*	
n-ʑæ	'to give birth'		*l^{h}jɨ^2*	
sæ	'to kill'	*sja^1*		
k^{h}uæ	'to cut'	*k^{h}jwɨ^1*		
sæ	'to die'	*sjɨ^1*		'loss'?
lmə	'to forget'	*mjɨ^2*		
ra	'to hit'		*rjijr^2*	
tæpæ	'to take out'			
st^{h}æ	'to finish'			
tje	'to become'			
rji	'to wake up'			
lxua	'to appear, go back'		*l^{h}jwo^1*	
ʑæ	'to come'			
ntɕ^{h}o	'to have'	*tś^{h}ju^1*		PFV
ndzo	'to stay'	*dźjiij^1*		
dʑi	EXV	*dźjij^2*		
wi	EXV (attached)	*·wjij^2*		
ŋuə	COP		*ŋwu^2*	
ma	NEG.EXV		*mjij^1*	
ŋgə	'to eat'			
t^{h}i	'to drink'		*t^{h}ji^1*	
və	'to do'		*·wji^1*	?
dæ	'to do'			
t^{h}o	'to build'			

Table 8: 27 most frequent Geshiza verbs with the preverb *dæ-* in Honkasalo (2019), with their Tangut cognates

scribed above (13, where *-ǎ* indicates a new situation – NSIT).[23] This cognate preverb (see Table 2) enforces the view expressed above. However, note that the translocative interpretation is not wrong from a diachronic point of view, as the cognate preverb *tʰo-*, in Munya, is a genuine translocative. The former analysis of Tangut 𘜶 *dja²-* as a translocative is then consistent with the fact that the general perfective found in Tangut, Horpa languages, and Ganluo Ersu probably *originated* from a translocative.

(13) *tʰɛ́ zámá dzə̣̀ tʰɛ́-dzə̣̀=ǎ*
3SG food eat PFV-finish=NSIT
'He finished eating the food.' (Ganluo Ersu, Chirkova & Wang 2017)

Orientational preverbs in Geshiza are more similar to Tangut orientational preverbs than any other orientational preverb system in documented languages of the area. Geshiza's system displays the same phenomenon of the addition of orientation to the general perfective aspect (i.e., orientational meaning is available only with certain verbs). Geshiza also has a preverb *dæ-* cognate with 𘜶 *dja²-*, whose behavior allows one to attribute to that same 𘜶 *dja²-* a coherent semantic value. The next subsection will show that the Tangut preverb system is derived in the same manner: optative from imperative and interrogative from perfective. A comprehensive analysis of Tangut 𗅁 *kjɨ¹-*, whose counterpart in Geshiza is the only preverb being used in an imperfective way, could potentially provide new insight into the Tangut orientational preverbs system.

1.2.5 Interrogative preverb

Tangut shares with other West Gyalrongic languages a cognate interrogative prefix (14). As in g.Yurong Horpa and Khroskyabs languages, this preverb is identical to that which encodes the direction 'upwards'. For the moment, the split observed in West Gyalrongic between Geshiza and Mazur Stau and the other West Gyalrongic languages for the direction 'upwards' is still difficult to account for from the point of view of Tangut. I believe *rə-* to be cognate with the unspecified orientational prefix 𗧓 *rjɨr²-* found in Tangut (Subsection 1.2.2). We should nevertheless be able to reconstruct a common interrogative preverb for proto-West Gyalrongic.

[23] Chirkova & Wang (2017: 6-7) write: 前綴 9（tʰɛ-）有專有構成完整體的功能，詞彙意義虛化，主要與終止型動詞（telic verbs）（瞬成動詞和達成動詞）相搭配，如 *tʰɛ̀-ʃó* '死'、*tʰɛ̀-bzɿ̣́* '敗'、*tʰɛ̀-tó* '折斷、斷掉'、*tʰɛ̀-mɛ́* '忘記'、*tʰà-tʃhá* '康復'、*tʰɛ̀-lí* '融化'）。*"The ninth prefix (tʰɛ-) has a function specifically encoding the perfective aspect. Its lexical meaning is null and it is mostly associated with telic verbs such as tʰɛ̀-ʃó 'to die', tʰɛ̀-bzɿ̣́ 'to be defeated', tʰɛ̀-tó 'to break', tʰɛ̀-mɛ́ 'to forget', tʰà-tʃhá 'to recover', tʰɛ̀-lí 'to melt')."* As a matter of fact, the verbs found with Tangut 𘜶 *dja²-*, and Geshiza *dæ-* are largely identical. Thus, we could see the perfective preverbs of those three languages as cognates going back to a proto-subgroup inside macro-Qiangic, including Ersu but excluding Munya. The original preverb was either already grammaticalized or displayed semantic features that conditioned the same grammaticalization pathway (which departed from a genuine translocative, attested by Munya). Another possibility would be to see in the grammaticalization an areal innovation shared by Tangut, Horpa, and Ersu at a stage of early contact.

orientation *TAM*	Geshiza	Tangut
meaning	IND, IMP / NACT / INTRG, OPT	IND, IMP / INTRG, OPT
undefined *perfective (secondary)*	/	rjɨr^{2} / rjijr^{2}
upwards *perfective (secondary)*	rə- / rə- / ri-	.a / .jij^{1}
interrogative	æ-	.a
downwards *perfective (secondary)*	næ- / nə- / ni-	nja^{1} / njij^{2}
outwards (T), dwnstr. (G) *perfective (secondary)*	wə- / wə- / wi-	·wji^{2} / .wjij^{2}
inwards (T), upstr. (G) *imperfective*	gæ- / gə / gi-	kji^{1} / kjij^{1} ?
perfective (primary)	dæ- / də- / di-	dja^{2} / djij^{2}

Table 9: Type-A/B orientational preverbs in Geshiza and Tangut

(14)

mjo^{2}=tja^{1} tśjow^{1} ·jɨ^{2}te^{1} ŋwu^{2} dzji^{1}dzeej^{1} lja^{1} kjir^{2}=mjijr^{2} ·a^{0}-dju^{1}
1HUM=TOP Zhang Yide COP compete come dare=NMLZ INTRG-EXV

「吾是張翼德，敢來決敵也！」(Shi et al. 1993: 316-317)

"I am Zhang Yide, will there be people who dare to compete with me?" (Leilin 09.15.A.6)

1.2.6 'Optative' preverbs

The most striking similarity between Tangut and Geshiza (and Stau also) regarding mood lies in a common alternation observed between the two series of orientational prefixes (see Table 9). In Geshiza and Stau, the series with *-i* is the result of the fusion of the Type-A prefix with an irrealis morpheme *-i-*.[24] That morpheme produces an interrogative mood when the prefix it attaches to has an underlying indicative mood (usually with aspectual implications), and an optative mood when the prefix it attaches to marks the imperative (Honkasalo 2019: 535). There is also a third series, labeled 'non-actual' (NACT), not reflected in Tangut.

This description should therefore allow us to consider the possibility that in Tangut the optative would also be a derivation of the imperative, and that there are interrogative occurrences derived from the indicative perfective, a possibility tested successfully below.

First of all, the optative prefix in Tangut does not lose its orientational meaning, as illustrated by (15). Here, the officiant has to remove the animal 'out of' the

[24] This similarity is identified by Lai et al. (2020), though without testing its semantic, aspectual, and modal implications for Tangut.

place it used to be.

(15)

kow^2 [l^{h}jwo^{1}dźiow^2 lja^1]=tś^{h}ja^1 dja^2-sji^2 sju^2
officiant return.official_document send_back=SUPE PFV-die$_2$ animal
wjij^2-k^{h}jwɨ^{1}tjij^1
OPT:OUT-remove

'The officiant, when sending back the official document, must remove the dead animal (from the record).' (Code, article 1362, Sacrificial Animals)

Then, arguing that optative preverbs are derived from the imperative requires finding examples of imperative sentences with the first series's preverb. Such uses exist, as (16) shows:

(16)

·jị^2 tsjiir^{1}lu^2 dja^2-k^{h}jow^1=·wjo^1 ·jir^1 dzji^1 śji^1=su^1
again rank IMP-giving=LV:do[B] emolument eat before=COMP
·a-bjij^1=·wjo^1
IMP:UP-raise=LV:do[B]

"Верните [Мэн-чан-цзюню] его прежние ранг и должность, а жалованье и пропитание дайте больше, чем прежде." (Solonin 1995: 40)

"Give him back his rank and raise his emoluments higher than those he had before!" (12K, 132.26.02)

If the system is similar, and as there are two moods possible for the first series in Geshiza, it also requires that one should be able to find examples of Type-B preverbs bearing interrogative meaning in Tangut. Such an example can precisely be found (17). Note that this example can leave no doubt about the interrogative nature of *djij^2*-, as the as the verb *·jir^1* 'to ask' introduces the question, the answer being introduced by *kụ^{2}dạ^2* 'respond'. Note also that the correlate Type-A prefix is used in the answer, in perfect parallel with the question.

(17)

nioow^1 ·jir^1 ·we^2 sə^{1}t^{h}u^1 djij^2-tś^{h}jɨ^1-l^{h}ew^2
after ask Wei situ INTRG-POT-liberate

kụ^{2}dạ^2 dja^2-l^{h}ew^2
answer PFV-liberate

Not translated by Solonin.

"Did you manage to free the Situ of Wei?"
'He answered: "I did."'(12K, 132.01.07)

As for the perfective above, the system found in Geshiza perfectly parallels that of Tangut; in fact, it gives the best explanation to date regarding the distribution of Type-A/Type-B preverbs one can observe in Tangut.

1.2.7 Negative preverbs

Table 10 gives an overview of negative preverbs in Tangut, Geshiza, Mazur Stau, and Wobzi. Other languages of the Qiangic family (including Gyalrong) have potential cognates, but the usages diverge significantly.[25] The comparison here focuses then on the West Gyalrongic subgroup.

	Neg.1	Neg.2	Neg.3	Neg.4
Tangut	mji^{1}-	mjij^{2}-	mji^{1}-	tji^{1}-
Geshiza	mi-	mɛ-	mə-	di-
Stau	mi-	mæ-	mə-	ti-
Wobzi	mə-			tə-

Table 10: Negative preverbs in Tangut, Geshiza, Stau, and Wobzi

All four languages have a default negative preverb, which is in the Table placed in the column 'Neg. 1'; and in all three languages, Neg. 4 derives prohibitive or negative jussive constructions.[26] Nevertheless, only Tangut and Horpa languages have a negative morpheme (Neg.3) preceding a subset of modal verbs similar to those found in Tangut.[27] For example, in Tangut and Geshiza, a cognate verb Geshiza *dʑo* 'to bear, be able to put up with' :: Tangut *dźioow²* 'to bear, be suited to' only occurs with modal negation, as illustrated in (18) and (19).

(18) [illegible]

tsji^{1} lụ^{1}ljị^{2} *ŋwu^{2}* *ljị^{1}* *mji^{1}-dźioow^{2}*
just rocky_land COP cultivate NEG.MOD-bear

[25]For example, Japhug has four negative preverbs *mɤ-*, *mɯ-*, *ma-* and *mɯ́j* which are manifestly related to Tangut's first three forms. Nevertheless, the distribution is very different to what one can see in West Gyalrongic: *mɯ́j-* is a sensory evidential negation; *ma-* occurs with prohibitive verb forms; *mɤ-* occurs on non-finite verbal forms without orientation preverb, in factual non-past, irrealis and when preceded by interrogative and proximative morphemes; *mɯ-* is seen elsewhere. In Munya, the system is on the whole very different, as four prefixes *tɕɯ-*, *tɕɛ-*, *nyɯ-*, and *mo-* form it.

[26]Munya also has a cognate *tɕɯ-*. This prohibitive can be traced back as far as proto-Trans-Himalayan (Matisoff 2003: 586).

[27]In Mazur Stau, *mə-* does not only appear with modals, though mostly with modals and Class 1 verbs (which tend to be stative). Gates (2021) calls it a negative imperfective. A dedicated modal negation marker is also documented for the Phoxiu variety of Stau, Central Horpa (Sami Honkasalo, personal communication).

猶石田也，不可種之 (Shi et al. 1993: 261)

'It is just a wasteland, it is not fit for cultivation.' (Leilin 03.21A.6)

(19) *ŋa mɲa-me dæ-dəu tɕʰu mdʑurtenme=nɔ stɕʰəkʰi-ʑæ*
1SG NEG.EXV-NMLS:P PFV-do.1 CONJ common.people=TOP.C watch-NMLS:P
mə-dʑoŋ-ræ
NEG.MOD-bear.1-SENS

'I did something wrong, so I cannot bear facing people.' (Honkasalo 2019: 648)

The functions of Geshiza's *mɛ-* (Neg.2) on the whole also matches Tangut [illegible] *mjij²-*, which is traditionally described as marking the past. In Geshiza, *mɛ-* marks also the past, except with the prospective aspect. This behavior might actually be identical in Tangut, but more investigation is still needed. In (20), [illegible] *mjij²-*, being prefixed to a verb marked with the telic/future [illegible] *.jij¹*, cannot be accurately glossed as 'past negation'. It appears the configuration corresponds here to the definition Comrie (1989: 64) gives of the prospective, i.e., an aspect expressing a "present state relative to a future event."

