PART 2. – DESCRIPTIVE APPROACH TO TIBETIC LANGUAGES
4. Proto-Tibetic and Old Tibetan

Proto-Tibetic is the common ancestor of Tibetic languages. In the literature, it is traditionally referred to as 'Proto-Tibetan', but in order to be consistent with our terminology, we will use Proto-Tibetic (hereafter PT). As a hypothetical language, PT can be reconstructed in two ways: by comparing the various modern languages and dialects; and by comparing them with Old Tibetan and Classical Tibetan (see Chapters 5 and 6).

Higher levels of Proto-languages for the ST macrofamily have been partly reconstructed. Many Proto-Tibeto-Burman roots have been reconstructed by various authors such as Benedict (1972), Matisoff (2003), and Thurgood and LaPolla (2003).

Some authors such as R. Sprigg (1972), Li Fang-Kuei (1987, 1993), G. Jacques (2004a and b) have proposed various reconstructed forms for PT but so far there is no systematic study of the common ancestor of the Tibetic languages.

Reconstructed PT forms are often similar or identical to the orthography of Classical Literary Tibetan. The reality is, of course, somewhat more complex. As Sprigg (1972: 556) points out "Shafer would have said that we already know what Proto-Tibetan looks like: it is embalmed in the orthographic forms of Written Tibetan." He adds though that "none of the dictionaries gives a reliable picture of the phonological structure of Written Tibetan during a given État de langue."

Sprigg further states that, for establishing PT forms, it is better "not to [accept] Tibetan orthographic forms without first testing them against constructions based on comparing contemporary spoken-dialect forms."

Whether Proto-Tibetic was a homogenous language or was a hybrid is a debatable issue. As we have seen in the previous chapter, during their history, Tibetic languages were sometimes in contact with both ST languages and many languages belonging to other genetic groups, such as Indo-Iranian, Mongolic or Turkic. However, even if we accept that some of the contact languages could have left a significant lexical and structural impact on neighboring Tibetic languages, it is clear that the great proximity of the modern languages points toward the existence of a common PT. Of course, this
does not exclude the fact that some processes of convergence could have taken place and contributed to the similarity of the modern languages.

**4.1. Methodology for the reconstruction**

As previously rementioned, the reconstruction of PT can be achieved essentially by comparing all the modern Tibetic languages and dialects with each other and then, together, with Classical Tibetan. Other sources can also be used. For example, many words found in Bodic, rGyalrongic and Qiangic languages have been borrowed from Tibetan at various stages of their history, and often they have preserved archaic forms.

Historical annals and chronicles in Chinese, Mongolian, Tangut or other languages mentioning Tibetan names or titles may also supply valuable information about the pronunciation of Old Tibetan or Tibetan spoken during the Middle Ages. The Tibetan orthography of foreign words from Sanskrit, Chinese, Mongolian, etc., also provides clues about ancient pronunciation.

The existence of a millennium-old literary tradition in Tibet is very useful for the reconstruction of PT. Many written documents from the eighth century onwards have been found and preserved. The reconstruction is also facilitated by the fact that Tibetans have always used a phonetic alphabet (or more precisely an *alphasyllabary*) to transcribe their language (see Chapter 5). Had the Tibetans used pictographic or ideographic systems such as Chinese, Naxi, Tangut, Yi, etc., the reconstruction would have been very difficult.

**4.2. A dialect of Old Tibetan used as a basis for the written language**

In written languages, which use a phonetic alphabet (or an alphasyllabary), the ancient orthography usually reflects the pronunciation of a given dialect.

There is thus little doubt that the written language was based on a given spoken dialect of Old Tibetan. However, it is difficult to have a precise idea as to which dialect served as the basis for the written language and there is so far no consensus on this issue among scholars.

Many scholars first thought that the Old and Classical orthographies were largely artificial because of the complexity of their syllable structures (see 5.2). During the last
three decades, however, the data collected from modern Tibetic languages such as Balti, Purik, Ladak, and Choča-ngača have shown that these languages preserve complex syllabic structures and that the pronunciation of certain words remains very similar to the corresponding transcriptions that we can infer from Classical orthography.

If we look at the historical background, it is clear that the emergence of a written language is related to the rise of the Tibetan Empire in the seventh century. Although the earliest Tibetan written document known at present dates from 764, Tibetan historical tradition states that the creation of the Tibetan alphabet occurred during the reign of King Songtsan Gampo སྲོང་བཙན་སྒམ་པོ (SRONG.BTSAN.SGAM.PO) in the seventh century. Given the empire’s extensive military conquests, it is likely that the script was already in use at the beginning of the Tibet’s military expansion, i.e. in the first part of the seventh century, and that it facilitated the administration of the empire, including communication between the various regions under its rule.

Songtsan Gampo’s father, Namri Songtsan གནམ་རི་སྲོང་བཙན་ (GNAM.RI SRONG.BTSAN), reigned at end of the sixth century, ruling over a small kingdom in the Chonggyä and Yarlung valleys. The capital of his kingdom was Chingwa Tagtseཕྱིང་བ་སྟག་རྩེ (PHYING.BA STAG.RTSE) in the Chonggyä valley. His son, Songtsan Gampo gradually moved the capital to Lhasa, less than two hundred kilometers to the north. Lhasa became the capital of a powerful empire until the fall of that empire in the mid ninth century. All of the tombs of the Tibetan emperors are located in the necropolis of Chonggyä. This history clearly demonstrates that Lhasa, as well as the Yarlung Shampo and Chonggyä valleys, were central places for the Tibetan Empire.

The ethnic diversity on the Plateau at the end of the sixth and the beginning of the seventh century is not known in detail. At that time, various ethnic groups had their own kingdoms, such as Zhangzhung in Ngari area, ’Azha (or Tuyuhun) in the Kokonor area, the Qiang in eastern Tibet or the Sumpa, and did not speak Tibetan dialects.

1. However, the Old Tibetan annals indicate that the Tibetan emperor (BTSAN.PO) was not settled permanently in Lhasa and had a mobile court.
Thus, the language that served as the basis for the first transcriptions of written Tibetan was presumably spoken in the Yarlung valley and Lhasa areas around the seventh century. This language, which was quite different from its modern form (especially its phonology), then spread to other areas of the Tibetan Plateau along with the expansion of the Tibetan Empire. The degree of dialectal diversity within Central Tibet in the seventh century is not easy to establish, but it is very plausible that dialectal diversification had already taken place.

If these hypotheses are correct, we can say that Central Tibetan and Dzongkha have undergone a massive phonological evolution, whereas languages spoken at the periphery of the Empire, such as Amdo, Balti, Purik, Ladaks or Choča-ngača have retained more archaic features of this original ‘Yarlung language’.

4.3. Archaic reflexes found in some modern languages

As mentioned above, some Modern Tibetic languages have preserved reflexes which are very similar or identical to Classical spellings.

Examples of initial consonant clusters

The initial consonant cluster LT has been well preserved in Balti, Purik and Ladaks dialects, as well as some Amdo dialects neighboring the rGyalrong area (the modern pronunciation appears in oblique bar):

\[\text{ལྟ་} \text{LT} /\text{l}t\text{a}/ \text{‘to look at’ (Ba, Pur),} \text{ལྟོགས་} \text{LT} /\text{l}t\text{oks}/ \text{‘to be hungry’ (Ba, Pur, La).}\]

The initial consonant clusters RG and RGY have been well preserved in Balti, Purik and some conservative Amdo dialects:

\[\text{རྒད་པོ} \text{RG} /\text{rgatpo}/ \text{‘old man’ (Am, Bal, Pur),} \text{རྒྱལ་པོ} \text{RG} /\text{rgyalpo}/ \text{‘king’ (Bal, Pur),} \text{/rgawo}/ \text{(Am).}\]

The initial consonant cluster SR has also been well preserved in Balti and Purik:

\[\text{སྲོག་} \text{SR} /\text{stroq}/ \text{‘life’ (Bal, Pur),} \text{སྲུང་} \text{SR} /\text{strung}/ \text{‘to keep, protect’ (Bal, Pur).}\]

2. About the notation of preinitials see chapter 7. We use here a phonological notation. From a strictly phonetic point of view, the preinitial T is pronounced as unvoiced. It is sometimes noted as a fricative [ɬ] or as voiceless [ʃ].
The initial consonant clusters KHR and GR have disappeared in most languages except for Balti, Purik and the archaic Kham dialects of Rongdrak (sProsnang) and Phongpa.

The initial consonant clusters SKR and SGR have usually not been preserved. However, they are attested in Purik:

The initial consonant clusters PHY and BY are still attested in some western and southern languages such as Balti, Purik or Choča-ngača:

The initial consonant clusters PR and BR are still attested in some western and southern languages such as Balti, Purik, Kyirong or Choča-ngača:

The preinitial consonant M has disappeared in nearly all the modern languages but is still present in some archaic Amdo dialects.