(20) [illegible]
[illegible]

ɕiə¹ tsʰew² ŋo² tʰjwə¹ kjɨ¹-mjij²-sjɨ¹-.jij¹ zjij¹ gji²=.jij¹
Shi Cao ill get PFV:DIR-NEG-die$_1$-TEL time[⊠] child=ANTIERG
lʰji²wẹ¹ ŋa² dja²-sjɨ¹ zjij¹ ɣur¹ djij¹.rewʳ²=kʰju¹ rjɨr²-tjọ¹
bequeath 1SG PFV-die$_1$ time[⊠] corpse stairs=SUBE PFV:DIR-put[B]

'Ши Цао заболел и, перед смертью, завещал своему сыну : "Когда я умру, положи моё тело на ступени дворца."' (Solonin 1995: 62)

'Shi Cao fell ill and as he still hadn't died [which he was going to], said to his son: "When I will be dead, put my body below the stairs."' (12R, 133.27.04)

Morphologically speaking, the vocalic alternations in the negation preverbs are a local retention of what could be a very ancient distinction. These alternations are seen in other Gyalrongic languages (although with different semantic values) and also in Tibetic, which has a མི་ *mi* / མ་ *ma* opposition (Tournadre & Dorje 2003: 423).

1.2.8 Modal preverbs

The slot which immediately follows negation can be filled in Tangut, Geshiza, and Wobzi with a modal preverb incorporated into the verbal template. These morphemes in Tangut are *tśʰjɨ*^{1} (potential), *ljɨ*^{1} (concessive) and *tsjɨ*^{1} (uncertain function).

1.2.7.1 tśʰjɨ^1 *(potential)*

Two possible candidates compete in Geshiza to establish cognacy with the Tangut 'potential' modal preverb *tśʰjɨ*^{1} (21).

(21)

nioow^1 *pʰo*^{2} *śioow*^{1} *ŋa*^{2}*=·jij*^{1} *so̜*^{1} *tśiej*^{2} *njij*^{2}*=do*^{2}
POSTE Bao Shu 1SG=ANTIERG three time king=TERM
·a-pow^1*=·wji*^{1} . *njij*^{2} *mjɨ*^{1}*-tśʰjɨ*^{1}*-nji*^{2}
PFV:UP-help=LV:do[A] . king NEG.MOD-POT-listen[B]

"Потом Бао Шу три раза помогал мне [встретиться] с ваном, а ван не стал меня слушать." (Solonin 1995: 38)

"After, Bao Shu helped me three times before the king. [But] the king could/would not [still] hear me." (12K, 132.19.06)

First, Geshiza has a bound deontic auxiliary *-tɕʰi* (22) that expresses acceptability ('can, be all right'). Phonologically speaking, the correspondence works, as Tangut /ɨ/ syllables with palatal affricate aspirated onsets are reflected as a high vowel in Geshiza, as in T. *tśʰjɨ*^{1} :: G. *tɕʰe* 'narrow'.[28] However, that morpheme, although bound to the verb as for Tangut *tśʰjɨ*^{1} (potential), is located in a suffixal position.

(22) *lmæ=ntshe smæŋa xo=zɔ gæ-ɕoŋ-tɕhi-ræ=je*
3=ASSOC.GEN girl DEM.LOC=only DIR-go.NPST.1-AUX-SENS=MOD
'(On a trip to Dandong,) you can go (to stay) in their daughter's place.' (Geshiza, Honkasalo 2019: 573)

Another possibility would be to see another auxiliary verb *tɕʰa* 'can' as a potential cognate in Geshiza (23). Like its putative Tangut cognate (the modal verb *tśʰjɨ*^{1}) this verb appears just after the negative slot.

However, there are three potential counterarguments to that second solution: first, *tɕʰa* 'can' in Geshiza is not part of the verbal template and usually acts as an independent verb (notably, it bears conjugation). Second, Tangut syllables with the surface form /tśʰja/ seem to correspond to Geshiza /tɕʰa/, as with the superessive T. *tśʰja* :: G. *tɕʰa* (postposition 'on'). Still, it should be kept in mind that the particular infixal position of *tśʰjɨ*^{1} would make it easy for the syllable to produce a neutralized form.

[28] The correspondence with unaspirated palatal affricate onsets is still unclear: T. *tśjɨ̣*^{1} :: G. *a̰tɕʰi* 'to move', and T. *tśjir*^{1} :: G. *rtɕʰe* 'to tie' (see also Wobzi Khroskyabs *rcʰê*), but T. *tśji̱r*^{2} :: G. *stɕær* 'to be afraid'. However, that last correspondence seems unlikely, as pre-Tangut preinitials *s- are usually reflected as a dot below the vowel in Gong Hwang-cherng's reconstruction.

Only further investigation will reveal the correct etymology, knowing that the two potential cognate verbs may also have a link in Geshiza itself. In fact, Tangut [illegible] *tśʰji¹* could also be the result of reanalysis due to the semantic and syntactical proximities of the two morphemes.

(23) *d-ə-vkə=ke=ræ* *rə-ro* *rə-nɕʰə*
PFV-NACT-get.full.NPST.3=SEQ=LNK DIR-ADV DIR-jump.INF
mə-tɕʰa-mə-ræ
NEG.MOD-AUX.can.NPST.3-EP-SENS
'After getting full, it could not jump up (and go away).' (Geshiza, Honkasalo 2019: 648)

1.2.7.2 [illegible] *ljɨ¹* (concessive)

Lai (2021) reports in Siyuewu Khroskyabs a formerly undocumented prefix *də-* 'even' (24a), related to another homonymous enclitic *=də* in the same language (24b).

(24) a. *kə-mə-də-sŋ-óŋ*
PST-NEG-even-sleep$_2$-1
'I didn't even sleep.'

b. *sŋə̂=də* *kə-mə-sŋ-óŋ*
sleep$_1$=even PST-NEG-sleep$_2$-1
'I didn't even sleep.'

As proposed by the author, the two morphemes could be cognate with the pair encoded by Tangut [illegible] *ljɨ¹*.[29] Both of the roles of [illegible] *ljɨ¹* can indeed be seen in similar configurations; first as a clitic with the meaning 'also' (inclusive focus marker) or 'even' (additive focus marker, 25), second as a preverb immediately attached to the verb with a unique additive focus function (26).[30]

(25) [illegible]

sjij¹ *ŋwuu¹=ljɨ̣¹* *mjij²-to²*
now word=even NEG.PFV=go.out

'Ныне он не сказал ни слова.' (Solonin 1995: 38)
'He did not even say a word.' (12K, 132.18.04)

(26) [illegible]

[29] Lai Yunfan, personal communication.
[30] The enclitic in Siyuewu Khroskyabs can be attached to nouns. Note that there are other examples of correspondence Tangut *l-* :: Khroskyabs *d-* with the numeral 'four', see Subsection 2.1.

tsʰji^{1}mjĩ^1 njij^2 śjij^1 jã^2 lʰjij=dźjɨ·wji^1 lʰjij
Qimin king reign Yan State=ERG State
djɨ^2-ljɨ̣^{1}-lʰjwi^1=·wji^1-djij^2 ·jị^2 lạ^1 kjɨ^1-ljị^1
PFV:AUTOB-CONC-taking=LV:do[A]-PROG again hand PFV:IN-sink.into

‘Во время циского Минь-вана царство Янь захватило царство [Ци], но [потом] оно снова вернулось в руки [Минь-вана].’ (Solonin 1995: 40)

‘During the reign of king Qimin, although Yan was seizing his country, he restored it (to Qimin).’ (12K, 132.24.03)

Lai (2021) favors the hypothesis of a grammaticalization pathway departing from the enclitic to explain the synchronic existence of the two morphemes. If this hypothesis is correct, it could be a clue to the cognacy between Geshiza deontic auxiliary *-tɕʰi* and Tangut potential *tśʰjɨ^1*, as it provides an example of grammatical prefixation with a suffixal origin.[31]

1.3 Suffixal positions

1.3.1 Agreement suffixes

In this subsection, after I give a brief overview of the Tangut person indexation system, well known since the seminal works of Kepping (1975) and Gong (2001), Then I provide a cross-analysis with verbal paradigms in several West Gyalrongic languages. I do not discuss here the dual suffix *-kjɨ^1* first revealed by Nishida (2004), and whose existence has been since confirmed by Arakawa (2018) with very convincing examples.[32] This morpheme, which occurs only in first person contexts, must be related to Siyuewu Khroskyab's first person dual agreement suffix *-ɣ* (Lai 2017: 347-349), which appears in similar configurations. However, more research is required on how and why these suffixes surface to propose a systematic comparison (which will have to include inverse constructions).

1.3.1.1 Agreement rules

In intransitive contexts, the agreement scheme is quite simple: the verb, with just one argument, agrees with that unique argument employing the suffix *-ŋa^2* for the first person singular, *-nja^2* for the second person singular, and *-nji^2* for the plural of both the first and second person. I focus next on the transitive conjugation, which is far more complex.[33]

[31] Geshiza has a scalar concessive conditional marker *=be* ‘too, even’, not attested at a prefixal position, which is not etymologically related.

[32] Neither do I discuss the dual suffix *tsjɨ^1* whose existence has been revealed very recently (Zhang Ms).

[33] Beaudouin (2022) has shown that agreement in Tangut is primarily mandatory, the exception being circumscribed in dependency patterns.

In local scenarios (between first and second person), agreement can be likened to an ergative-absolutive system, with agreement always targeting the patient (P) ([illegible] *pʰji*^{1}-*ŋa*^{2} 'You send me', [illegible] *pʰji*^{1}-*nja*^{2} 'I send you'). However, in mixed contexts (i.e., if the interaction takes place between first or second and third person), agreement occurs with the first or second person, regardless of its syntactic role (i.e., agent or patient), the only difference being the vowel of the root, stem A for 3 → 1/2 configurations, stem B for the 1/2 → 3 ones. In example (27) (mixed context, 2 → 3), the agreement is with the agent, in contrast with example (28) (mixed context, 3 → 1), where the agreement is with the patient.

(27) [illegible]

xwã^{1} *kow*^{1} *·jir*^{1} *dạ*^{2} *dźjɨ* *wji*^{2} *·wa*^{2} *·wjọ*^{2}-*nja*^{2}
Huan Gong ask say perform skill INTRG can[B]-2

'Хуань-гун спросил: "Какое искусство [ты] покажешь?"' (Solonin 1995: 36)

'Huan Gong asked : "What fine art are you good at?"' (12K, 132.12.03)

(28) [illegible]

·ju^{2} *sjiij*^{2} *źjɨ̣*^{1} *tśier*^{1} *dzjwo*^{2}=*dźjɨ·wji*^{1} *mja*^{1}-*sja*^{1}-*ŋa*^{2}
often think left right people=ERG IRR-kill-1

常慮左右圖己 (Shi et al. 1993: 267)

'He was often thinking : "My waiters could kill me."' (Leilin 04.03A.4)

I give in Table 11 the attested paradigm of the Tangut verb, first described by Gong (2001: 32-34). 1 → 3 and 2 → 3 contexts produce the emergence of stem B (Σ^2, marked in red).

<table>
<tr><th>A \ P</th><th>1SG</th><th>1PL</th><th>2 SG</th><th>2PL</th><th>3</th></tr>
<tr><td>1SG</td><td rowspan="2" colspan="2"></td><td rowspan="2">Σ^1-nja^2 [illegible]</td><td rowspan="2">Σ^1-nji^2 [illegible]</td><td>Σ^2-ŋa^2 [illegible]</td></tr>
<tr><td>1PL</td><td>Σ^1-nji^2 [illegible]</td></tr>
<tr><td>2SG</td><td rowspan="3">Σ^1-ŋa^2 [illegible]</td><td rowspan="3">Σ^1-nji^2 [illegible]</td><td rowspan="2" colspan="2"></td><td>Σ^2-nja^2 [illegible]</td></tr>
<tr><td>2PL</td><td>Σ^1-nji^2 [illegible]</td></tr>
<tr><td>3</td><td>Σ^1-nja^2 [illegible]</td><td>Σ^1-nji^2 [illegible]</td><td>Σ^1</td></tr>
</table>

Table 11: Agreement paradigm of the Tangut verb

This alternation, which makes it possible for the speaker to differentiate 3 → 1 and 3 → 2 configurations from these contexts, has been shown by Jacques (2009) to originate from the merger of an ancient third-person patient suffix *-w (whose presence is widespread in the Sino-Tibetan family, see for example DeLancey 2010).

Gong (2017) revealed a remnant showing striking similarities between mixed contexts of Tangut and the stem 3 of a core Gyalrong language, Zbu, which also results from the merger of *-w with the stem. I give an adaptation of his discovery in Table 12.[34]

	3 = P	3 = A	
1SG	*^{n}dzo-ŋʔ*	*və-^{n}dzé-ŋʔ*	Zbu
	dzjo^{1}ŋa^{2}	*dzji^{1}ŋa^{2}*	Tangut
	‘I eat it’	‘It eats me’	
2SG	*tə-^{n}dzoʔ*	*tə-və-^{n}dzéʔ*	Zbu
	dzjo^{1}nja^{2}	*dzji^{1}nja^{2}*	Tangut
	‘You eat it’	‘It eats you’	
3SG	***^{n}dzoʔ***	*və-^{n}dzéʔ*	Zbu
	dzji^{1}	*dzji^{1}*	Tangut
	‘He eats it’	‘It eats him’	
1PL	*^{n}dzé-jə*	*və-^{n}dzé-jə*	Zbu
	dzji^{1}nji^{2}	*dzji^{1}nji^{2}*	Tangut
	‘We eat it’	‘It eats us’	
2PL	*tə-^{n}dzé-ɲə*	*tə-və-^{n}dzé-ɲə*	Zbu
	dzji^{1}nji^{2}	*dzji^{1}nji^{2}*	Tangut
	‘You eat it’	‘It eats you’	

Table 12: Zbu Gyalrong stem 3 and Tangut stem B (‘to eat’)

1.3.1.2 Cross-analysis of Tangut, Khroskyabs and Horpa verbal agreement paradigms

I give in Table 13 the paradigm of Wobzi Khroskyabs, in Figure 5 the paradigms of Geshiza and Dgebshes Stau, and in Figure 6 the paradigms of Gexi and Khang.gsar.[35] In those figures, I mark in red all the forms for which I can establish biunique correspondences with Tangut, and in blue those for which I can establish biunique correspondences but where one of the Tangut forms is missing in the modern language (i.e., when a merger occurred).

The motive for displaying all of these paradigm charts is twofold. First, trivially enough, a comparison using a more comprehensive ensemble of paradigms can lead to better accuracy in the analysis. Second, the internal variation revealed in Horpa forces us to remain cautious when drawing conclusions. Indeed, as one can see, Khang.gsar has its history (the disappearance of all the suffixes when

[34]The preverb *və-* corresponds to the inverse discussed in the next paragraph.

[35]Wobzi Khroskyabs data is from Lai 2017, Geshiza Horpa data from Honkasalo 2019, Dgebshes Stau Horpa data from Gates 2017 and Sun & Tian 2013, and Khang.gsar Stau Horpa data from Jacques et al. 2014.

<table>
<tr><th>A \ P</th><th>1SG 1PL</th><th>2SG 2PL</th><th>3</th></tr>
<tr><td>1SG
1PL</td><td></td><td>Σ-n</td><td>Σ-ŋ
Σ-j</td></tr>
<tr><td>2SG
2PL</td><td rowspan="2">u-Σ-ŋ u-Σ-j</td><td></td><td>Σ-n</td></tr>
<tr><td>3</td><td>u-Σ-n</td><td>u-Σ</td></tr>
</table>

Table 13: Verbal paradigm of Wobzi Khroskyabs

the P is a second person, for example), but the distance from Tangut gradually diminishes when one compares Tangut with Geshiza and Dgebshes.

<table>
<tr><th>A \ P</th><th>1SG 1PL</th><th>2SG 2PL</th><th>3</th></tr>
<tr><td>1SG
1PL</td><td></td><td>Σ-n</td><td>Σ-w
Σ-ŋ</td></tr>
<tr><td>2SG
2PL</td><td rowspan="2">v-Σ-ŋ</td><td></td><td>Σ-i
Σ-n</td></tr>
<tr><td>3</td><td>v-Σ-n</td><td>v-Σ</td></tr>
</table>

<table>
<tr><th>A \ P</th><th>1SG 1PL</th><th>2SG 2PL</th><th>3</th></tr>
<tr><td>1SG
1PL</td><td></td><td>Σ-n</td><td>Σ-u
Σ-ã</td></tr>
<tr><td>2SG
2PL</td><td rowspan="2">v-Σ-ã</td><td></td><td>Σ-i
Σ-n</td></tr>
<tr><td>3</td><td>v-Σ-n</td><td>v-Σ</td></tr>
</table>

Figure 5: Paradigms of Geshiza (left) and Dgebshes as in Gates 2017 (right)

<table>
<tr><th>A \ P</th><th>1SG 1PL</th><th>2SG 2PL</th><th>3</th></tr>
<tr><td>1SG
1PL</td><td></td><td>Σ-n</td><td>Σ-u
Σ-ŋ</td></tr>
<tr><td>2SG
2PL</td><td rowspan="2">v-Σ-ŋ</td><td></td><td>Σ-i
Σ-n</td></tr>
<tr><td>3</td><td>v-Σ-n</td><td>v-Σ</td></tr>
</table>

<table>
<tr><th>A \ P</th><th>1SG 1PL</th><th>2SG 2PL</th><th>3</th></tr>
<tr><td>1SG
1PL</td><td></td><td>Σ</td><td>Σ-w
Σ-ã</td></tr>
<tr><td>2SG
2PL</td><td rowspan="2">v-Σ-ã</td><td></td><td>Σ-j</td></tr>
<tr><td>3</td><td colspan="2">v-Σ</td></tr>
</table>

Figure 6: Paradigms of Dgebshes as in Sun & Tian 2013 (left) and Khang.gsar Horpa (right)

Of all modern languages, the person indexation system of Tangut (Figure 11) is for once most similar to that of Wobzi Khroskyabs (Table 13). Apart from the differences due to the absence of an attested inverse in the Tangut verbal paradigm and a merger between second person singular and plural forms in Wobzi, the paradigms are nearly identical. However, some features of the Tangut agreement system exist only in Horpa languages. As stated above, Stem B results from the fusion with an ancient third-person patient suffix *-w, an affix we can still find in a lot of varieties of Horpa (*-u* and *-w* being different notations of the same suffix).

If one considers the reflexes observed in all the West Gyalrongic languages listed above, this cross-comparison enables the reconstruction given in Table 14 for proto-West Gyalrongic.

This reconstruction explains the occurrence of stem alternation in Tangut (Σ^2 < *Σ-w), plural markers in Tangut (*-nji²* < *-jna ∨ *-jŋa), and the nasal reflexes of plural forms where the P is third-person in Horpa (< *-jna ∨ < *-jŋa). Except for Tangut, where a merger occurred in all plural forms, *-jna has always simplified

A \ P	1SG	1PL	2 SG	2PL	3
1SG 1PL			*Σ-na	*Σ-j-na	*Σ-w-ŋa *Σ-j-ŋa
2SG 2PL	*w-Σ-ŋa	*w-Σ-j-ŋa			*Σ-w-na *Σ-j-na
3			*w-Σ-na	*w-Σ-j-na	*w-Σ

Table 14: Reconstruction of the person-marking paradigm for proto-West Gyalrongic

to *-n* in West Gyalrongic, which caused this suffix to merge with the reflex of the singular form *-n* (< *-na). As for the first person suffixes, they evolved differently in Khroskyabs and Horpa: Horpa languages, except when a third person P *-w- interfered, simplified both *-ŋa and *-jŋa into *-ŋ*, whereas Khroskyabs conserved the distinction with *-ŋ* (< *-ŋa) and *-j* (< *-jŋa). The only point difficult to explain is the *-i* second person singular of Geshiza and Dgebshes, which must have resulted from fusion, through palatalization, of *-w- with *-n.

The system depicted brings West Gyalrongic closer to core Gyalrong, as shown in Table 15, which is an adaptation of the Kyomkyo Situ verbal paradigm from Prins (2016: 357).

A \ P	1SG	1DU	1PL	2SG	2DU	2PL	3SG	3DU	3PL
1SG 1DU 1PL				*ta-Σ-n*	*ta-Σ-n-dʒ*	*ta-Σ-j-n*	*Σ-ŋ* *Σ-dʒ* *Σ-j*		
2SG 2DU 2PL	*ko-Σ-ŋ*	*ko-Σ-dʒ*	*ko-Σ-j*				*tə-Σ-w* *tə-Σ-n-dʒ* *tə-Σ-j-n*		
3SG 3DU 3PL	*wu-Σ-ŋ*	*wu-Σ-dʒ*	*wu-Σ-j*	*to-Σ-n*	*to-Σ-n-dʒ*	*to-Σ-j-n*	*Σ-w* *Σ-n-dʒ* *Σ-j-n*		

Table 15: Paradigm of the Kyomkyo Situ verb

1.3.2 Telic (future) suffix

In Mazur Stau, the slot immediately following the agreement suffix can be filled by a telic suffix *-jæ* (Gates 2021: 324). This suffix exemplified in (29) is cognate with the Tangut suffix *-jij¹*. Apart from the position of the two morphemes in the template, which is identical, the correspondence Mazur Stau *-æ* :: Tangut *-ij* can already be seen for past negation (see Subsection 1.2.7).

(29) *ŋæ χæji dʐa-jæ mjæ-rə*
1SG still well-TEL COP.NEG-SENS
'I'm still not completely well.' (Mazur Stau, Gates 2021: 337)

The suffix [illegible] *·jij¹* in Tangut is traditionally analyzed as indicating future. However, a telic interpretation is also possible. The telic aspect expresses that an action is, or will be, completely finished when a point is reached. In (30) already given above (see subsection 1.2.7), the death of Shi Cao is the ineluctable point towards which is directed the process encoded by the verb [illegible] *sjɨ¹* 'to die'.

(30) [illegible]

ɕiə¹ tsʰew² ŋo² tʰjwə¹ kjɨ¹-mjij²-sjɨ¹-·jij¹ zjij¹ gji²=·jij¹
Shi Cao ill get PFV:DIR-NEG-die$_1$-TEL time[⊠] child=ANTIERG
lʰji²wẹ¹ ŋa² dja²-sjɨ¹ zjij¹ ɣur¹ djij¹.rewʳ²=kʰju¹ rjɨr²-tjọ¹
bequeath 1SG PFV-die$_1$ time[⊠] corpse stairs=SUBE PFV:DIR-put[B]

'Ши Цао заболел и, перед смертью, завещал своему сыну : "Когда я умру, положи моё тело на ступени дворца."' (Solonin 1995: 62)

'Shi Cao fell ill and as he still hadn't died [which he was going to], said to his son: "When I will be dead, put my body below the stairs."' (12R, 133.27.04)

In (31), the telic analysis also fits perfectly, and is coherent with the presence of the inferential [illegible] *-sji²*. This morpheme induces a completion of the action, in which the speaker evaluates (including by anticipation) the source of knowledge after the action or event has taken place.

(31) [illegible]

mə¹ ·wu²-nja²-·jij¹-sji²
sky help-2-TEL-IFR

天將助矣 (Shi et al. 1993: 290)

"Heaven will help you." (Leilin, 06.15B.7)

1.3.3 Inferential (mirative) suffix

In Geshiza Horpa, a suffix *-sʰi* denoting inferential access to the information (also present in Stau as *-sə* and in Wobzi as *-si*) has a cognate in Tangut with a suffix [illegible] *-sji²*, formerly glossed in the literature of Tangut grammar as a perfective.[36]

From a homonymic and distributional point of view, this is perfectly coherent. As in Geshiza, where a nominalizing perfective suffix *-sʰi* can be observed, the Tangut [illegible] *-sji²* has a determinative counterpart [illegible] *sji²*, whose role is analyzed

[36] This cognacy is mentioned by Lai et al. (2020), who give other examples, though without mentioning the homonymous nominalizer.

below in Section 2.[37] Thus, the question here is not to discuss the cognacy of the two morphemes, which is beyond doubt, but to be confident about the accuracy of the 'inferential' value in Tangut.[38]

An interesting two-fold fact supports the view that *-sji²* is related to a category that pertains to modality or evidentiality. The suffix is absent from the legal texts glossed by Downes (2018) (which do not contain reported speech), and it usually occurs in (semi) direct speech. Such an absence would be difficult to account for by considering *-sji²* as a perfective, as one should expect such a category to appear regardless of the document type. Conversely, the label 'inferential' matches the distribution well, as depicted by the examples given in (32).[39]

(32) a. *<dàngāo>=tə vluvzâŋ=ɣə u-dzí=si*
cake=DEF Blobzang=ERG PST.INV-eat$_2$=IFR
Blobzang a mangé le gâteau (Blobzang ate the cake). (Wobzi, Lai 2017: 495)

b. *ʐə noŋ wrə dæ-lvo-sʰi*
field in water PFV-freeze-IFR
The water in the field has frozen. (Geshiza, Honkasalo 2019: 606)

c. *pʰio²=tja¹ kʰjɨ¹ lạ¹ mjij¹ nja² pʰio² mjij²-śjij¹-sji²*
snake=TOP leg arm COP.NEG 2SG snake NEG-achieve-IFR
"У змеи нет ног и рук. Ты нарисовал не змею." (Solonin 1995: 43)
"A serpent does not have arms and legs! You didn't draw a snake." (12K, 132.37.06)

The 'inferential' label usually refers to a grammatical category encoding conscious knowledge that the information was not obtained through first-hand eye witness (Tournadre & LaPolla 2014). In (32a) the speaker did not see Lobsang eat the cake; he realizes the cake has been eaten when returning to the place where it was placed. The same goes for (32b): the speaker did not witness the process by which the water froze (a process marked by the presence of the perfective). In (32c), each painter occupies his place, and we can assume that the speaker has to use the inferential, as he did not witness the process by which the other painter drew the wrong animal.

Tangut, like other languages, can employ the inferential when indicating the surprise of the speaker, with a mirative overtone:[40]

[37] The same homonymy between the inferential and a perfective nominalizer *-sə* can also be observed in Stau.

[38] In Lizu (Chirkova 2017), a cognate *=sæ* 非親見標記 'inferred' can also be found.