The preinitial consonant * has not only been preserved in many eastern dialects of Kham and Amdo, but also in To Ngari dialects.
**Examples of final consonants**

Some final consonants, such as $S$, $L$, and $D$, which have generally disappeared, are still heard in certain languages or dialects.

For example, the final $S$ is still present in the western languages of Balti, Purik and Ladaks: $NaN/AS$ 'barley' /nas/ (Ba, Pur), $LAN/AS$ 'work' /las/ (Ba, La, Pur).

The final consonant $L$ is still present in many western and southern languages, such as Balti, Purik, Ladaks, Choča-ngača, Sherpa: $BN/LL$ 'wool' /bal/ (Bal, Pur, Lad, Cho), /pal/ (Sh); $CN/KHAL$ 'score' (Bal, Pur, Lad, Cho, Sh), $SN/SLPA$ 'frog' /slpa/ (Ba, Pur).

The final $D$ is still present in the western languages of Balti, Purik, Ladaks and Zanhar as well as some conservative Amdo dialects and Choča-ngača: $RMD-PO/gatpo/ 'old man' (Am, Bal, Pur), /gatpo/ (Cho).

The final consonantic cluster $GS$ has disappeared in nearly all the modern languages except for Ladaks, Purik and some Balti dialects:

$ECL/GS$ 'iron' /čaks/ or /čaqs/ (La, Pur), $MFIN/PHYG/S'cattle' /phuks/ 'goat and sheep' (Bal), $EL-DWAGS/Ladakh' /ladaks/ (La).

The second suffix $D$ has left some traces in the Classical orthography of bound morphemes such as $TO$ 'the final particle', and the terminative case $TU$. For example $NYR-TO$ 'has changed' instead of 'NYR-DO', $NYR-TU$ 'entirely' instead of 'NYR-DU'. See Tournadre & Dorje (1998, 2003: 468).

To our knowledge, segmental traces of the second suffix $D$ are not attested in the modern languages but some rare suprasegmental traces (tone changes) are found (see 7.3.1).

No single modern language has managed to preserve all the consonant clusters found in Old Tibetan and in the Classical orthography. However, if we put together the most conservative languages and dialects, such as Balti, Purik and Ladaks in the western area, Amdo in the east, Choča-ngača in the south and some archaic Kham dialects (as if they were the pieces of a linguistic jigsaw puzzle), we find that nearly all the consonantic clusters have been preserved.
4.4. Characteristic features of Proto-Tibetic

We will now briefly examine various characteristic features of PT. But first, let’s emphasize the necessity of distinguishing between PT, the common ancestor of the family and Pre-Tibetic, a stage of the language, which immediately preceded the emergence of the Proto-Tibetic language. Note that some authors actually use the term Proto-Tibetan to refer to forms that we call Pre-Tibetic (cf. Jacques 2004b; Sprigg 1972).

A few authors, such as Uray (1953), Li Fang-kuei (1987, 1993), Coblin (1976), Beyer (1996), and Jacques (2004a-b) have proposed a reconstruction for some PT roots.3

The main phonological features that characterize PT are:

▪ the preservation of the prefixes inherited from Proto-TB;
▪ the palatalisation of dental and alveolar before y;
▪ the change from lateral to dental after m;
▪ the emergence of distinctive aspirated initial plosives.

Preservation of prefixes

The numerous prefixes of Proto-TB are still clearly pronounced, most probably with an epenthetic vowel in PT.

On this topic, Matisoff (2003: 97) gives the following comment:

“We cannot be sure from the WT [Written Tibetan] orthography how the Tibetan combinations of prefixes and initials were pronounced in ancient times; but judging by their excellent state of preservation in WT, we may surmise that they were pronounced with a following unstressed schwa-type vowel,4 which served to protect them from too close contact with the root initial. That is most words with prefixes must have been pronounced sesquisyllabicity.” Matisoff (2003: 97)

It is not excluded that in some cases, prefixes were in fact the result of a metathesis as suggested by Zeisler (pers. comm. 2020).

3. These authors use the term Proto-Tibetan.
4. The term ‘schwa’ which denotes the vowel [ǝ] used in general linguistics is borrowed from Hebrew classical terminology and corresponds more or less to the French “e muet.”
The main prefixes found in PT are: *s(ǝ), *d(ǝ)/g(ǝ), *m(ǝ) and *b(ǝ).

The words for numbers give a good illustration of this phenomenon.

* g(ǝ) - *gyi 'one', *g(ǝ)-nyis 'two', *g(ǝ)-sum 'three', *b(ǝ)-zi 'four', *l(ǝ)-na 'five',
* d(ǝ)-ruk 'six', *b(ǝ)-dun 'seven', *b(ǝ)-rgyat 'eight', *d(ǝ)-gu 'nine', *b(ǝ)-teu 'ten'.

Concerning the numbers '6', '7', '8', it is possible that Pre-Tibetic forms had two prefixes: d(ǝ)-k(ǝ)-ruk 'six', *b(ǝ)-d(ǝ)-nis 'seven', *b(ǝ)-r(ǝ)-g yat 'eight'.

The presence of high tones in southern Kham for 'six' /ʈɔɁ/ could be a trace of the voiceless prefix k(ǝ). In written Burmese, the form for 'six' is /khruk/.

The prefix 's' is used for animals and body parts.

* s(ǝ)-dik-pa 'scorpion', *s(ǝ)-bal 'frog', *s(ǝ)-tak 'tiger', * s(ǝ)-b-rul 'snake', * s(ǝ)-pr a 'monkey', * s(ǝ)-kra 'hair', * s(ǝ)-nyiŋ 'heart', s(ǝ)-na 'nose'.

However, other prefixes such as 'd', 'm' and 'r' are also used for the body parts:

* d(ǝ)-myik 'eye', * m(ǝ)-go 'head', * r(ǝ)-na 'ear'. For a discussion about the status of the morphological prefix /d/, see Jacques (2001, 2008) and Hill (2011a).

Concerning the prefixes * d(ǝ) and * g(ǝ), Li Fang-kuei (1933) noticed that they occur in a complementary distribution. The dental prefix * d(ǝ) occurs before the labials and the velar (g) whereas the velar prefix * g(ǝ) occurs before dentals (t, d, n) and the lateral (l).

**Palatalisation of dentals and alveolars before y**

Palatalisation is one of the main features of PT. The combinations * ty, * ly, * sy, * tsy were not palatalised in Pre-Tibetic (see Jacques 2004b and Gong 1977, for these reconstructions), but all these combinations have undergone a palatalization in PT, which is recorded in the orthography of Literary Tibetan. All the modern languages and dialects have now developed reflexes of these palatalised forms.

5. Sagart and Jacques propose the reconstruction * s-m-rul for some earlier stage (pers. comm., also compare Hill 2011: 448).
In the following examples, we propose a reconstruction for Pre-Tibetic forms, then provide the reconstructed form for PT and the classical orthography, and in some cases the archaic orthography.

- \*ty > tɕ; \*dy > dʑ
- \*g(ǝ)-tyik 'one' > PT: \*g(ǝ)-tɕ(h)ik > OT: GCIG/GCHIG ལེགས་ / ལེགས་
- \*tye 'big' > PT: \*tɕ(h)e > OT: CHE ཆེ་
- \*b(ǝ)-tyu 'ten' > PT: \*b(ǝ)-tɕu > OT: BCU/BCHU མ་ / མ་
- \*tyi 'what' > PT: \*tɕ(h)i > OT: CI/CHI མ་ / མ་

Beyer (1992: 78) has proposed the following evolution for 'flea' in PT, postulating the metathesis: zli > lzi.
- \*z-li 'flea' > \*lzi > \*ldi > PT-ldyi > LJI ལིའི / ལིའི

Proto-Tibeto-Burman had non-palatalised forms: \*g(ǝ)-tyik 'one', \*b(ǝ)-tyu 'ten'.

In many Tibeto-Himalayan, and even in Bodish languages closely related to Tibetan, we do not find palatalised forms of \*t+y. See for example Bake (Basum lake) /ti/ 'what', /tɨʔ/ 'one' which reflects a stage close to Pre-Tibetic.

- \*sy > c
- \*sya 'flesh' > PT: ca > CT: SHA མ་
  < syes 'know' > PT: ces > CT: SHES མ་
  < sying 'wood' > PT: cin > CT: SHING མ་

Many Bodish languages such as Tamangic (Tamang, Gurung, etc.) and East Bodish (Kurtö, Bumthang, etc.) have not undergone this change.