[39] A fact recalled in a statement given by Lai (2017: 496) for Wobzi Khroskyabs: "dans les histoires et les récits, le marqueur *=si* est très fréquemment attesté. *(In stories and tales, the marker =si is widely attested).*"

[40] The term 'mirative', coined by DeLancey (1997), then re-employed by Aikhenvald (2004), is

(33) a. *jdə̂=tə pə̂ kə-rpʰə̂m=si*
water=DEF all PFV-freeze$_2$=MIR
"Le fleuve est complètement gelé !" ("The water has completely frozen!") (Lai 2017: 497)

b. *ləspə rə-lxua-sʰi*
body PFV-gain.weight.3-IFR
"You have gained weight!" (Honkasalo 2019, p.606)

c. [illegible]
tʰja¹ lew¹.kjɨɨr².dzjwo² ·jij¹gu² dạ²jə² tʰji² ·o²
DEMDS servants altogether say DEMPR alcohol
dja²-ɣiə¹njij²-sji²
PFV-scant-IFR
'Слуги решили так: "Этого вина очень мало."' (Solonin 1995: 43)
'The servants said altogether: "There is now very little of this wine!"' (12K, 132.37.01)

1.3.4 Modal enclitics

Here I examine two modal particles found in Tangut, which also have counterparts in Horpa: the uncertainty marker [illegible] *=mo²* and the interrogative [illegible] *·jaa¹*. The cognacy of this last marker with Geshiza *-jɔ* seems to be plausible, but in contrast with [illegible] *=mo²* and [illegible] *-sji²*, this candidate does not benefit from any distributional argument. In any event, the existence of potential cognates elsewhere in Qiangic tends to point more to a retention than to an innovation.

1.3.3.1 Irrealis [illegible] =mo²

The syllable [illegible] *mo²* in Tangut represents three different morphemes: it can either be a conjunction indicating a choice (Duan 2015), a question marker, or a particle seen in the apodosis of conditional constructions, indicating the incredulity of the speaker facing a consequence, glossed here as irrealis (34).

(34) [illegible]
[illegible]

used here due to the benefit it brings of referring to an overtone attested in Tangut, as one can see. However, the justification for a separated grammatical category is still not evident, as pointed out by Hill (2012). Regarding (32a), the author states that the sentence is uttered in an exclamative mood, making the label 'mirative' also applicable here. Conversely, in (33a), a basic inferential interpretation is also possible, as the speaker did not witness the process by which the river froze; likewise in (33c), where the servants did not see how wine became so scarce.

sjij¹ nja²=·jij¹ tśjị¹ ŋjir¹ dja²-lʰew²-pʰji¹-nja²=tja¹
now 2SG-GEN sorrow misfortune PFV-liberate-CAUS-2=TOP
mjɨ¹-lew¹=mo² *nioow¹ xja¹* *śjɨ¹-ŋa² ·jɨ²-nja² thjij²-ɣiej¹*
NEG.MOD-satisfied=IRR after quickly go₁-1 say-2 INTRG-true

"Ныне вы недовольны тем, что я избавил вас от беды и горестей и теперь говорите: "Я быстро уйду". Почему это?" (Solonin 1995: 37)

"Today I have delivered you from your misfortunes, but you seem not satisfied with it and then say: "I will leave quickly" How can it be?" (12K, 132.15.07)

As already mentioned by Lai et al. (2020: 195-196) and Gates (2021: 338-339), Mazur Stau Horpa has a clause level irrealis enclitic *=mo* which is cognate with [illegible] *mo²*. This cognate occurs in three different contexts which are similar to those of the Tangut [illegible] *mo²*. In (35), *=mo* marks the apodosis of the conditional construction.

(35) *tɕhəgɛ ŋæ=tɕhæ mtshere və-rə ɲi thi pi kɛ-qhəmæ vi regɛ*
then 1SG=on scold do-SENS 2SG DEMDS like INTENS-bad do.2 and
smeze=ji mo=ɲi qə-re də-rə=mo jə-rə
girl=GEN eye=PL go.blind-NMLZ have-SENS=IRR say-SENS

"Then she scolded me, "To do like this is bad and the girl's eyes will be blinded."" (Mazur Stau, Gates 2021: 339)

1.3.3.2 Interrogative [illegible] =·jaa¹

There is in Tangut an interrogative [illegible] *·jaa¹* found in rhetorical questions. I found a similar morpheme in the same contexts in Geshiza (1.3.4). Example (36) in Geshiza illustrates a situation where the speaker knows that the recipient will eat, just like the Tangut example (37), in which the speaker knows that he is not going to meet with Fu Zijian.

(36) *dzi mi-ŋgi=jɔ*
food NEG-eat.2=Q

Don't you eat?! (Honkasalo 2019: 611)

(37) [illegible]

ŋa² tʰja¹ ·wjɨ¹ tsə¹tsʰja²=rjir² tʰjij²sjo² ber¹-ŋa²=·jaa¹ ·jɨ²
1SG DEMDS Fu Ziqian=COMIT how meet-1=INTRG.RTH say

"[Почему я...] встрече с Фу Цзы-цянем радуюсь?" (Solonin 1995: 49)

"How could I meet with that Fu Zijian?" (12K, 132.65.07)[41]

[41]My translation differs slightly from that of Solonin in this sentence.

A cognate of that interrogative can be found in Lizu, as demonstrated by example (38):

(38) *jô ǣ-bǣ mê̂=dʒo tê jê*
self VOC-father NEG=EXV.ANM one Q
"Am I the one without a father?" (Lizu, Chirkova 2017)

1.4 Summary

As a conclusion to the present section, I give in Table 16 an overview of the cognates found in languages of the macro-Qiangic family for each position of the Tangut verb template. The data is from Honkasalo 2019 (Geshiza), Gates 2021 (Mazur Stau), Lai 2017 (Wobzi), Prins 2016 (Kyomkyo), Zhang 2013 (Yuexi Ersu), and Bai 2019 (Munya). Three Lizu forms (in parentheses) are from Chirkova & Wang (2017) and Chirkova (2017): the inferred *=sæ*, the interrogative *jê*, and the uncertainty particle *mɐ*.[42] The coloring in red indicates a cognacy fostered by external information, either semantic or distributional. For the orientational preverbs, it implies *direct systemic and semantic correspondences* (i.e., the preverbs function by pairs consistent with those found in Tangut and/or display an identical behavior). For the negative and modal preverbs, the coloring indicates that their roles are similar to that found in Tangut. For agreement suffixes, relatedness with the stem is marked in blue, and relatedness with the suffix in red. Finally, the cognacy for the inferential *-sʰi* of Geshiza is supported by homonymy with another morpheme, i.e., Mazur Stau's irrealis enclitic *=mo*.

The agreement system is the feature that is the most suggestive of Tangut's Gyalrongic affiliation, whereas cognacies with other tokens of the Tangut verbal slots strongly point to a West Gyalrongic affiliation.[43] I include the reconstructed form of the proto-West Gyalrongic verbal paradigm to allow a clearer view of the correspondences.[44] The proximity is particularly manifest with Geshiza and Mazur Stau, which are the only languages displaying a second series of orientational preverbs with similar usages (see subsection 1.2.6), three negative preverbs with nearly identical behavior, and more post-verbal cognate morphemes.[45]

[42] In Ganluo Ersu, the forms for the orientational preverbs would be *nɛ-*, *ŋɛ-*, *kʰɛ-*, and *tʰɛ-*, respectively. The Lizu forms for which cognates can be found in Ersu are not indicated.

[43] The second person patient forms of Mazur Stau are innovative. In second person patient contexts, the agreement system of Mazur Stau ceased to index the patient

[44] I reconstructed these forms without taking into account the Kyomkyo Situ paradigm, whose proximity should thus be seen as no more than a confirmation.

[45] Even if it is beyond the scope of the present study, phonetically speaking, the proximity between Geshiza and Tangut is also more evident, with correspondences in articulation between the vowels, the exception G. *-æ* :: T. *-ɨ* occurring only in post-velar contexts. Knowing that Tangut *ɨ* and *ə* are in complementary distribution, the former only occurring after the yod *-j-* (i.e., one of the controversial elements of Gong Hwang-Cherng's reconstruction), the correspondence T. *ɨ* :: G. *ə* could actually be seen as T. *ə* :: G. *ə*.

	Tangut	pWR	Geshiza	Maz. Stau	Wobzi	Kyomkyo	Y. Ersu / (Lizu)	Munya
1	·a- / ·jij^1-				æ-			
	nja^1- / njij^2-		næ- / ni-	nə- / né-	næ-	na-	nə-	no-
	·wjɨ^2- / ·wjij^2-		wə- / wi-	ɣə- / ɣé-	və-		ŋə-	ɛ-
	kjɨ^1- / kjij^1-		gæ- / gi-	kə- / ké-	kə-	kə-	kʰə-	ɣɤ-
	dja^2- / djij^2-		dæ- / di-	tə- / té-			tʰə-	tʰo-
	rjir^2- / rjijr^2-		rə- / ri-	rə- / ré-	rə-			
	·a-		æ-	æ-	æ-		a-	
2	mji^1-		mi-	mi-	mə-	ma-	ma-	
	mjij^2-		mɛ-	mæ-				mo-
	mjɨ^1-		mə-	mə-				
	tji^1-		di-	ti-	tə-		tʰa-	tɕɯ-
3	tśʰjɨ^1-		-tɕʰi					
	ljɨ^1-				də-			
	tsjɨ^1-							
6	1SG:3 (Σ_2) -ŋa^2	*-wŋa	-w	-u	-ŋ	-ŋ		
	2/3:1SG Σ_1 -ŋa^2	*-ŋa	-ŋ	-ã	-ŋ	-ŋ		
	2SG:3 (Σ_2) -nja^2	*-wna	-i	-i	-n	-w		
	1/3:2SG Σ_1 -nja^2	*-na	-n	-u / -ã / ø	-n	-n		
	1PL:3 Σ_1 -nji^2	*-jŋa	-ŋ	-ã	-j	-j		
	2/3:1PL Σ_1 -nji^2	*-jŋa	-ŋ	-ã	-j	-j		
	2PL:3 Σ_1 -nji^2	*-jna	-n	-n	-n	-jn		
	1/3:2PL Σ_1 -nji^2	*-jna	-n	-u / -ã / ø	-n	-jn		
7	-jij^1			-jæ				
8	-sji^2		-sʰi	-sə	-si		(=sæ)	
9	-djij^2							
M1	=mo^2			=mo			(mɐ)	
M2	=jaa^1		=jɔ				(jê)	

Table 16: Overview of the cognacy with each Tangut verbal slot in macro-Qiangic

2 Noun phrase

Many similarities between Tangut and Horpa languages, especially with Geshiza Horpa, can also be observed in the domain of the nominal phrase.

First of all, Tangut, as Horpa and Khroskyabs languages, has lost the possessive proclitics still seen in core Gyalrong.[46] There are also specific morphemes and functions of those morphemes that suggest Tangut's placement within the Horpa subgroup specifically and not just in West Gyalrongic. I analyze here two kinds of evidence.

Firstly, I present numeral cognates shared by Tangut and Geshiza (Subsection 2.1). Secondly, I give examples of cognate nominalizer enclitics in West Gyalrongic languages (Subsection 2.2).

2.1 Numerals

Table 17 lists Tangut numerals and their cognates in Geshiza Horpa, Mazur Stau Horpa, Siyuewu Khroskyabs, Wobzi Khroskyabs, Japhug Gyalrong, and Munya (Qiangic). I have colored innovations in red, and retentions in blue. The phonetic correspondences will be discussed in detail in a future paper.[47]

Num.	Tangut	Geshiza	Stau	Siyuewu	Wobzi	Japhug	Munya
1	lew^{1}	rəu	ro	rǽ̂ɣ	rɑ̂ɣ	*ci* / tɤɣ	to-
pref. 1	·a-	æ-	æ-	ə̂-	ə̂-	tɯ-	
2	n^{j}ɨɨ^{1}	wne	ɣne	ɣnǽ̂ɣ	jnǽ̂	ʁnɯz	nə-
3	sọ^{1}	ws^{h}u	ɣsu	xsə̂m	çsə̂m	χsɯm	sɔ-
4	l^{j}ɨɨr^{1}	wʑæ	ɣɮə	vdə́	vdə́	kɯβde	rə-
5	ŋwə^{1}	ŋuæ(< *w-)	ɴɢvɛ	mŋɑ́d	mŋɑ́	kɯmŋu	ŋɑ-
6	tś^{hj}iw^{1}	wtɕ^{h}əu	ɣtɕ^{h}o	xtɕéɣ	ftɕú	kɯtʂɤɣ	tɕ^{h}ü-
7	ś^{j}ạ^{1}	sɲe	zɲe	sɲǽ̂	sɲê	kɯɕnɯz	nyü-
8	·jar^{1}	rjɛ	rjɛ	vjɑ́d	vjɑ́	kɯrcat	ɕo-
9	g^{j}ɨɨ^{1}	ŋgæ	ngə	ŋgə́d	ŋgə́	kɯngɯt	ngɯ-
10	ɣạ^{2}	zɣa (ɣæ)	sʁɑ	sjə̂d	sjə̂	sqi	-ɣɑ́ (ɣɔ-)

Table 17: Numerals in Tangut, Geshiza, Stau, Khroskyabs, Japhug and Munya

The numeral 'one' has a prefixed version consisting of a simple vowel and is usually used with classifiers, which I included in the second row of Table 17. This monovocalic morpheme is not an innovation, as it can be found in Qiang (LaPolla & Huang 2003), where it is subject to vowel harmony.[48] However, in

[46]This feature may be areal, as pointed out by Sun (2019), as Tibetan works in a same way.

[47]The preinitial x- in Siyuewu is an allophone of ɣ- in an unvoiced context; as one can see, correspondences between Siyuewu and Wobzi regarding preinitials are not only systematic but also proportional.

[48]This could actually point towards the phenomenon of root suppletion for 'one' in Qiangic as a retention, the innovation being the reanalysis from the CVC morpheme.

the case of Tangut and West Gyalrongic languages, the cognacy is confirmed by the homophony found in all the languages between this numeral prefix and the interrogative prefix, which is not the case for Qiang.[49] Munya only has a prefixed numeral which occupies the two first lines.

Tangut lost codas, like Wobzi (except for the rhyme *-əm*), Geshiza and Munya. But some developments are specific to Tangut and Horpa, in particular Geshiza. First, in Tangut and Geshiza the rhymes for the numerals 1 and 6 underwent the same reduction of a former *-ɣ into *-w*, with a scalar raising (*-ew* → *-iw*) of the vowel based on the presence or absence of -j- in Tangut. Second, Tangut 𗏁 *ŋwə¹*, Geshiza *ŋuæ* (numeral 5), and Stau *nɢvɛ* went through a metathetic loss of a former preinitial *w- (*wŋ-* → *ŋw-*).[50]

Some *w-* preinitials in Geshiza are retentions (see 'four'), while others are innovative (see 'two', 'three'). The proximity between velars and labio-velars makes it sometimes difficult to identify the place of the innovation (in terms of place of articulation, as the consistant voicing of the preinitial brings Geshiza and Stau together). If Geshiza reanalyzed all its numerals with *w-* ('two' and 'three' were probably influenced by the form for 'four') and Stau with *ɣ-* (see 'four'), there is a discrepancy in Khroskyabs between Wobzi and Siyuewu which impedes a straightforward understanding of what happened for 'six'. For that numeral, two scenarios seem possible. Geshiza and Wobzi forms could be innovative, as could be Stau and Siyuewu forms. I favor the second hypothesis, as Japhug's form for 'four' shows that a labial preceding a stop should be kept, which probably indicates for that language the loss of a fricative velar (a loss made easier by the overall back vocalic context of the prefix preceding the affricate). The innovative *f-* of Wobzi should not automatically be analyzed as a common innovation with Geshiza, though. Wobzi has itself a tendency to simplify velar preinitials, as show the palatalized forms of 'two' and 'three' before continuants, and this innovation could be independent.

This chart offers new insights into Tangut phonology. The tenseness observed in the numeral 𗍫 *śjạ¹* 'seven' cannot come here from a former *S-, and could be the result of the simplification of a former complex onset *śɲ-, as the coronal fricative is still present in Tangut.[51] An interesting distributional fact could also interest the value of vowel lengthening in Tangut, which is still a controversial topic. As one can see, all Tangut numerals with a long vowel correspond to Geshiza words with sonorant preinitials, either nasal before a plosive voiced initial, or approximant before a continuant. It is actually not impossible that the form seen in Geshiza corresponds to the actual pronunciation in Tangut, a subject which will be explored in future work.

[49] Moreover, in Qiang number counting causes *a* to appear even without a classifier. However, for numeral + classifier combinations, West Gyalrongic languages use the CVC form from the first row in Table 17.

[50] There are cognates in Ersu and Duoxu, where the numeral 'five' is respectively *ŋwá˞* and *ŋo³²* (Chirkova 2014).

[51] Most of the tense vowels of Tangut are usually seen as remnants of former *S- (see Gong 1999).

That is not all. The proximity between Geshiza and Tangut is apparent, once again, in common specific uses of numerals. In Geshiza, the prefixed alternation for the numeral ‘one’ has meanings ranging from the use with a classifier (*æ-ɣi*, one-CLF.PERS = ‘one person’) to the use as a prefix deriving collective nouns (*æ-stɕəpa*, one-villager = ‘all villagers’) or indicating approximation in a succession of numerals (*æ-wnæ-sq^{h}a*, one-two-ten = ‘twenty-something’). One can find similar usages in Tangut. First, as in Geshiza (and also Stau: *æ-ʁe* and Siyuewu Khroskyabs: *ə̂-ʁe*, one-CLF), the bound form [illegible] *.a-* is used with classifiers (39).[52]

(39) [illegible]

lṳ^1 kjɨ^1-ljị^1 ·a-kjiw^1 rar^2 nioow^1 [ɣu^1 lṳ^1 ljị^1
stone PFV:DIR-plant one-year pass POSTE [before stone plant
ljị^2]=do^2 ·juu^1=śji^2
place]=TERM see=go$_2$

經一年，往所種地看 (Shi et al. 1993: 303)

‘One year passed after he had planted the stone; he went to look at the place he had planted the stone before.’ (Leilin 07.26.B.1)

Even if [illegible] *·a-* does not derive a collective noun of formerly separated entities as *æ-* in Geshiza (as does also Siyuewu Khroskyabs for that matter: *ə̂-rgænrgən*, one-wife.and.husband), it nonetheless has the same meaning of ‘all’. In (40), [illegible] *·a-* indicates the entirety of the defined country (differing from the prototypical indefiniteness usually observed with classifiers):[53]

(40) [illegible]

·a-l^{h}jij zji^2 niəj^1=k^{h}a^1 lew^1 ŋa^2 tjij^1 sej^1 ·ji^1-dzjwo^2 zji^2
one-country all mired=INTERE only 1SG alone clear lot-people all
lia^2=k^{h}a^1 lew^1 ŋa^2 tjij^1 gji^1=jɨ^2
drunk=INTERE only 1SG alone be_awake=say

一國皆濁，惟我獨清，眾人皆醉，惟我獨醒。(Shi et al. 1993: 261)

“In the whole country, while all are mired, there is only me to be clear; while a lot of people are drunk, only I am awake.” (Leilin 03.22B.6)

Finally, Tangut makes use of the same strategy to encode approximate number by a succession of numerals (41), a behavior not seen in Khroskyabs (Lai Yunfan, personal communication):

[52] This suppletion pattern for ‘year’, seen also in Stau, has already been mentioned by Jacques (2014: 213).

[53] This behavior is also seen in Siyuewu Khroskyabs: *ə̂-dɣəm*, one-family (Lai Yunfan, personal communication).

(41)

la² *·ju²* *źjiw¹sji¹* *·a-njɨɨ¹-ɣạ²* *pʰu²* *·o¹*
tomb ANTE cypress one-two-ten CLF EXV.on

墓前有數柏樹 (Shi et al. 1993: 264)

'There were around twenty cypresses in front of the tomb.' (Leilin 03.30A.5)

2.2 Nominalizers

Nominalizers form another grammatical category in which morphemes from Tangut and Horpa languages display similarities exceeding mere resemblance. Some nominalizers embed clauses inside other clauses, primarily to construct relative clauses.

2.2.1 The subject/agent nominalizer *-mjijr²*

The subject/agent nominalizer *mjijr²* is a very productive morpheme which, when attached to a verb it derives, conveys the meaning 'one who (verb)'. Probably related to the noun *mjir¹* 'people', it is also cognate to core Gyalrong Japhug *turme* 'human being, someone else' and Zbu *tərmé?* 'man (masculine), someone else.' The morpheme's nominal origin could explain why the nominalized form can often stand by itself as a noun, even if an interpretation as a headless relative clause is sometimes possible, as in (42).

(42)

tʰji² *ɣwie¹=do²* *dzjwo²* *gjɨ²* *dźjiij¹*
DEMpr strength=TERM man INDF EXV
·jij¹-luu¹=·wji¹=mjijr²=tja¹ *rejr²=ljɨ¹*
OPT-digging=LV:do[A]=NMLS:A=TOP be_numerous=EXCLAM

'Чтобы человек пребывал в этом могуществе, тех, кто копает, должно быть много.' (Solonin 1995: 54)

'In order for a man to have such a strength (= to be in that strength), those who need to dig are many.' (12K, 133.02.04)

The behavior of *mjijr²* is in every respect similar to a nominalizer *-me* found in Geshiza.[54] Both of these highly productive markers encode the subject or the agent of the verb's action, and follow the same trinomial pattern.[55] Table 18 gives examples of nominal derivations for stative (S-NMLZ, first pattern) and transitive

[54]Considering the relatedness between Tangut and Geshiza regarding that nominalizer, the etymology points in Geshiza towards a native element of the lexicon and not a loanword from Tibetan མི་ *mi* 'person'.

[55]In Siyuewu Khroskyabs, a cognate *=mə* can also nominalize the patient in real configurations. I do not find such uses in my corpora.

verbs, either realized in a simple way (A_0-NMLZ, second pattern) or by incorporation of the object (A_1-NMLZ, third pattern). Here, Tangut and Geshiza display exceptional morphological stability across centuries.[56]

	source	*output*
	S-NMLZ	
Geshiza	*ŋo* 'to be sick'	*ŋo-me* 'sick person'
	sʰæ 'die'	*sʰæ-me* 'dead person'
Tangut	[illegible] *tʰjwị¹* 'to be young'	[illegible] *tʰjwị¹-mjijr²* 'young person'
	[illegible] *lia²* 'to be drunk'	[illegible] *lia²-mjijr²* 'drunk person'
	A_0-NMLZ	
Geshiza	*mdʑəska* 'to watch'	*mdʑəska-me* 'spectator'
	v-ræ 'write'	*ræ-me* 'writer'
Tangut	[illegible] *djị²* 'to cure'	[illegible] *djị²mjijr²* 'healer'
	[illegible] *rjar¹* 'to write'	[illegible] *rjar¹-mjijr²* 'writer'
	A_1-NMLZ	
Geshiza	*ʈʂhetsə læ* 'to drive a car'	*ʈʂhetsə-læ-me* 'driver'
	dzi və 'to cook food'	*dzi-və-me* 'cook'
Tangut	[illegible] *rjijr¹ śioow¹* 'to raise horses'	[illegible] *rjijr¹-śioow¹-mjijr²* 'esquire'
	[illegible] *źju² lju²* 'to catch fishes'	[illegible] *źju²-lju²-mjijr²* 'fisherman'

Table 18: Subject/Agent nominalization with Geshiza *-me* and Tangut *-mjijr²*

As stated above, nominalization with [illegible] *mjijr²* is highly productive. Sometimes the operation exceeds the scope of mere object incorporation, to produce a fully lexicalized output. In (43a) [illegible] *ʲjwɨr²-la¹-mjijr²* 'text-transcribe-NMLS:A') can be translated as 'scribe'. With the nominalized output of this example, it is actually the frequency of use of the nominalized form that can support the attribution of genitive value to [illegible] *=jij¹*, as examples where [illegible] *=jij¹* marks a recipient can also be found, like the one given in (43b).

(43) a. [illegible]

tśjiw¹ sjwa¹ njij²=jij¹ jwɨr²-la¹-mjijr² ŋwu²
Zhou Xuan king=GEN scribe COP
周宣王史官也 (Shi et al. 1993: 311)
'He was king Zhouxuan's scribe.' (Leilin 08.21.A.5)

[56] The technique by which this derivation occurs seems to be areal. The Ersu/Lizu/Duoxu nominalizers *su* / *ɕu* have exactly the same functions and are also derived from 'man/person' (Katia Chirkova, personal communication). Here though, apart from the phonetic proximity with Geshiza, plus the ascertainable Gyalrongic etymology for Tangut (the *-r* coda of [illegible] *mjijr²* reflects a former preinitial *r-* still present in core Gyalrong), the trinomial patterns of Geshiza and Tangut fit perfectly.

b. [illegible]

tjɨ^1*tjij*^{1} *bji*^{2}*mjijr*^{2} *ljə̣*^{1} *tśju*^{1}*ljij*^{1}*dzjwo*^{2}*=kʰa*^{1} *njɨɨ*^{2}*njɨɨ*^{2} *ŋa*^{2}*=·jij*^{1}
CONJ officials and servants=INTERE everyday 1SG-ANTIERG
dźjar^2*-tsʰjiij*^{1}*-mjijr*^{2} *dju*^{1} *ku*^{1} *tʰja*^{2}*=·jij*^{1} *bjuu*^{2} *kʰwej*^{2} *kʰjij*^{1}*-ŋa*^{2}
fault-say[⊠]-NMLS:A EXV then 3SG=ANTIERG reward great give[B]-1
·ji^2
say

"Если сановники и должностные лица ежедневно станут [мне] говорить о моих проступках, то я дам им большую награду." (Solonin 1995: 45)

"Were there among the officials and servants someone reporting my misconducts everyday to me, I would reward him." (12K, 132.53.06)

2.2.2 The determinative nominalizer [illegible] *sji*^{2}

Geshiza and Tangut have a cognate evidential suffix and nominalizing/determinative morpheme, which are homophonous in each language. This suffix behaves very similarly in the two languages, except that the nominalizer *-sʰi* is described in Geshiza as being attached to perfective/past stems (vs. *-me*, attached to non-past stems), whereas in Tangut, no aspectual restriction seems to be required. The determinative in Tangut also requires a nominal head to its right.[57]