- \*tsy > tɕ
- \*b(ǝ)-tsyat 'to cut' (past stem) > PT: \*b(ǝ)-tɕat > CT: BCAD བཅད་
- \*m(ǝ)-tsyl-ma 'spittle' > PT: \*m(ǝ)-tɕ(h)il-ma > CT: MCHIL.MA མཆིལ་མ་
- \*m(ǝ)-tsin-pa 'liver' > PT: \*m(ǝ)-tɕ(h)in-pa, CT: MCHIN.PA མཆིན་པ་

6. Both 'ji-ba and lji-ba are found in CT. The TDCM gives two slightly different meanings for 'ji-ba and lji-ba, but they have certainly a common etymology. See also Beyer (1992: 78).
For some words, we have to postulate a metathesis. That is the case for the term *LCAGS ‘iron’. 7

\[
\begin{align*}
&\text{x(s)-lak(s) }^\prime \text{‘iron’ } > \text{x-l-sak(s)} > \text{x-l-tesak(s)} > \text{PT: } \text{x-teaks} > \text{CT: } \text{LCAGS } ლྕགས  \\
&\text{x ly } > \text{x z}.
\end{align*}
\]

Among the important innovations of Proto-Tibetic is the palatalization of the lateral /l/ in front of y (see Jacques 2004a). This sound law has been dubbed ‘Benedict’s law’ by Hill (2011: 445). 8

\[
\begin{align*}
&\text{x b(ǝ)-lyi }^\prime \text{‘four’ } > \text{PT: } \text{x b(ǝ)-zi} > \text{CT: BZHI भी} \\
&\text{x zingen }^\prime \text{‘field’ } > \text{PT: } \text{x zin} > \text{CT ZHING झी}.
\end{align*}
\]

The lateral of the sequence *bli is also preserved in many other Bodish or even ST languages, such as Kurtö, Tshona (mtsho-sn), rGyalrong, and Old Chinese (see Jacques, 2004).

The change from lateral to dental after m

The change from *ml to *md that occurs in PT and its reflex is found in all the modern Tibetic languages.

Thus PTB (Matisoff) *b/m-la ’arrow’ > PT: *mda > CT: MDA म्दा.

Some Bodish languages closely related to Tibetic did not undergo this mutation. Cf. Kheng (see also Michailovsky and Mazaudon, 1994).

Emergence of distinctive aspirated initial plosives

Another characteristic of PT is the emergence of distinctive aspirations for initial plosives, as was shown by Li Fang-kuei (1993). In Old Tibetan, the status of the aspiration gradually became phonemic. The fact that the aspiration of initial plosives appeared at a relatively late stage can easily be proven by the fluctuation found in the

---

7. Matisoff (2003: 317) proposes for Proto-TB the reconstruction *l-tesak. This reconstruction is perfect for Tibetan. However it does not match some data found in Bodish languages such as Kurtoep /läʔ/ and Bumthap /łeʔ/ (Michailovsky and Mazaudon, 1994) or Bake /hrʔ/.

8. This phenomenon also occurred in some romance languages such as Spanish: lla > /ʎ/ as in llamar, llorar, which have derived into [ʃ] in Portuguese: chamar, chorar (and in Argentina Spanish).
orthography of OT between aspirated and non aspirated consonants (Bialek 2018a ou b?).

Ex: THE /CHI/ ‘what’, etc. CU-DMYIG ཆུ་མྱིག་ /CU-DMYIG ཆུ་མྱིག་ /CHU-DMYIG ཆུ་མྱིག་ ‘(water) spring’.

Fluctuation of aspirated versus non-aspirated consonants in the orthography of OT is also attested for second syllables and for non-initial consonants. This type of phenomenon is still attested in some Amdo dialects with the aspirated initial consonants of second syllables which are sometimes pronounced as unaspirated.

Ex: PHYIN-CHAD/PHYIN-CAD གྱིན་ཆད་ /གྱིན་ཅད་ ‘from now on’
Ex: GCIG གཅིག་ /GCHIG གཆིག་ ‘one’

Cases of non-coincidence between PT and Classical Tibetan

For some words, the classical orthography does not correspond to PT. For example for the word ཉིག ‘eye’, a few dialects of Amdo have a form such as /ŋnyax/ or /mnyax/. The PT form could be reconstructed *d(ɔ)-myik. Fortunately, an archaic orthography ཆུ་མྱིག་ [ dimeyik] is attested in some old documents.

But this is not always the case. For the word བོམ་པོ ‘big, thick’ (for rope), the reconstruction based on some dialects of To and Amdo (Ngaba [rNgawa]) which yield /sboompo/ should be PT: *sbrom-po (*sboompo). This form is not attested in Literary Tibetan but the form /domp/ is found in Spiti. This is a perfect reflex of the reconstructed form *sbrom-po *sboompo. An astonishing confirmation of this hypothesis is found in Purik where the word for ‘fat (person)’ is /brompo/, which is also a cognate meaning ‘thick in circumference’.

On the basis of many languages (To, Balti, Ladak, Sherpa, Gyalsumdo, Lhoke, etc.), we should reconstruct for ‘flower’ PT *mentok, whereas Classical Tibetan has གཉིས་ ག་ཁོང་ ME,TOG. Fortunately, the form MENTOG གཉིས་ ག་ཁོང་ is also attested in Old Literary Tibetan (see Hill 2007: 480 note 8).

Bidmeier gives a similar illustration for Balti:

“In a number of cases the comparative evidence of the dialects does not lead back directly to the Written Tibetan etymological equivalent. Either the evidence leads to a
form previous to the Written Tibetan etymological equivalent, making the Written Tibetan equivalent thus not the ‘ancestor’ but simply an ‘older relative’, or else we have to accept that certain morphonological or grammatical changes took place within individual dialects. To give an example, we have GYANG ‘wall’ in Written Tibetan with regular etymological correspondences in all dialect groups of Tibetan, but in Balti we find rgyang ‘wall’. In such a case we have either to start from a Common Tibetan [PT] *RGYANG of which Written Tibetan gyang is a later offshoot, comparable to Purik guang, or else we may assume that there was an internal Balti development from Common Tibetan [PT] *GYANG to Balti RGYANG by prefixing r-. We would then have to explain the reason for this prefixing.

A last example of discrepancy between Classical Tibetan and PT is the word for ‘silver’. The word for ‘silver’ is DNGUL in Classical Tibetan and most of the Tibetic languages have a reflex of this form. However, a couple of languages and dialects of Western Tibet and Baltistan have forms as /xmul/ or /mūl/, so we could reconstruct a PT form: *dmul (※DNGUL), which would have undergone a change from labial /m/ to velar /ŋ/ in most Tibet languages. Thus PT *dmul > dŋul.

We should keep in mind that due to contact with other languages as well as reanalysis and analogies, it is quite possible that some of the forms that look ‘archaic’ do not necessarily point toward a PT origin and may well be later developments.

In some cases, one could be tempted to reconstruct the PT form, however a phonological innovative rule may provide a better account for the phenomenon. For example, the word KHANG.PA ‘house’ is sometimes prenasalised in some dialects of Amdo, Khöpokhok, Minyag Kham and Baima (Zhang 1997) and thus we could have proposed to reconstruct PT: *khangpa. But in this case, the phonological environment may provide a better explanation. A few words with this type of prenasalisation (e.g. TSHANG ‘nest’, PHRENG ‘beads’, TSHANG.MA ‘all’) attested in these dialects originally had an aspirated obstruent initial with -ng final which was omitted and caused a prenasalisation instead (except Amdo: Machu, mGolog). Thus, we have to be careful

9. Website of the CTDT: www.isw.unibe.ch/tibet/CDTD.htm. R. Bielmeier’s use of ‘Common Tibetan’ here refers to a reconstructed form of language equivalent to our Proto-Tibetic. It should not be confused with our definition of the Modern Common Tibetan.
with the reconstruction of PT forms when they are based on a single region of the
Tibetic area and never rule out the possibility of local areal innovations.

**Grammatical features**

It is beyond the scope of this book to present a reconstruction of the PT grammar,
which has not been described so far, but we can provide some of its essential features.
The verbal morphology (see Chapter 6) inherited from proto-TB has been replaced in
PT by a system of auxiliary verbs used with nominalized forms of the verb (see DeLancey,
2011b). The reconstruction of a pronominalization system in proto-TB has generated a
hot debate. Whether or not it existed in TB, there is no trace of it in OT.

This system is already clearly present in Old Literary Tibetan in the first attested
documents. In modern languages, auxiliary verbs have become verb suffixes and convey
a number of tense-aspect, evidential and epistemic meanings. However the various
verb stems (present, past, future and imperative) as well as the causative derivation in
*r*s have been relatively well preserved in Classical Tibetan and to a certain extent in
many modern Tibetic languages. Negation in all modern Tibetic languages is always
marked by reflexes of *MA* and *MI* or *MYI* and thus PT negation should be
reconstructed as *ma* and *myi.