(44) a. *rgævæ gæ-jəu-sʰi æ-lə də-ræ*
stone IPFV-grow.3-NMLZ one-CLF.INDF EXV-SENS
'There is a stone that grows.' (Honkasalo 2019: 688)

b. [illegible]

sjij^2*ju*^{2}*=do*^{2} *kji*^{1}*-kʰuu*^{2} *·o*^{2}*-kjur*^{1} *sji*^{2} *gju*^{2}
common.people=TERM PFV:IN-search alcohol-pour NMLZ ustensile
gji^2 *·a-rjir*^{1}
INDF PFV:UP-get[A]

於百姓家搜得酒具 (Shi et al. 1993: 276)

'He came to search among the people and got a wine ladle (= an ustensile that ladles wine).' (Leilin 04.06B.6)

There is a particularity attached to Tangut [illegible] *sji*^{2} though, which can be seen in relative clauses. [illegible] *sji*^{2} usually requires a head noun positioned to its right, as (45) shows by displaying a pattern similar to (44b) above. In this way, this morpheme

[57] As does another past nominalizer of the same phonological form found in Dgebshes Horpa (see Tian & Sun 2019 §2.4).

behaves like a determiner – like the particle 的 *de* in Modern Mandarin Chinese, also used to construct relative clauses – even if it is still a nominalizer, as it always follows a verb.[58]

(45)

śja̱¹ kjiw¹ ·we² *sji²* *mjij¹=·jij¹ lʰju¹=ŋwu² lu̱¹dzjwo²* *tji̱¹*
seven year become NMLZ girl=GEN milk=INS stone.person make.eat

ku¹ *kji¹djij²* *lej²* *·ji²*
then certainly change say

須七歲女子以乳之，則當變 (Shi et al. 1993: 294)

'If one makes a man of stone drink the milk of a seven-year-old girl, he will necessarily change.' (Leilin 06.28B.4)

3 Locative case markers

Jacques et al. (2017) list some grammatical cases with common origins in Gyalrongic, including Tangut. This present section focuses on locative cases, the category displaying the most significant number of striking similarities. Core cases like the antiergative / oblique *·jij¹* have counterparts in West Gyalrongic, as illustrated by Lai et al. (2020).[59] However, the exact behavior of these morphemes in Tangut still need further description and analysis before they can be used to ascertain whether Tangut is closer to Horpa or Khroskyabs.

3.1 Overview

I give in Table 19 an overview of some potential cognate cases in Tangut, Geshiza Horpa, and Wobzi Khroskyabs. To help to visualize the proximity between Tangut and West Gyalrongic, I provide below, when possible, related morphemes in Kyomkyo Situ (Prins 2016), Ersu (Zhang 2013), and Munya (Bai 2019). Once again, one can observe some similarities between Tangut, Horpa, and Ersu, as was the case for orientational preverbs in Section 1 (Subsection 1.2). Apart from a palatalized postposition *tɕʰo* 'on', Ersu also has a multi-functional postposition *kə* with

[58] This homophony between a nominalizer and an inferential morpheme is also found in several Qiangic languages (including Muya, Queyu) and Tibetan dialects of Sichuan (Katia Chirkova, personal communication). It could be a feature of an areal nature but could also come from a cognitive alignment resulting from an areal feature (indeed, the postverbal position of the two morphemes might enable one category to influence the pronunciation of the other). The use of the same character in Tangut to transcribe the two morphemes could be a clue for such an interpretation: the nominalizer and the inferential, at least for the scribes who created the Tangut script, shared a certain unity. In any event, this does not perturb the assignment of cognacy between the Geshiza and Tangut morphemes, as cognacy can already be established separately.

[59] This term was coined by LaPolla (1992) to indicate the antagonist of the ergative case in ergative languages, i.e., the oblique argument of the verb. It can refer in Tangut to a semantic object (accusative), a recipient (dative), or a beneficiary (benefactive).

meanings overlapping with the interessive *kʰa*^{1}. There is a Kyomkyo cognate *wəkʰa* ‘among, through’, which is derived from *təkʰa* ‘mouth’.

Tangut	ɣa^{2} LOC	tśʰjaa^{1} SUPE	kʰa^{1} INTERE	/ nioow^{1} POSTE
Geshiza	ɣa LOC	tɕʰa LOC ‘on’	(kʰa APPROX)	ɲo ‘after, because’
Khang.gsar	ʁa ALL	tɕʰa LOC ‘on’	kʰa INS	
Wobzi	ʁɑ LOC	tʰɑ LOC ‘on’		
Kyomkyo			wəkʰa ‘among, through’	
Yuexi Ersu		tɕʰo ‘on’	kʰə ‘inside, among (etc.)’	ɲo ‘outside’
Munya				

Table 19: Potential cognate cases in Tangut, Geshiza and Khang.gsar Horpa, Wobzi Khroskyabs and other macro-Qiangic languages

The semantic shift observed in Table 19 for the cognates of the first column (a general locative in Tangut, Geshiza, and Wobzi; an allative in Khang.gsar Horpa) can be easily explained by a characteristic of locative cases of Tangut and Gyalrongic languages in general, namely the absence of contrast between location with motion and static location in the use of case markers. In Tangut, the only information given is the position in reference to the head, as seen in example (46), where the superessive can be translated both by ‘on the top of’ and ‘from the top of’. The notion of motion from/towards the speaker being inherent to the verb, a semantic shift resulting in the specialization as an allative in Khang.gsar Horpa is then easy to imagine.

(46)

sọ^{1} *tśiẹj*^{2} *ko*^{1}*=tśʰjaa*^{1} *·wjɨ*^{2}*-dzuu*^{2} *sọ*^{1} *tśiẹj*^{2} *ko*^{1}*=tśʰjaa*^{1}
three time vehicle=SUPE PFV:OUT-sit three time vehicle=SUPE
nja^{1}*-lʰjị*^{2}
PFV:DOWN-descend

‘…три раза садился в колесницу и три раза сходил с неё.’ (Solonin 1995: 39)

‘He sat three times on the vehicle, and three times got down from it.’ (12 R, 132.21.02)

3.2 *ɣa*^{2} : general locative

The first cognate locative case to be discussed, found in the Tangut, Wobzi Khroskyabs, and Horpa languages, is the general locative. Examples of this cognate are given below, first in Wobzi (47), then in Geshiza (48), and Tangut (49).[60]

[60] There is in Shuhi (Naic) a locative *ʁɔ̃* ‘on’ that could be related (Katia Chirkova, personal communication).

(47) *ænarêsi <piàn pian>=tə=ʁa jʊŋsâ lbé=çsærpa=tə dzâɣ ró*
but fabric=DEF=LOC:general again urine=new=DEF EXV$_1$ must$_1$
"Mais il faut qu'il y ait de l'urine fraîche sur le tissu." ("But there must be fresh urine on the fabric.") (Lai 2017: 189)

(48) *mtɕʰærten-ɣa skærva dæ-van*
stupa-LOC circumambulation PFV-LV:do.1PL
"We circumambulated the stupa." (Honkasalo 2019: 386)

(49) 

ljụ^2=ɣa^2 kʰwa^{2}lji^1 ·o^1
body=LOC pants EXV.on
（良妻）身著布裙 (Shi et al. 1993: 309)
'[She - Wang Liang's wife] had pants on her.' (Leilin 08.11.A.4)

The case has become unproductive in Geshiza, but there are other varieties of Horpa such as Stau where it is still widely used, with different degrees of productivity and functions. Described previously as an allative (Jacques 2017: 604), its functions are actually more numerous. In Mazur Stau (Gates 2021: 307-308), apart from the allative function, it can express the semantics of 'on a vertical location' (50a). It can also mark the oblique argument (50b) or the object (50c) of a transitive verb, indicating thus an accusative function.[61]

(50) a. *ɟõ=ʁa pɛr xi-rə*
wall=ALL picture EXV.on-SENS
'There is a picture on the wall.' (Mazur Stau, Gates 2021: 307)

b. *stændzətɕʰuscə rɟɛzo=ʁa kɛ-rcu rgæ-rə Bstan*
'dzin.chos.skyid potato=ALL INTENS-much like-SENS
'Bstan.'dzin.chos.skyid likes potatoes a lot.' (Mazur Stau, Gates 2021: 308)

c. *ŋæ tʰɛ=ʁa ɣə-tʰu*
1SG DEM=ALL PFV-drink.1
"I drank it." (Mazur Stau, Gates 2021: 308)

[61]These usages seen elsewhere in West Gyalrongic show that the allative function is probably secondary and results from characteristics discussed in Subsection 3.1. The emergence of the accusative function could be a side effect of the progressive specialization of the case as an allative. In many accusative languages (like Russian or Sanskrit), allative constructions require the use of the accusative case.

3.3 *tśʰjaa^1* : superessive, time enclitic

The second of the locative cases listed in Table 19 is the superessive, for which an example was given above in (46). It is used in Tangut after a VP or after a demonstrative pronoun (cf. 51 and 52), with a temporal meaning instead of a locative meaning. Kepping (1985) described this phenomenon, seeing in it the marking of the iconic temporal precedence of the first NP to the second (usually translated 'when' in a perfective way for post-VP occurrences, and 'then' for post-pronominal ones).

(51)

·u^2 la^2 źjɨ̣^{1}=bju^1 ·jow^1 sjo^2=·jij me^2=tśʰjaa^1
Middle aunt jealous=because Wang Xiang=ANTIERG sleep=SUPE
tʰaa^1.bjɨ^1=ŋwu^2 sja^1 kiẹ^2
suppress.dagger=INS kill desire

後母患之，乃持刀往祥所斫之 (Shi et al. 1993: 263)

'As his middle aunt was jealous, she wanted, when Wang Xiang would fall asleep, to kill him with a sword.' (Leilin, 03.29A.6)[62]

(52)

gji^{2}bjij^2 ·wjɨ^2-lʰo^0 ·wa^1 ·juu^1 ljij^2 tʰja^1=tśʰjaa^1 tśʰjɨ^{2}rjar^2
wife PFV:OUT-go_out pig see come DEM=SUPE immediately
dja^2-lju^2
PFV-catch

其妻乃出看豬，遂擒之。 (Shi et al. 1993: 289)

'The wife went out, saw the pig; then she immediately caught it.' (Leilin 06.09B.1)

The same uses of the superessive can be found in Horpa languages, e.g., (53 and 54). In Geshiza and Stau, however, it indicates simultaneity, not iconic succession. This interpretation could also be more accurate for Tangut; in (52), the enclitic, attached to a perception verb, strongly indicates immediacy, calling into question an analysis of iconic succession.

[62] The interpretation 'bedroom' for the word *me^2* seems unlikely, as the locution *me^2=twụ^1* 'bedroom' occurs in the next sentence, with a clear verbal value for *me^2* in the tatpuruṣa compound ('the place to sleep'). Near nominal interpretations of the verb do exist though, e.g., in association with the medessive postposition in *me^2=gu^2*. In that case, *me^2* can be analysed as an action nominal, i.e., a non-finite form of the verb.

(53) *ŋa ʈʂaʃi jo ʁ̥e-ŋ=tʃʰa tʰə=ɣu dʒa v-tʰi-gə*
1SG Trashi home arrive$_2$-1=LOC 3SG=AGT tea MARK.TR-drink-NMLZ
ɟi-rə
EXV-MED
"When I got to Trashi's home, he was drinking tea." (Dgebshes, Sun 2019)

(54) *tʰi tɕʰa*
DEM.GEN on
'on this/that (spatial meaning), then' (with the temporal meaning 'at that time') (Honkasalo 2019: 335)

Such similarities in the locative marker's morphology *and* its uses (same morpheme to encode both specific spatial and temporal meanings diverging from each other) strongly suggest considerable closeness between Tangut and the Horpa languages.[63]

3.4 𗂧 *kʰa¹*: interessive, instrumental, superlative

The enclitic 𗂧 *=kʰa¹* usually controls nouns referring to things not having definable existence, typically natural elements without boundaries (Beaudouin to appear). In such cases it means 'in'. It can also be attached to collective nouns, with the meaning 'among'.[64] These meanings, proper to Tangut, are not far from the label 'approximative' label of Geshiza. However, the surface form *=kʰa* in that language is not a part of the case system, even if it is close to the instrumental of Stau varieties.[65] Attached to a VP, 𗂧 *=kʰa¹* also has a temporal meaning 'when'.[66]

The status of 𗂧 *=kʰa¹* is sometimes unclear if one only uses the traditional interpretations of that morpheme in Tangut. Comparative evidence again permits some new conjectures. There are some scenarios where an instrumental interpretation is possible, like in Khang.gsar Horpa (close to the meaning 'through' of the cognate form in Kyomkyo Situ), as in the examples presented in (55).[67]

[63] It is indeed the cooccurrence of these two similarities (both the semantic and the morphological) that tend to indicate cognacy, as the temporal use of locative words is otherwise typologically well attested (cf. French 'sur ce', English 'thereupon', etc.). Note here that in both languages the surface form is identical regardless of the function, in contrast with French and English.

[64] Yuexi Ersu Zhang (2013: 285) has a bound enclitic *=kə* which matches strikingly well this description.

[65] This proximity is even more evident in Stau itself (Mazur variety), which has an approximate time enclitic *=kʰæ* related to Geshiza *=kʰa*, and an instrumental *=kʰæ* related to Khang.gsar *=kʰa*. Even if these two enclitics differ in functions, they could still be specializations of a unique original form, in different contexts.