There are no traces in PT of the verb agreement found in many ‘pronominalised
languages’ such as Kiranti, West Himalayish, Qiangic or rGyalrongic (see van Driem
2001; LaPolla 1992; DeLancey 2010, 2011b). Thus, modern Tibetic languages do not
exhibit any verb affixes (prefixes or suffixes) related to personal or directional marking.
Instead, Tibetic languages have developed a nominalization strategy associated with

Classical Tibetan has a system of ten nominal cases (see Tournadre 2010; Hill
2012a). It is difficult to have a clear picture of the original PT nominal case system. Most
of the modern languages have preserved to some extent the nominal cases inherited from
Classical Tibetan but in many languages, the number of cases is reduced.

All the Tibetic languages exhibit a form of nominal ergative marking, with one or
two exceptions such as Baima, however the modern languages differ in the type of
ergativity (see Chapter 8). Systems of classifiers, which are found in many ST languages,
are not found in CT nor in the modern languages, although a few rare classifiers are encountered. Hence, it is probable that there was not any system of classifiers in PT.
5. The Tibetan script

Classical Tibetan is closely related to modern Tibetic languages and its orthography allows us to reconstruct many ancient forms and understand the evolution of these modern languages. The first forms of written Tibetan (for the early written sources in Tibetan, see 6.3) are intimately linked to the elaboration of a specific Indic script around the seventh century A.D. No form of Old Tibetan is attested in any other script.

The Tibetan script has also been used to transcribe some other Tibeto-Burman languages such as Nam, Zhangzhung, rGyalrong and Indic languages such as Sanskrit.

We will briefly present below the script and its history as well as some important features of Tibetan syllable structure. Further information about the Tibetan alphabet can be found in Tournadre & Sangda Dorje (2003: 29-52).

5.1. The script and its origin

According to Tibetan tradition, the Tibetan script was created in the seventh century during the reign of King སྲོང་བཙན་སྒམ་པོ་ Songtsän Gampo by one of his ministers, ཐོན་མི་སམ་བྷོ་ཊ Thönmi Sambhoṭa. This minister was sent to India and is purported to have created not only the Tibetan alphabet but also to have written eight grammatical treatises (six of which were subsequently lost) as well as translations of various Buddhist sutras. There is a great deal of uncertainty about the historicity of Thonmi Sambhota and his composition of two grammatical treatises that are still well known to Tibetans today: the SUM.CU.PA (ཐུམ་ཅུ་པ་) and the RTAGS.KYT JUG.PA (ཧྲིག་ཏུ་རྒྱུད་པ་). First, the name of this minister is not mentioned even once in the Dunhuang documents, in which all the important ministers of Songtsän Gampo are

---

1. In a marginal way, transcriptions of the pronunciation of some rGyalrongic languages with the Tibetan script were conducted in 18th century. The documents are named Xifan Yyun (Chinese-Tibetan vocabulary), Cf. Nishida 1973 and Nishida & Sun 1990.

2. For instance cf. Miller 1976, 1993; Róna-Tas 1985: 183-303; Zeisler 2006b. Kesang Gyurmé is one of the few Tibetan linguists and grammarians who take a critical approach to the tradition. In his view, Thonmi Sambhota may not have written the SUM-RTAGS at all, or at least not its present version (Kesang Gyurmé. pers. comm.). Such an opinion is not likely to be accepted by most Tibetan scholars, since it contradicts tradition. Thonmi is for most religious people a sacred figure and his work cannot be questioned. There is even more uncertainty about the six lost treatises.
listed. Second, there is linguistic evidence suggesting that the ‘current’ versions of the SUM.CU.PA and the RTAGS-KYI.JUG.PA treatises were composed not in the seventh century but in the ninth century, or perhaps even later. For example, the grammatical rules of gender agreement explained in the text correspond to the rules of the second orthographic reform which took place during the reign of another Tibetan king, Thri Tsukdetsän also known as Rälpčän.

What is clear, however, is that the Tibetan script is directly derived from a script used in the Gupta Empire of Northern India. The various types of Gupta script are themselves considered as late forms of Brāhmi script, the ancestor of all modern Indic scripts.

It is hard to tell precisely which variant of Gupta (or closely related scripts such as the Siddham Khotanese and Śāradā alphabets) the Tibetans used as a model to develop their own script. Below is a chart comparing the shape of Gupta letters with Tibetan letters.

Whatever the precise model was, it is clear that the Tibetans slightly transformed the shape of some letters and elaborated them into a very elegant graphic system. But above all, the Tibetan philologists and translators of that time adapted the Indic script to the phonology of their own language. First, they did not incorporate into the basic consonant alphabet the letters corresponding to retroflex and voiced aspirated sounds, because the Tibetan language did not have such sounds.

Second, they invented at least seven letters in order to render the Tibetan affricate series (ts, tsh, dz) as well as some sounds (w, zh, z, ɦ) that did not exist in the Indic

3. This term refers to orthographic and euphonic rules explained in the traditional treatise RTAGS-KYI.JUG.PA and many later commentaries.

4. See R. Hoernle 1916, Manuscript remains of Buddhist literature found in Eastern Turkestan. About the various scripts at the origin of the Tibetan alphabet, see e.g. van Schaik 2011; Scherrer-Schaub 1999, 2002; and Saejj 2010.

5. They did, however, create a way to transcribe the Indic letters used for the specific sounds of Sanskrit. See below the ‘six reversed letters’ and the ‘five thick letters’.
script they had taken as a model. It is interesting to note that the Tibetans in their system of transcribing Indic languages used the Tibetan affricates to render the Indic palatal.

6. The sounds either did not exist in the Indic languages or the were perceived as different from their Tibetan equivalents. These missing letters appear in grey in the chart above.

7. The chart includes the Gupta letters with their Tibetan correspondences. The Gupta letters are taken from Wikipedia. (http://en.wikipedia.org/wiki/Gupta_script). The chart does not display the Gupta retroflex and voiced aspirated letters, which have no Tibetan correspondence.

<table>
<thead>
<tr>
<th>Gupta</th>
<th>Tibetan Correspondence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka</td>
<td>མ ka</td>
</tr>
<tr>
<td>kha</td>
<td>ཙ kha</td>
</tr>
<tr>
<td>ga</td>
<td>མ ga</td>
</tr>
<tr>
<td>nga</td>
<td>ཙ nga</td>
</tr>
<tr>
<td>ca</td>
<td>མ ca</td>
</tr>
<tr>
<td>cha</td>
<td>ཙ cha</td>
</tr>
<tr>
<td>ja</td>
<td>མ ja</td>
</tr>
<tr>
<td>nga</td>
<td>ཙ nga</td>
</tr>
<tr>
<td>tsa</td>
<td>མ tsa</td>
</tr>
<tr>
<td>tsha</td>
<td>ཙ tsha</td>
</tr>
<tr>
<td>dza</td>
<td>མ dza</td>
</tr>
<tr>
<td>wa</td>
<td>ཙ wa</td>
</tr>
<tr>
<td>ta</td>
<td>མ ta</td>
</tr>
<tr>
<td>tha</td>
<td>ཙ tha</td>
</tr>
<tr>
<td>da</td>
<td>མ da</td>
</tr>
<tr>
<td>na</td>
<td>ཙ na</td>
</tr>
<tr>
<td>pa</td>
<td>མ pa</td>
</tr>
<tr>
<td>pha</td>
<td>ཙ pha</td>
</tr>
<tr>
<td>ba</td>
<td>མ ba</td>
</tr>
<tr>
<td>ma</td>
<td>ཙ ma</td>
</tr>
<tr>
<td>zha</td>
<td>མ zha</td>
</tr>
<tr>
<td>za</td>
<td>ཙ za</td>
</tr>
<tr>
<td>fia</td>
<td>མ fia</td>
</tr>
<tr>
<td>ya</td>
<td>ཙ ya</td>
</tr>
<tr>
<td>ra</td>
<td>མ ra</td>
</tr>
<tr>
<td>la</td>
<td>ཙ la</td>
</tr>
<tr>
<td>sa</td>
<td>མ sa</td>
</tr>
<tr>
<td>ha</td>
<td>ཙ ha</td>
</tr>
</tbody>
</table>

The sign ཙ A (see our transliteration system in 5.9) which was originally a vowel in Gupta was interpreted as a consonant probably because the Tibetans perceived
some initial glottal stop [ʔ]. The sound 'A probably corresponded to a sound close to a voiced glottal fricative [ɦ]. The designer(s) of the Tibetan alphabet came up with a script consisting of thirty basic consonants sälje sumcu (GSAL.BYED SUM.CU) and four diacritic vowels or ‘vowel accent’ yangshi (DBYANGS BZHI). The vowel /a/ is a default vowel, which appears with all the consonants. The chart below gives the thirty consonants and four diacritic vowels of the Tibetan alphabet in the block letter style.

**Chart V.2. – The 30 consonants and the transliteration**

<table>
<thead>
<tr>
<th>Consonant</th>
<th>Transliteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka</td>
<td>ཀ བ ཇ ཉ ཊ</td>
</tr>
<tr>
<td>kha</td>
<td>ཀ བ ཇ ཉ ཊ</td>
</tr>
<tr>
<td>ga</td>
<td>ཀ བ ཇ ཉ ཊ</td>
</tr>
<tr>
<td>nga</td>
<td>ཀ བ ཉ ཊ</td>
</tr>
<tr>
<td>ca</td>
<td>ཀ བ ཇ ཉ ཊ</td>
</tr>
<tr>
<td>cha</td>
<td>ཀ བ ཇ ཉ ཊ</td>
</tr>
<tr>
<td>ja</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>nja</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>ta</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>tha</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>da</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>na</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>pa</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>pha</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>ba</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>ma</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>tsa</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>tsha</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>dza</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>wa</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>zha</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>za</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>zia</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>ya</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>ra</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>la</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>sha</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>sa</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>ha</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td>hia</td>
<td>ཀ ཉ ཊ</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Chart V.3. – The four diacritic vowels: vowel sounds, other than the inherent a, are indicated by diacritical marks above (i, e, o) or below (a) the consonant.**

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Transliteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>བ ཇ ཉ ཊ</td>
</tr>
<tr>
<td>u</td>
<td>བ ཉ ཊ</td>
</tr>
<tr>
<td>e</td>
<td>ཉ ཊ</td>
</tr>
<tr>
<td>o</td>
<td>ཉ</td>
</tr>
</tbody>
</table>

Apart from the basic thirty consonants and the four accent vowels, the Tibetan alphabet makes use of a few additional letters.

---

8. The situation is very similar to the Greek adaptation of the Phoenician alphabet. The Greeks did not have some glottal and laryngeal consonantal sounds used in Phoenician, but they needed vowels that the Phoenicians did not have. So they took various signs corresponding to consonantal sounds alien to the Greek language in order to note the vocalic sounds.
The six ‘reversed letters’ ལོག་པ་དྲུག LOG.PA.DRUG: མ་ Ta, ཀ་ Tha, མ་ Da, ད་ Na, ཐ་ Sha, ་ kSha. and the five ‘thick letters’ མཐུག་པོ་ལྨ་ MTHUG.POLNGA (breathy-voiced aspirated sounds).

These additional letters correspond to the voiced aspirated (or breathy) sounds and the retroflex sounds of Sanskrit. They are essentially used for the transcription of mantras or foreign loanwords.

In various Indic scripts, the letters corresponding to retroflex and aspirated voiced sounds have specific shapes. The ‘inventors’ of Tibetan script did not copy the shape of their Indic equivalent letters. Instead, they used two original and simple methods to render these sounds, which were absent in their phonology. To transcribe the retroflex sounds, they reversed the shape of the dental series (see the above chart). So that མ་ Ta → མ་ Ta, མ་ Tha → མ་ Tha, མ་ Da → མ་ Da, མ་ Na → མ་ Na. They did the same for the fricative sounds: མ་ Sha → མ་ Sha, མ་ kSha.

For the aspirated voiced sounds, they simply affixed a H letter under the various plosive sounds. མ་ ga → མ་ gha, མ་ da → མ་ dha, མ་ ba → མ་ bha, མ་ dza → མ་ dzha, མ་ Da → མ་ Dha.

Tibetan script also has been widely used to transcribe Sanskrit and Zhangzhung texts. In the Tibetan Buddhist tradition (Vajrayana), mantras usually are not translated into Tibetan but, instead, are rendered in a phonetic form of Sanskrit transliterated in Tibetan script. In the Bön tradition, many mantras are reportedly in Zhangzhung language with some Sanskrit or Tibetan words.

The transliteration of mantras is not entirely standardized but the online converter of the Tibetan and Himalayan Library (www.thdl.org) provides an automatic conversion from the Tibetan script to extended Wylie transliteration (and vice versa). One peculiarity of Sanskrit transcription in Tibetan script is that the intersyllabic dot is often dropped within a word.

9. The link between the historically documented Zhangzhung language (in the Dunhguang manuscripts) and the language reported by the later Bön texts is not clear.
Here are some examples of mantras in Sanskrit and their transliteration:

\[ \text{oM ma Ni pa d+me hU~M} \]
\[ \text{oM aHhU~M` badz+ra gu ru pad+ma sid+dbi hU~M':} \]
\[ \text{oM badz+ra sa twa ma ya/_ma nu pA la ya/_badz+ra satwa tue no pa/ _tiSh+bTha} \]
\[ \text{dri D+ho me b+hA wa/_su to Sb+yo me b+hA wa/_su po Sb+yo me b+hA wa/_a nu} \]
\[ \text{rak+to me b+hA wa/_sarba sid+d+bi m+me pra yats+tsba/ sar+b+ba karma su tsa} \]
\[ \text{me/_tsit+taM sbrI yaM ku ru bUM/_ba ha ha boH_b+ha ga wA na/_sarba ta_thA} \]
\[ \text{ga ta/_badz+ra mA me muny+tsa/_badz+rI b+hA ba ma hA sa ma ya satwa AH__oM} \]
\[ \text{badz+ra sa twa bUM/} \]

Many philosophical and religious texts – even those written up to the present day – are headed by Sanskrit titles in Tibetan script. This sometimes occurs when the text was originally written in Sanskrit and then translated into Tibetan, but with the original Sanskrit title preserved. Alternatively, a Sanskrit title may be contrived to accompany a text originally written in Tibetan.

To accommodate Sanskrit sound combinations, several combinations have been created for the Tibetan script. The consonants are usually stacked vertically. These combinations are used in the transcription of some Sanskrit mantras. For an exhaustive list of these combinations, refer to the 'Tibetan Sanskrit' fonts of the Tibetan and Himalayan Library.
5.2. The Syllable Structure

In written Tibetan, the morphological unit is the syllable. It can be identified in most cases as a letter or groups of letters between two intersyllabic dots called TSHEG. The TSHEG which is also referred to as SROG ‘soul’ is fundamental for the reading of Tibetan.

As mentioned earlier, written Tibetan was first used to write down Old Spoken Tibetan, a language spoken at the time of the Tibetan Empire. This language had a rich phonological system and complex syllable structure as we will see now.

The syllable spelling template is \((C)(C)\)\(^{11}\)\((C)(C)\)\(^{11}\)\((C)(C)\)\(^{11}\).

Thus the initial consonant cluster can theoretically comprise up to five consonants (the two postradicals include glides) and the coda two consonants. However, no combination has eight letters. The most complex attested syllable has seven letters CCCCCVCC (see below) and is attested in a few words, but syllables such as CCCCCVCC, CCCCCVC, CCCCVC, CCCCC and CCCCC are frequent. The simplest syllable unit is made of one consonant and one vowel CV.

The rhyme structure (vowel nucleus and coda) is straightforward but the consonants appearing in the onset may have different properties depending on their positions (see Jacques 2004).

In written Tibetan, the basic syllable consists of a single radical consonant, or ‘root letter’ ར ༊ (MING GZHI) and a vowel བྱངས ༊ (DBYANGS): CV.

**Example:** ར ༊. This basic syllable is made of a root letter ར ༊ with the diacritic sign ༊. As noted above, the vowel ༊ is inherent to all radicals, unless a different vowel sound is indicated by a diacritic either above or below the radical.

**Example:** ར ༊. This basic syllable is simply made of a root letter ར ༊ with no diacritic.

This nuclear structure can be expanded upon by combining other consonant sounds with the radical consonant, or by adding consonant sounds after the vowel.

---

11. The only exceptions are the beginning and the end of a sentence. As mentioned above, the intersyllabic dot is often elided in the mantra transcriptions of Sanskrit.
Consonant sounds are referred to as preradicals when they precede the radical, and as postradicals when they follow the radical. Final consonant sounds following the vowel (inherent or diacritical) are referred to as suffixed consonants.

Thus, the maximal syllable consists of seven sounds noted by six letters and a diacritic or inherent vowel. This can be represented by the following scheme:

\[(\text{ANTE})+(\text{PRE})+\text{RAD}+(\text{POST})+(\text{POST2})+\text{VOC}+(\text{SUFF})+(\text{SUFF2})\]

The preradical, radical and postradical correspond to the onset of the syllable whereas the vowel corresponds to the nucleus and the suffixed consonant(s) correspond to the coda of the syllable.

In written Tibetan, the 'radical' is in most cases easily identified since it bears the diacritic vowel. With the vowel ā, there is no diacritic sign, and the identification of the radical must be deduced from its position in the chain of sounds.