[66] Both Lizu and Ganluo Ersu also have a form *kʰæ* and *xa*, meaning 'when', which is most likely derived from a locative (Katia Chirkova, personal communication). The Yuexi Ersu bound enclitic *=kə* can also be employed with temporal meanings.

[67] The primary use of 𘓷 *ɣwej¹* is verbal, with the prefix 𗍊 *ʼa-* and the comitative 𗧓 *=rjir²*; [A] 𗧓 𗍊𘓷 = 'fight against/with [A]'.

(55) a. [illegible]

ŋwu^{2}kwar^1=k^{h}a^1 t^{h}ja^1 ·we^2 nja^1-kiew^1
cry(ing)=INTERE DEMDS city PFV:DOWN-collapse

Not translated by Solonin.

'The city collapsed in/with/through tears.' (12K, 132.47.07)

b. [illegible]

tjij^1 dja^2-bej^1-nja^2 ku^1 ɣwej^1=k^{h}a^1 kạ^1.bja^2-nja^2
if PFV-loose-2 then fight(ing)=INTERE life.break-2

"...если же потерпите поражение, то в сражении закончите жизнь." (Solonin 1995: 43)

"If you lose, then you will perish in/by/through fighting." (12K, 132.38.04)

There are also occurrences (56) of a superlative derivation of [illegible] *k^{h}a^1*, a pattern of derivation found in Geshiza for postpositions (*zə-tɕ^{h}a*, SUPL-on = '(on) the highest above', Honkasalo 2019: 333).[68] However, this derivation seems to be fossilized in Tangut, as it is only attested with [illegible] *k^{h}a^1*.

(56) [illegible]

ljij^2·jwĩ^1 tsjiir^1=tja^1 zji^2-k^{h}a^1 k^{h}wej^2 ŋwu^2 t^{h}ja^2=tś^{h}jaa^1 l^{h}u^1=tjị^2
lingyin rank=TOP SUPL-INTERE big COP DEMDS=SUPE add=NMLZ
mjij^1
COP.NEG

'Ранг линъинь - самый большой, [какой можно измыслить]. Выше него прибавить нечего.' (Solonin 1995: 43)

'The rank of lingyin is the highest of all and there is nothing to add above it.' (12K, 132.36.06)

3.5 [illegible] / [illegible] *nioow^1* : time postessive, causal

In Geshiza, a postposition *ɲo* (Honkasalo 2019: 341) can either encode (after an NP or a VP) the precedence of the marked constituent to the main clause (as 'after' in English, 58), or a causal relation (57).[69] Interestingly enough – and conversely to other time postpositions – that morpheme does not convey any spatial meaning.[70] Mazur Stau has a postposition *ʁoɲu* 'behind' (Gates 2021: 228) which does convey spatial semantics. However, in most compound words, the component *ɲu* (cognate to Tangut and Geshiza's forms) has temporal meanings, e.g., *lɛɲu* 'afternoon', *ɲusɲi* 'next day' (Gates 2021: 484).

[68] Stau postpositions can also take the superlative prefix, see Gates (2021: 227).

[69] This causal use is relatively rare, but still exists. (Sami Honkasalo, personal communication). In Tangut documents, [illegible] is also more employed than [illegible].

[70] In example 58, *tɕ^{h}u* is a cognate of Tangut [illegible] *tś^{h}jwo^1*, also seen in causal constructions.

(57) *rgæn=t^{h}ə t^{h}ə læsær ɲo <sæn-jyefən>=ke rgæn*
early.corn=TOP RED New.Year after three-month=DAT early.corn
g-ə-ʐoŋ.
IPFV-NACT-plant.1
"Then (lit. after that) there is the early corn. We plant the early corn after the New Year in March." (Honkasalo 2019: 341)

(58) *ɲi smæn-ræ tɕ^{h}u ɲi ɲo bəra dæ-ʐan.*
2SG like-2-SENS CONJ 2SG because TOPN PFV-come.1
"I like you, so I came to Balang Village for your sake/because of you." (Honkasalo 2019: 341)

In Tangut, there is an enclitic [illegible] *=nioow^1* 'because', which governs the entire clause (after an NP or a VP, see examples below), and which has a homonym postposition [illegible] *nioow^1* POSTE (59) expressing the same notion of time as Geshiza *ɲo*.

(59) [illegible]
[illegible]

p^{h}o^2 śioow^1 ŋa^2=·jij^1 dźju^{2}ɣu^1 mji^1-sej^1=tja^1 ŋa^2 mja^1 dźjij^2=
Bao Shu 1SG=ANTIERG weak NEG-consider=TOP 1SG mother have=
nioow^1 nioow^1 p^{h}o^2 śioow^1 ŋa^2=·jij^1 sọ^1 tśiẹj^2 njij^2=do^2
because POSTE Bao Shu 1SG=ANTIERG three time king=TERM
·a-pow^1=·wji^1
PFV:UP-helping=LV:do[A]

"Бао Шу не считал меня слабым и трусливым, ибо [он знал], что у меня есть мать [и я боюсь погибнуть]. Потом Бао Шу три раза помогал мне [встретиться] с ваном." (Solonin 1995: 38)

"If Bao Shu didn't see me as weak, it was because (he knew that) I had a mother (reason because of which the character was afraid to die). Afterwards, Bao Shu three times helped me before the king." (12K, 132.19.06)

(60) [illegible]

sjij^1 ·wjị^{1}dźjwi^1 gji^2=nioow^1 t^{h}ji^2=sju^2 ŋwu^{2}kwar^1-nja^2=tja^1 ljọ^2-tjij^2
now [friend INDF]=because DEM=as cry-2=TOP INTRG-manner

Какая причина, что вы так убиваетесь из-за какого-то человека? (Solonin 1995: 38)

You are now crying now because of a friend; how can that be (=where is that manner)? (12K, 132.19.03)

The morpheme [illegible] *nioow*^{1} can also convey spatial meaning, but in a very unusual way. Usually, time morphemes establish a cognitive link with a spatial notion of being 'after' a perceived object from a subjective point of view, or 'after' a conceptualized object from an objective perspective. However, the spatial [illegible] *nioow*^{1} looks more like an abessive than a spatial postessive and is strictly limited in our corpus to sentences where the marked NP is [illegible] *ɣa*^{1} 'door' (61). This uniqueness suggests that this spatial meaning in Tangut reached a certain stage of fossilization. There is in Lizu and in Ersu a locative *ɲo* meaning 'outside' which must be cognate to Tangut [illegible] *nioow*^{1}. With such links, one could conjecture that the fossilized abessive and the postessive/causal are from different origins. Only new fieldwork can potentially provide the data to allow us to know what happened.

(61) [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]

njij^{2}	*·a-zow*^{2}	*ɣa*^{1}	*nioow*^{1}	*·wjɨ*^{2}*-lʰo*
weapon	PFV:UP-hold	door	ABE	PFV:OUT-go_out

'Держа меч, вышел за ворота.' (Solonin 1995: 35)

'He took the weapon and went out of the gates of the city.' (12K, 132.09.06)

3.6 Summary

As I did for the verbal template, I list in Table 20 the cognates found in the two preceding sections in the domains of the noun and locative phrases. As in Table 16, the data is from Honkasalo 2019 (Geshiza), Gates 2021 (Mazur Stau), Lai 2017 (Wobzi), Prins 2016 (Kyomkyo), Zhang 2013 (Yuexi Ersu), and Bai 2019 (Munya). The Lizu form is from Chirkova (2017).[71]

Some elements belong to the West Gyalrongic taxon (as the alternation between a CVC plain numeral and a monovocalic prefix), others only exist in Tangut and Geshiza.[72] The nominalizers are maybe the more striking commonality, even if these morphemes have lexical cognates elsewhere in Gyalrongic (see Subsection 2.2). The ressemblances between Geshiza, Tangut, and Ersu/Lizu are quite thought provoking – especially the metathesis described above for the numeral 'five' – although at the present state I can provide no further insights into these similarities.

[71]The Lizu forms for which cognates can be found in Ersu are not indicated. Lizu 'on' is *tɕʰòtɕʰó*.

[72]It is worth to recall that in all three languages this numeral prefix is a perfect homonym to the (non-orientational) interrogative preverb.

Tangut				Geshiza	Maz. Stau	Wobzi	Kyomkyo	Y. Ersu / (Lizu)	Munya
numerals	1	[illegible] [illegible]	*lew*^{1} *·a-*	*rəu* *æ-*	*ro* *æ-*	*râɣ* *ə̂-*	*kərek*	*tə(-)*	*to(-)*
	5	[illegible]	*ŋwə*^{1}	*ŋuæ* (< *w-)	*nɢvɛ*	*mŋá*	*kəmŋi*	*ŋuà*	*ŋa-*
nominalizers		[illegible] [illegible]	*-mjijr*^{2} *-sji*^{2}	*-me* *-sʰi*	 *-sə*				
case	LOC	[illegible]	*=ɣa*^{2}	*-ɣa*	*=ʁa*	*=ʁa*			
	SUPE	[illegible]	*=tśʰjaa*^{1}	*tɕʰa*	*tɕʰæ*	*=tʰa*		*=tɕʰo*	
	INTERE 'when'	[illegible]	*=kʰa*^{1}	*=kʰa* 'about'	$= k$ { INS / 'about' }		*wəkʰa*	*=kə* *(kʰæ)*	
	POSTE 'because' ABE	[illegible] [illegible] [illegible]	*nioow*^{1} { 'after' / 'because' / 'outside' }	*o* { 'after' / 'because' }	*ɲu* 'after', 'behind'			*ɲo* 'outside'	

Table 20: Overview of the cognacy with Tangut in macro-Qiangic (Noun & Loc. Phrases)

4 Conclusion

The amount of common morphosyntactic similarities shared between Tangut and Geshiza Horpa is too important to be the result of mere chance. Language contact between Tangut and Horpa languages is not a plausible explaination since the Horpa languages are separated greatly from Tangut by time and space. Instead, the similarities between Tangut and Horpa suggest a close genetic relationship. Geshiza is particularly close to Tangut, and offers insights into Tangut grammar.

Apart from the fact that Geshiza Horpa and Tangut share the same distribution of orientational preverbs – a fact which points to the loss of preverbs seen in Khroskyabs and core Gyalrong (dir. 4 and dir. 5, cf. Table 2) as a shared innovation – morphological, semantic, and distributional comparison of verbal templates allows an interpretation of the so-called orientational preverb 𗧓 *dja²-* as a perfective. Simultaneously, comparison gives new insights that could be valuable for the history of the macro-Qiangic preverb system. The analysis of optative preverbs in the light of Geshiza Horpa then gives the key to the process of derivation occurring in Tangut from Type-A preverbs to Type- B preverbs, with a binary aspectual basis for the stem where the same process derives the optative from the imperative and the interrogative from the indicative perfective.

Tangut and Geshiza behave very similarly with regards to negation, both systemically and semantically; there is an etymological link between two modal preverbs in Tangut and their cognates in West Gyalrongic.

My analysis of verb agreement, made across each paradigm, shows common retentions abandoned in Wobzi Khroskyabs, even if the parallelism of the agreement system of Wobzi seems at first sight closer to Tangut. The analysis of suffixes and enclitics also revealed many morphemes behaving in the same way, not only for one use, but a whole range of uses observed in Geshiza.

Some aspects of noun and locative phrases display cognacy, in form and function, which is unlikely to be due to mere chance. Geshiza Horpa is now the closest known relative to Tangut. This proximity has already led to new hypotheses in the understanding of nominalizers, and to new clues which will be followed soon in other works, which should rely on data from a larger range of Horpa languages.

Methodologically speaking, the present paper also exploits an original way of establishing cognacy by using cognatic homonymy as a comparative tool. Such a tool cannot ascertain cognacy by itself, but can still be a valuable member of what is known in judicial practice as 'a set of concordant items of evidence' (French 'faisceau d'indices'). More investigation will be needed to improve our understanding of the relationship between the Ersu/Lizu/Duoxu cluster and Gyalrongic, the proximity of the former with Tangut having already been pointed out, in the case of Duoxu, by Nishida.

5 Abbreviations

5.1 Glosses

Σ: verb stem; Σ[⊠]: stem 1 of transitive verbs; Σ[ʙ]: stem 2 of transitive verbs); Σ_1: stem 1 of intransitive verbs; Σ_2: stem 2 of intransitive verbs; 1: first person agreement; 2: second person agreement; 3: third person agreement; 1SG: first person singular pronoun; 2SG: second person singular pronoun; ABE: abessive; ADV: adverbializer; AGT: agentive (→ ergative); ANM: animate; ANTE: antessive; ANTIERG: antiergative; ASSOC: associative; ASSERT: assertive; AUTOB: autobenefactive; COMIT: comitative; COMP: comparative; CONC: concessive; CONJ: conjunction; DEM_{DS}: distal demonstrative; DEM_{PR}: proximal demonstrative; PFV:DIR: first series orientational preverb; EP: epistemic suffix; EXV: existential verb; EXCLAM: exclamative; HUM: humilific; IFR: inferential; INTERE: interessive; INV: inverse; LNK: linker; LV: light verb; MED: mediative (→ sensorial) MEDE: medessive; MIR: mirative; MARK.TR: marked transitive (→ inverse); MOD: modal; NACT: non actual; NAT: nativity and source suffix; NPST: non past; NSIT: new situation; OPT: optative; PART: particle; POSTE: postessive; POT: potential preverb; RED: reduplication; RHT: rhetorical; SENS: sensory evidential; SUBE: subessive; SUPE: superessive; SUPL: superlative; TEL: telic; TERM: terminative; TOP.C: contrastive topic; TOPON: toponym;

5.2 Other abbreviations

1) Documents: 12K: *The Twelve Kingdoms*; Cxz: *New collection on parental love and filial piety*; Avtṃsk: *Avataṃsaka Sūtra*;
2) Directions: centrf: centrifugal; centrp: centripetal; unspec.: unspecified; inw.: inwards; outw.: outwards; upstr.: upstream; downstr.: downstream; h.alt.: high altitude; l.alt: low altitude; obl.down: oblique down; obl.up: oblique up; N, S, W, E: North, South, West, East;
3) Others T. : Tangut; G. Geshiza; * : unattested; !: unexpected; ?: unknown; X: unknown pronunciation; TAME: tense, aspect, mood, evidentiality; tr.: transitive; intr.: intransitive; NP: Noun Phrase; VP: Verb Phrase

References

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.