Not all the thirgy radical consonants may appear as preradicals or postradicals. One only finds the following possibilities:

1. **antepreradicals**: B
2. **preradicals**: G, D, B, M, ā, R, S, L
3. **postradicals**: Y, W, ā, L
4. **second postradical**: W
5. **suffixed consonants**: G, NG, D, N, B, M, R, ā, S, ā
6. **second suffixed consonants**: D, S

See the chart V.4., *The eight preradical consonants and the ten final consonants*.

---

12. In the some ancient and even recent works describing Literary Tibetan, the term of suffix is also used. Note that the term 'suffix' refers then to suffixed consonant letters, and not to a morphological unit.

13. Concerning the sound R, it generally corresponds to a postradical when it occurs in the second position of the syllable onset, but in some rare words it may be interpreted as a preradical. For example, in the word *BRAG* 'rock, cliff', the B is a radical and the R is a postradical, while in the word *BRIS* 'to write', the letter B is a preradical and the R is the radical. For the consonant L, the traditional grammar treats it as postradical when it occurs in the second position of the syllable onset but from a phonotactic point of view, we will see in Chapter 7 that it behaves like a radical in the Tibetic languages.
There are constraints on the combinations of radical consonants with preradicals and postradicals. Thus, for example, the preradical R is compatible with only twelve radicals out of the thirty consonants: RKA, RTA, RTSA, RGA, RJA, RDA, RBA, RDZA, RNGA, RNYA, RNA, RMYA.

Similarly, the postradical y is only compatible with seven radicals out of the thirty: KYA, PYA, KHYA, PHYA, GYA, BYA, MYA.

There are even more constraints on the combination of both preradical and postradical occurring together. For example, only three radicals (K, G, M) out of the thirty can combine with both the preradical r and the postradical y: RKYA, RGYA, RMYA.

Many combinations not permitted by the Tibetan phonotactics do not have a standard written form. They include SKLA, SBLA, RKLA, RBLA, MBLA, MLA, MLA, MLA, DLA, SDRA, LDRA, GBRA, etc.

The traditional terminology to describe Tibetan letters and syllable structure is based on a graphic point of view rather than a phonological one. Although Tibetan writing is oriented from left to right, some of the preradical letters, postradical letters, and vowels are stacked vertically.
CHART V.A. – The thirty root-letters, eight preradical consonants and ten final consonants

Note: our transliteration differs from the Wylie transliteration for the notation of the letter a, which we transcribe as ʔa. For the transliteration of ʰa [ʰ], we use the Wylie symbol ('). See 5.9.

The five preradicals (G, D, B, M, ') that are written horizontally are called SNGON, JUG ‘prefixed (letter)’ or PHUL, YIG. The three preradicals (R, S, L) that are written vertically, i.e. above the radical are called MGO, CAN ‘the head (letters)’ or superscript. The four postradicals (Y, W, R, L) are called DOGS, CAN ‘attached letter’ or subscript because they are written vertically and attached under the radical. The vowel is called DBYangs ‘melody’ and the first final consonant are called RIES-JUG ‘suffixed (letter)’ and the second final consonant is called YANG, JUG ‘again suffixed’. The five prefixed SNGON, JUG (G, D, B, M,')
and the ten suffixed letters རྗེས་འཇུག་ \textit{RJES, JUG} (G, NG, D, N, B, M, R, L, S) are called འཕུལ་རྟེན་ \textit{PHUL, RTEN}, but there is no general traditional name to call the preradical sounds.

The matrix below represents the graphical display of complex syllables:

The left diagram corresponds to the vowels written above the radical and the right diagram corresponds to the vowel written under the radical.

\[
\begin{array}{cccc}
\text{VOC} & (\text{PRE}_{1}) & & \\
(\text{PRE}_{2}) & (\text{PRE}_{3}) & \text{RAD}^{+} & (\text{SUFF}_{1}) & (\text{SUFF}_{2}) \\
(\text{PRE}_{3}) & \text{RAD}^{+} & (\text{SUFF}_{2}) & (\text{SUFF}_{3}) & (\text{POST}) \\
(\text{POST}) & & & \text{VOC} \\
\end{array}
\]

We give below examples of the graphical display of the written syllable:

BSGRIGS 'arranged' written

\[
\begin{array}{cccc}
\text{I} & \text{B} & \text{S} & \text{G} & \text{S} \\
\text{G} & & & & \\
\text{R} & & & & \\
\end{array}
\]

BSGRUBS 'achieved' written

\[
\begin{array}{cccc}
\text{B} & \text{S} & \text{B} & \text{S} \\
\text{G} & & & \\
\text{R} & & & \\
\text{U} & & & \\
\end{array}
\]

In rare cases, a second postradical text is attached vertically as a subscript:

\[
\begin{array}{cccc}
\text{RAD} & (\text{POST}_{1}) & & \\
(\text{POST}_{2}) & & & \\
\end{array}
\]

For example:

GRWA 'college' written

\[
\begin{array}{cccc}
\text{G} & \text{R} & \text{W} \\
\end{array}
\]
5.3. Calligraphic styles

It seemed important to include a section about calligraphic styles because the calligraphy plays a significant role in the cultural tradition of the Tibetic areas.

The Tibetan script is remarkably conservative in its form. The shape of the letters has not undergone any significant change during the last 1,250 years.

Thus, the knowledge of the alphabet in the block print style or ཆུ་ཅན་ DBU CAN 'Učän' (lit. 'headed') presented above allows for the reading of Old Tibetan texts without any difficulty.

The Učän styles are opposed to ཆུ་མེད་ DBU MED 'Ume styles' (lit. 'headless').

"The first script is characterized by short horizontal lines (the 'head') along the tops of many letters, like the serifs of the Latin script, while the second script dispenses with these lines. There are numerous different styles within the headless script [...]." (van Schaik 2014)

There is a rich tradition of calligraphic styles (see e.g. BOD LJONGS DGA' SKYID GLING 2006). The historicity of some of these styles is problematic and more research must be carried out to establish their historical status. There are many block print styles which include the 'Crawling Black Frog' ཛྷུང་ཐུགས་ཅེན dba’i, the 'Square-Brick' style ཀྲུང་གཏོང་ ཱ་ ཕེ་པ་ rdo rje, the 'Rooster' style གའི་དོན་པ་ rdo rje, the 'Green-Barley-Scattered-on-White-Felt style' རུ་སུ་ཕུག་དཀར་མ་ sbyin dbag pa, the 'Pearl String' style རྒྱུགས་སྐོད་ srog sugs, the 'Black-Beetle-Crawling' style, འབྲུ་ཅེན་ sbyin dbag pa, and many others.

Over the course of history, the Tibetans have also developed many calligraphic Ume styles for cursive, semi-cursive or elegant hand-writing, as well as official or religious purposes, such as གནས་པ་ (KHYU.YIG) Khyuk ‘fast letters’, དཔེ་ཚུགས་ DPE.TSHUGS 'Petshuk', རྫུ་ཁྲུང་ ལྡུ་ (TSHUGS.TSHUNG) Tshugthung.
Nowadays, only the Ucän style is common to all the Tibetic-speaking areas and is used in all the schools, universities and monasteries in the five countries, but some Ume styles are found in various areas from Ladakh to Amdo. Some regions such as Amdo or Spiti, do not use cursive styles very often. In Amdo and Kham Ume styles are usually called རྒྱུན་ཁར། yigkar ‘white letters’. In Ladakh, Ume style is referred to as རྒྱུན། thrayik ‘thin letters’. In Bhutan, the cursive style is called རྒྱོགས་ཡིག་ mgyog concessions ‘fast script’. Certain cursive styles also may be specific to some areas. For example, the Dzongkha cursive style is particular to Bhutan and differs from its equivalent in Central Tibet. However, literate people can usually decipher the various calligraphic styles without difficulty since the variations in the letter shapes are limited.

The first Unicode fonts were developed around 2005. They now include various styles. The main Unicode fonts are: Jomolhari, Microsoft Himalaya, Monlam, Qomolangma Betsu, Qomolangma Chuyig, Qomolangma Drutsa, Qomolangma Tsuring, Qomolangma Tsutong, Qomolangma Sarchen, Qomolangma Sarchung, Qomolangma Suring, Qomolangma Sutung, Qomolangma Tsumachu, Qomolangma Uchenbiaoti, Qomolangma Uchenxiaobiaoti, Qomolangma Dunhuang, Qomolangma Edict, Qomolangma Art, and Qomolangma Woodblock. The font series named Qomolangma has been developed by Tashi Tsering and copyrighted by China Tibetology Research Center (CTRC). (These fonts are available at: www.yalasoo.com)

Here are some examples of Unicode fonts corresponding to various Ucän and Ume styles. These samples illustrate the significance of the calligraphy and its development on the internet during the last 15 years.

The text is the same in the various scripts:

15. See e.g. ZLABA TSHERING 1983, 1999; Tournadre & Sangda Dorje 2003; GO.BADGYIG & HRI.ZHOD.LI 1990.
SNOD KYI SKYON GSUM

RNAB AMI GTOD KHASUB LTABU'T SKYON

YID LAM'I'DZIN ZHAR RDO LITA BU'T SKYON

NYON MONGS DANG DRES DUG CAN LTABU'T SKYON

Jomolhari

སྣོད་ཀྱི་སྐྱོན་གསུམ།

རྣ་བ་མི་གཏོད་ཁ་སྦུབ་ལྟ་བུའི་སྐྱོན།

ཡིད་ལ་མི་འཛིན་ཞབས་རྡོལ་ལྟ་བུའི་སྐྱོན།

བཉོན་མོངས་དང་འདེས་དུག་ཅན་ལྟ་བུའི་སྐྱོན།

Microsoft Himalaya

སྣོད་ཀྱི་སྐྱོན་གསུམ།

རྣ་བ་མི་གཏོད་ཁ་འཇིག་ཤིག།

ཡིད་ལ་མི་འཛིན་ཞབས་རྡོལ་འཇིག་ཤིག།

བཉོན་མོངས་དང་འདེས་དུག་ཅན་འཇིག་ཤིག།

Monlam

སྣོད་ཀྱི་སྐྱོན་གསུམ།

རྣ་བ་མི་གཏོད་ཁ་འཇིག་ཤིག།

ཡིད་ལ་མི་འཛིན་ཞབས་རྡོལ་འཇིག་ཤིག།

བཉོན་མོངས་དང་འདེས་དུག་ཅན་འཇིག་ཤིག།

Qomolangma-Uchen Sarchen

སྣོད་ཀྱི་སྐྱོན་གསུམ།

རྣ་བ་མི་གཏོད་ཁ་འཇིག་ཤིག།

ཡིད་ལ་མི་འཛིན་ཞབས་རྡོལ་འཇིག་ཤིག།

བཉོན་མོངས་དང་འདེས་དུག་ཅན་འཇིག་ཤིག།

Qomolangma-Uchen Sutung

སྣོད་ཀྱི་སྐྱོན་གསུམ།

རྣ་བ་མི་གཏོད་ཁ་འཇིག་ཤིག།

ཡིད་ལ་མི་འཛིན་ཞབས་རྡོལ་འཇིག་ཤིག།

བཉོན་མོངས་དང་འདེས་དུག་ཅན་འཇིག་ཤིག།
Qomolangma-Uchen Sarchung

Qomolangma-Uchen Suring

Qomolangma-Uchen Biaoti

Qomolangma-Uchen Xiaobiaoti

Qomolangma-Dunhuang
Qomolangma-Woodblock

Here are some examples of Unicode fonts corresponding to various cursive and ornamental styles: 'KHYU,YIG', 'BRU,TSHA', 'TSHUGS,THUNG', 'DPE,TSHUGS,RING', 'DPE,MA,KHYUG', etc.

Qomolangma-Betsu

Qomolangma-Chuyig 'KHYU,YIG

Qomolangma-Drutsa 'BRU,TSHA
Qomolangma-Tsuring

Qomolangma-Tsitong

Qomolangma-Tsumachu

Qomolangma-Art
5.4. Ornamental scripts

One should also mention that for ornamental and religious purpose, the Tibetans have used a number of other scripts (see Bod Ljongs Dga’ Skyid Gling 2006) such as the so called Persian style tagzik punyik ཤུག་གཟིག་སྤུང་ཡིག་ (STAG, GZIG SPUNG YIG), Lantsa also called Rañjanā བཀ་པོ་ (LANDZA), Wartu སྐྲ སྭ (Lantsa and Wartu are used only for Indic languages), Marchen རྡོ་རྗེ་ (SMARCHEN), Marchung རྡོ་རྗེ་ (SMARCHUNG), Lhabab Yige ‘the script descended from the God(s)’ ལྷ་བབ་ཡི་གེ (LHA BAB YI GE), the ‘Agate-style script’, ziyik ཤིབ་ཡིག (GZI YIG), and Yangden Yiksar དབྱངས་ལྡན་གསར་ (DBYANGS LDAN YIG GSAR) (See http://www.tibetan-blockstyle.at/). Very few Tibetan monks or experts can read these ornamental scripts and they are usually not used to write texts.

5.5. Graphic abbreviations

Some Tibetan texts make an intensive use of abbreviations. They are frequent in texts written in Ume styles, particularly (but not only) ritual texts. They aim at reducing the number of syllables and the space of the manuscript. The abbreviations are not entirely standardized but they are usually easy to guess. The general principles are to compact two or more syllables syllables into one and to use diacritic signs for some vowels or consonants. Here are some samples of abbreviations written in a Petshuk style. (Examples provided by Thubten Rigzin, pers. comm.):
The above abbreviations respectively correspond to: BYANG,CHUB ‘awakening’, SNYING,PO ‘essence’, DE,BZIHIN,GSHEGS,PA, THUGS,RJE CHEN,PO ‘the great compassion’, Dkon,CHOG GSUM ‘the three jewels’, PHYAG TSHAL LO ‘(I) prostrate’, GZHAL,YAS,KHANG ‘the palace’, KUN,TU,BZANG,PO ‘Samatobhadra Buddha’, SDUG,BSNGAL ‘suffering’. The abbreviated forms may include a grammatical case (see the example of THUGS,RJE CHEN,PO ‘the great compassion + genitive’.

Here are two additional examples of graphic abbreviations: མུ་མོ་ BCUIG (for BCU,GCIG) ‘eleven’ མཱ་མོ་ BCUIS (for BCU,GCIG) ‘twelve’.

The number of abbreviated terms may be very high in some texts and essentially depends on the frequencies of the terms.

5.6. Printing techniques

The various texts written in OT were mainly found on paper manuscripts, wooden tablets, bells or stone pillar (see Chapter 6). A significant part of the Classical literature appears in a printed form.

There are mainly two types of traditional printing techniques: xylographic (requiring a single carved block for each page); and typographic, made with interchangeable pieces of movable type (Gutenberg’s technique). The first printed documents probably appeared in China in the second half of the ninth century as mentioned by Pelliot (1953: 47):

« Le plus ancien imprimé daté qui soit un véritable livre, formé de la réunion de feuilles tirées sur un certain nombre de planches, est le Jingang jing, c’est-à-dire une version chinoise de la Vajracchedikā prajñāpāramitā ‘The Diamond Sūtra’ qui se termine par un colophon imprimé daté du 11 mai 868. »

16. “The oldest printed work that is dated, which is truly a book, made up of a collection of pages printed with print blocks, is the Jingang jing, i.e. a Chinese version of the Vajracchedikā prajñāpāramitā, which concludes with a colophon dated 868, May 11.”
In India, the Jesuits had established printing houses by the second half of the sixteenth century in Goa, but printing production developed during the seventeenth century (see Robin 2003: 156).

The carving and production of xylographic blocks in Tibetan script probably began during the thirteenth century (Jackson 1990; Robin 2003) in China and in Xixia. This time-tested technique is still used in Tibet, Bhutan, India, and Nepal, essentially for topics such as religion, philosophy, medicine or astral sciences. However, most readers now favor modern printed editions.

The diffusion of the Literary language is certainly linked to the xylographic technique and the multiplication of Tibetan printing houses parkhang བར་ཁང་ (PAR.KHANG). Interestingly, during the twentieth century, Tibetans developed a few models of typewriters with Tibetan keys.

Since the 1990s, various Tibetan fonts have been created which allow printing Tibetan texts from computers. As mentioned earlier, there has also been a considerable development of on-line publications in Tibetan after 2006, thanks to the elaboration of Unicode fonts. The coexistence of a living xylographic tradition and computer fonts is a rare phenomenon among the languages of the world.

5.7. Reading and spelling styles

The reading pronunciation of Literary Tibetan has many variations, depending on the native phonology of the readers. The same literary text may be read with tones e.g. in Ü, Tsang, Spiti, Kham, Dzongkha, and Lhoke or without tones in Ladaks and Amdo. There are also many segmental differences. In Amdo and Ladaks, the preradicals R, S, L, D, G, B are more or less pronounced, while they are absent in Central Tibetan, Kham, Dzongkha, and Lhoke. The preradicals ْ and M are read as prenasals in Eastern Tibet but usually not pronounced in Ü, Tsang or in the Western Tibetic areas of Ladakh. Other variations are related to the reading pronunciations of the final consonants (G, NG, D, N, B, M, R, L, S). These nine final sounds are still clearly realized in Ladaks reading pronunciation, but only partially in some Amdo dialects; and they have entirely disappeared in some south-eastern languages of the Kham area.
Note that in Amdo, there is no difference in the reading pronunciation of some letters, such as ཀ་ VS. ཉ་, ར་ VS. ལ་. So in order to distinguish these letters, various expressions describing the shape of the letters are used. For example, (DPAL. JAM 1999: 142; Jangbu Dorje Tshering, pers. comm.) ཁ་ KA = ག་, 'KA with opened mouth'; ག་ KA = ར་, 'KA with mouth shut'; ལ་ CA = ཝ་, 'CA with round (shape)', etc. In the same way, Central Tibetan does not distinguish between ད་ 'A' VS. ཨ་ 'A', so the former is usually referred to as སྒྲུལ་ 'small A', the latter as སྒྲུལ་ 'big A' (Beyer 1992: 43) or, more often, simply ཨ་ 'A'.