Arakawa, Shintarō 荒川慎太郎. 2011. プリンストン大学所蔵西夏文華厳経巻七十七訳注 Purinsuton daigaku shozō seikabun Kegongyō kan 77 yakuchū [An annoted translation of the 77th volume of the Tangut version of Avataṃsaka Sūtra from the University of Princeton]. *Journal of Asian and African Studies* 81. 147–305.

Arakawa, Shintarō 荒川慎太郎. 2012. On the Tangut Verb Prefixes in “Tiansheng Code”. In Irina Popova (ed.), *Тангуты в Центральной Азии: сборник статей в честь 80-летия проф. Е.И.Кычанова [Tanguts in Central Asia: a collection of articles marking the 80th anniversary of Prof. E. I. Kychanov]*, 58–71. Moscow: Oriental Literature.

Arakawa, Shintarō 荒川慎太郎. 2014. On the Tangut Verb Phrase in *The sea of Meaning, Established by the Saints*. *Central Asiatic Journal* 57. 15–25.

Arakawa, Shintarō 荒川慎太郎. 2018. 西夏語の双数接尾辞について Seikago no sosu setsubishi ni tsuite [On the ‘dual’ suffix of Tangut]. In ユーラシア諸言語の多様性と動態ー20 号記念号ー追悼庄垣内正弘先生 *[Diversity and Dynamics of Eurasian Languages: The 20th Commemorative Volume Dedicated to Prof. Dr. Masahiro Shōgaito]* Contribution to the studies of Eurasian languages series 20, 69–83. The Consortium for the Studies of Eurasian Languages.

Bai, Junwei. 2019. *A Grammar of Munya*: James Cook University dissertation.

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

Beaudouin, Mathieu. 2018. *Etude du verbe tangoute dans le* [illegible] *Djij^{1} bo^{1}*. Inalco MA thesis.

Beaudouin, Mathieu. 2022. Tangut verb agreement: optional or not? *Linguistics of the Tibeto-Burman Area* 45(1).

Beaudouin, Mathieu. to appear. On locatives [illegible] *·u^{2}* and [illegible] *=kha^{1}*. In *Festschrift in Honor of Pr. M. Sofronov*, publication prochaine.

Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, Cambridge: Cambridge University Press.

Chirkova, Katia. 2014. The Duoxu Language and the Ersu-Lizu-Duoxu relationship. *Linguistics of the Tibeto-Burman Area* 37(1). 104–146.

Chirkova, Katia. 2017. Lizu (Ersu). In Graham Thurgood & Randy LaPolla (eds.), *The Sino-Tibetan Languages (second edition)*, chap. 42, 823–839. London: Routledge.

Chirkova, Katia & Dehe Wang. 2017. 甘洛爾蘇語動詞的體貌範疇 Gānluò ěrsū dòngcí de tǐmào fànchóu [Verbal aspect in Ganluo Ersu]. *Language and Linguistics* 18(3). 355–382.

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. Oxford: Blackwell (2nd edition).

DeLancey, Scott. 1997. Mirativity: The grammatical marking of unexpected information. *Linguistic Typology* 1. 33–52.

DeLancey, Scott. 2010. Towards a history of verb agreement in Tibeto-Burman. *Himalayan Linguistics* 9(1).

Downes, Alan. 2018. *How Does Tangut Work?*: Macquarie University dissertation.

Duan, Yuquan 段玉泉. 2015. 西夏語中的選擇連詞 Xīxiàyǔ zhōng de xuǎnzé liáncí mo². 語言研究 *Yǔyán yánjiù* 35(1). 123–126.

Gates, Jesse. 2017. Argument indexation in Stau from a cross-dialectal perspective. Academia online paper.

Gates, Jesse. 2021. *A Grammar of Mazur Stau*: École des Hautes Études en Sciences Sociales dissertation.

Gong, Hwang-cherng 龔煌城. 1999. 〈西夏語的緊母音及其起源〉 Xīxiàyǔ de jǐn mǔyīn jí qí qǐyuán [The tense vowels of Tangut and their origin]. *Journal of the Institute of History and Philology of the Academia Sinica]* 70(2). 531–558.

Gong, Hwang-cherng 龔煌城. 2001. 〈西夏語動詞的人稱呼應與音韻轉換〉 Xīxiàyǔ dòngcí de rénchēng hūyìng yǔ yīnyùn zhuǎnhuàn [Rhyme alternation and personal agreement in the Tangut verb]. *Language and Linguistics* 2(1). 21–67.

Gong, Hwang-cherng 龔煌城. 2006. 西夏語文研究論文集 *Xīxià yǔwén lùnwénjí [Collected papers on the Tangut language].* Taipei: Institute of Ethnology of the Academia Sinica.

Gong, Xun 龔勳. 2017. Verb stems in Tangut and their orthography. *Scripta* 9.

Hill, Nathan W. 2012. Mirativity does not exist: *ḥdug* in 'Lhasa' Tibetan and other suspects. *Linguistic Typology* 16(3). 389–433.

Honkasalo, Sami. 2019. *A grammar of Eastern Geshiza : A culturally anchored description*: University of Helsinki, Faculty of Arts dissertation.

Jacques, Guillaume. 2007. Textes tangoutes I. Nouveau recueil sur l'amour parental et la piété filiale. *München: Lincom Europa* .

Jacques, Guillaume. 2009. The Origin of Vowel Alternation in the Tangut Verb. *Language and Linguistics* 10(1). 17–28.

Jacques, Guillaume. 2014. *Esquisse de phonologie et de morphologie historique du tangoute.* Leiden: Brill.

Jacques, Guillaume. 2017. Sketch of Japhug. In Graham Thurgood & Randy LaPolla (eds.), *The Sino-Tibetan Languages (second edition)*, London: Routledge.

Jacques, Guillaume. 2021. *A Grammar of Japhug* Series in Diversity Linguistics. Language Science Press.

Jacques, Guillaume, Anton Antonov, Yunfan Lai & Nima Lobsang. 2014. Person marking in Stau. *Himalayan Linguistics* 13(1). 82–92.

Jacques, Guillaume, Anton Antonov, Yunfan Lai & Nima Lobsang. 2017. Stau. In Graham Thurgood & Randy LaPolla (eds.), *The Sino-Tibetan Languages (second edition)*, 597–613. London: Routledge.

Kepping, Ksenija. 1975. Subject and object agreement in the Tangut verb. *Linguistics of the Tibeto-Burman Area* 2(2). 219–231.

Kepping, Ksenija. 1985. *Тангутский язык: Морфология [The Tangut Language: Morphology].* Moscow: Nauka.

Lai, Yunfan. 2017. *Grammaire du khroskyabs de Wobzi*: Paris: Université Sorbonne Paris Cité dissertation.

Lai, Yunfan. 2021. Betrayal through obedience: on the history of the unusual inflectional chain in Siyuewu Khroskyabs. *Linguistic Typology* 25(1). 79–122.

Lai, Yunfan, Jesse Gates, Xun Gong & Guillaume Jacques. 2020. Tangut as a West Gyalrongic language. *Folia Linguistica Historica* 41(1). 171–203.

LaPolla, Randy. 1992. 'Anti-ergative' marking in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 15(1). 1–9.

LaPolla, Randy & Chenglong Huang. 2003. *A Grammar of Qiang: With Annotated Texts and Glossary.* Berlin: Mouton de Gruyter.

Laufer, Berthold. 1916. The Si-hia Language. A Study in Indo-Chinese Philology. *T'oung Pao* 17. 1–126.

Li, Fanwen 李範文. 2008. 夏漢字典 *Xià Hàn zìdiǎn [A Tangut-Chinese Dictionary].* Beijing: Social Sciences Academic Press.

Lin, Ying-chin & Kesih An. 1992. 《類林》西夏譯本之研究 Lèilín Xīxià yìběn zhi yánjiù [Study of the Tangut translation of Leilin]. 大陸雜誌 *Dàlù zázhì* 84(5). 1–8.

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman*, vol. 135. Berkeley and Los Angeles: University of California press.

Morisse, Georges. 1904. Contribution préliminaire à l'étude de l'écriture et de la langue Si-Hia. In *Mémoires présentés par divers savants à l'Académie des inscriptions et belles-lettres de l'Institut de France. Première série: Sujets divers d'érudition*, vol. 11 2, 313–379. Imprimerie nationale.

Nishida, Tatsuo 西田龍雄. 1973. 多續譯語の研究 : 新言語トス語の構造と系統 *Tosu yakugo no kenkyū: shin gengo tosugo no kōzō to keitō [A study of the Tosu-Chinese vocabulary, Tosu i-yu: the structure and system of Tosu, a new language].* Kyoto: Shokado.

Nishida, Tatsuo 西田龍雄. 1976. Hsihsia, Tosu and Lolo-Burmese Languages. *Studia Phonologica* 10. 1–15.

Nishida, Tatsuo 西田龍雄. 2004. 西夏語文法新探 Xīxiàyǔ wénfǎ xīntàn. In Lin Ying chin 林英津 (ed.), 漢藏語研究－龔煌城先生七秩壽慶論文集 *Hànzàngyǔ yánjiù – Gong Hwang-cherng xiānshēng qīzhì shòuqìng lùnwénjí [Sino-Tibetan Studies – collected paper for the anniversary of Prof. Gong Hwang-cherng]*, 353–381. Taipei: Institute of Linguistics of the Academia Sinica.

Prins, Marielle. 2016. *A Grammar of rGyalrong, Jiǎomùzú (Kyom-kyo) Dialects. A Web of Relations.* Brill.

Shi, Jinbo 史金波, Huang Zhenhua 黄振華 & Nie Hongyin 聶鴻音. 1993. 類林研究 *Lèilín yánjiù [Study of Leilin].* Ningxia Renmin Press.

Solonin, Kirill. 1995. [illegible] / *Двенадцать царств [The Twelve Kingdoms].* Saint Petersburg: Petersburg Oriental Studies.

Sun, Jackson T.-S. & Qianzi Tian. 2013. 霍爾語格西話動詞對協 Huòéryǔ Géxīhuà dòngcí duìxié [Verb Agreement in Gexi Horpa]. *Bulletin of Chinese Linguistics* 7(2). 203–223.

Sun, Jackson T.-S. 孫天心. 2000a. Parallelisms in the Verb Morphology of Sidaba rGyalrong and Lavrung in rGyalrongic. *Language and Linguistics* 1(1). 161–190.

Sun, Jackson T.-S. 孫天心. 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. 孫天心. 2003. Caodeng rgyalrong. In Graham Thurgood & Randy LaPolla (eds.), *The Sino-Tibetan Languages*, 490–502. London: Routledge.

Sun, Jackson T.-S. 孫天心. 2007a. The irrealis category in rGyalrong. *Language and Linguistics* 8(3). 797–819.

Sun, Jackson T.-S. 孫天心. 2007b. Morphological causative formation in Shangzhai Horpa. *Bulletin of Chinese Linguistics* 2(1). 211–231.

Sun, Jackson T.-S. 孫天心. 2019. The Ancestry of Horpa: Further Morphological Evidence. In *The Ancestry of the Languages and Peoples of China* (Monograph Series 29), 24–43. Journal of Chinese Linguistics.

Tian, Qianzi & Jackson T.-S. Sun. 2019. 格西霍爾語動詞的時與體 Géxī Huòéryǔ dòngcí de shí yú tǐ [On tense and aspect in the Gexi Horpa verb]. *Language and Linguistics* 20(3). 451–468.

Tournadre, Nicolas & Sangda Dorje. 2003. *Manuel de tibétain standard*. Paris: L'Asiathèque.

Tournadre, Nicolas & Randy LaPolla. 2014. Toward a new approach to evidentiality. *Linguistics of the Tibeto-Burman Area* 37(2). 240–263.

Tunzhi, Sonam Lhundrop. 2019. *Outline of Bra'go Variety of rTa'u (Horpa)*: La Trobe University dissertation.

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*: Inalco dissertation.

Zhang, Sihong. 2013. *A reference grammar of Ersu: a Tibeto-Burman language of China*: James Cook University dissertation.

Zhang, Yongfu 張永富. Ms. 西夏语第一、二人称双数后缀与人称范畴再探讨 Xīxiàyǔ dì yī, dì èr rénchēng shuāngshù hòuzhuì yǔ rénchēng fànchóu zài tàntǎo [New inquiries on the first/second person dual suffixes and the agreement in Tangut].