The various reading styles are often associated with various spelling techniques. In Amdo, spelling is called ཞེས་དུད་ TSHEGSUD and in Central Tibet སྦྱོར་གློག་ SBYORGLOG.

Differing pronunciations of the alphabet and spelling styles have been recorded by 'BRONG RWA ME BAR in 145 Tibetic varieties (see the site of 'BRONG RWA ME BAR in the webography).

In the Central Tibetan method of spelling, the prefixed letter is followed by the syllable གོག་ 'OOG/oa/ lit. 'under' (or 'after') while in eastern spelling styles (Amdo and Kham), the final consonant is followed by the syllable བཞག་ 'BZHAG/zh/ lit. 'to put'. Both system use the syllable བཏགས་ 'BTAGS/tch/' (Amdo) and /ta/ (Central Tibet).

For example:

བཏབ་ BTAB 'to plant' (past) is spelled in Amdo: /wa ta wa'tha/ > /tap/ and in Central Tibet: /p'ao ta pa/ > /tap/.

སྲུལ་ SBRUL 'snake' is spelled in Amdo: /sa wa'ta/ 'ba ra'ta 'dpa zham chö 'dpa la 'zhag 'dol/ and in Central Tibet: /sa pata pa rata ta shapkyu tu la tül/.

Additionally, within the same area, one should make a distinction between reading and colloquial pronunciations. For example, one can distinguish reading and colloquial pronunciation in Central Tibetan, Amdo, Kham, Dzongkha, and Ladakhs, etc.

Let's give some examples of differences between the various reading and colloquial styles in central Tibetan dialect (Lhasa) and Amdo. In the chart, we indicate the

17. Xinghai (Tségorthang).
tones because the variation between the reading and colloquial pronunciation also involve suprasegmental changes. See Table V.5.

<table>
<thead>
<tr>
<th>Literary Tibetan</th>
<th>Lhasa</th>
<th>Amdo</th>
</tr>
</thead>
<tbody>
<tr>
<td>བལ་</td>
<td>'wool'</td>
<td>/'p′ä:/</td>
</tr>
<tr>
<td>མིག་</td>
<td>'eye'</td>
<td>/'mi/</td>
</tr>
<tr>
<td>ཚབླུག་ 'snake'</td>
<td>/'tül/</td>
<td>/'ṭü:/</td>
</tr>
<tr>
<td>མཐོང་</td>
<td>'to see'</td>
<td>/'t'ong/</td>
</tr>
<tr>
<td>མ་ཕྱིན་ 'I did not go'</td>
<td>/'ma 'c'in/</td>
<td>/'ma 'c'in/</td>
</tr>
<tr>
<td>སྣ་མ་ 'bean'</td>
<td>/'sänma/</td>
<td>/'ṭänma/</td>
</tr>
<tr>
<td>བད་མ་ 'lotus'</td>
<td>/'pänma/</td>
<td>/'wänma/</td>
</tr>
<tr>
<td>གུད་ 'to expell'</td>
<td>/'p′i/</td>
<td>/'p′i/</td>
</tr>
</tbody>
</table>

| Sanskrit mantras and texts are also read according to local traditions. The pronunciation of Sanskrit words may also be specific for a given language or dialect.<sup>21</sup> |

### 5.8. Adaptation of the script to modern Tibetic languages

During the twentieth century, the necessity of writing down modern spoken Tibetic languages has been advocated for educational reasons in order to facilitate access to a written language.

Vernacular Tibetic languages are usually referred to as "phälkä PHAL.SKAD" and they are usually opposed to "BOD.YIG 'written Tibetan' and CHOS.SKAD lit. 'Dharma language', which refers to the Classical language of the

---

18. The verb MTHONG does exist in southern Amdo dialects such as Ngapa and Dzorge but it is not used in many northern Amdo dialects. Instead the verb RIG is used for the same meaning. So 'not seen' /ma rəχ/.  
19. In colloquial Amdo, the verb PHYIN is not used and replace by the verb SONG: /ma song/.  
20. The reading pronunciation /palma/ and the colloquial pronunciation /warma/ is also attested (e.g. Golok and Ngawa).  
21. Thubten Rigzin gives examples of Sanskrit words read with a very different pronunciation in Ü-Tsang and Amdo (pers. comm.).
Buddhist canons. As mentioned in Chap. 1.3, it should be emphasized that the two terms བོད་ཡིག་ BOD.YIG and ཆོས་སྐད་ CHOS.SKAD are sometimes used as synonyms, which is problematic. The reason is that written Tibetan is used for purposes other than Buddhism, such as historical texts, medicine, modern science and technologies, etc.; and also because it is utilized for another religion in Tibet, namely མོན་ BON Bön. The term ཆོས་ CHOS, which originally meant ‘dharma, phenomenon’, has acquired the meaning of ‘religion’ in a more general sense and is also used by Balti and Purik Muslims to refer to Islam. The Bönpos however are reluctant to use the term ཆོས་ CHOS which they perceive as Buddhist and use instead the term མོན་ BON Bön. For example, they don’t say ཆོས་ ‘dhammakaya’, but མོན་ BON ‘dhammakaya’, but མོན་ BON.

The Bön canons are written in བོད་ཡིག་ BOD.YIG but Bönpos would not say that they are written in ཆོས་ ‘CHOS.SKAD!

It is worth mentioning here that many of the Classical Buddhist texts translated from Sanskrit begin with the following phrases: རྒྱ་གར་སྐད་དུ་ RGYA.GAR SKAD-DU… བོད་སྐད་དུ་ BOD.SKAD-DU’ in Sanskrit [it is called …] and in Tibetan […].’ The term ཆོས་ ‘CHOS.SKAD-DU ‘in Dharma language’ is never used in these contexts. Concerning the confusion between ‘written language’, ‘script’ and ‘religion’ and their negative consequences, see 2.6.2.

In Bhutan the term ཆོས་ CHöke (CHOS.SKAD) is used for political reasons to avoid the use of བོད་ཡིག་ BOD.YIG ‘written Tibetan’. Chöke is mainly used in the monasteries whereas Dzongkha, the national language of Bhutan, is used for all other social purposes.

Some Buddhists from Tibet, as well as from other Tibetic regions such as Bhutan, Sikkim or Ladakh, traditionally considered (and still consider) written Tibetan as sacred. In some cases, this great prestige and holiness extends to the script itself.22

22. In some cases, the sacredness of the written language could apply indeed to any script. Once, in Paris, a lama saw his disciples sitting on telephone books and was shocked: ‘Tibetans would never walk over a book, much less sit on one! Two anecdotes reported by a Ladakhi called Gelek (pers. comm. Leh 2017) confirm the sacredness of the Tibetan script: in Lamdon (LAM.SGRON) Tibetan school in Leh,
Not only old prayer books or prayer flags, but also secular books or even sometimes ordinary papers with Tibetan script would not be thrown away but burned or preferably buried into caves, stupas or mani walls.\(^{23}\) Coming back to the Tibetic languages, we have just seen that the contrast is not only between written language and spoken vernacular languages, but also between lay spoken languages and written ‘sacred language’. Thus, the idea that only written Classical Tibetan is worth teaching is still rooted in the minds of many people.\(^{24}\)

The problem is that Classical Tibetan is quite different in its phonology, vocabulary and grammar from the modern Tibetic languages spoken in and outside Tibet. The younger generations find Classical Tibetan quite difficult to learn and often complain about it.\(^{25}\)

Thus, the severe diglossia between Classical Literary Tibetan (or even modern Literary Tibetan) and the vernacular languages has created a lot of obstacles for the democratisation of literacy and the development of a modern curriculum.

Students would have to justify their absence by writing a note in Bod-yig. They would get punished if they threw on the ground these papers with Tibetan script on them (but not if it was a note in Latin or Nāgarī script). The other anecdote is about car plates. A lama from Ladakh tried to prevent people from writing SGER KHOR ‘private car’ in Tibetan script on their plate. He considered that written Tibetan could not be used for any lay purpose.

\(^{23}\) This again is not unique to Tibetans. In the Jewish tradition for example, old or defective copies of Torah and other scriptures are kept in a special ‘cemetery’ called geniza. Similar remarks could also apply to Hinduism and texts written in Sanskrit.

\(^{24}\) Similar attitudes toward the spoken languages are found in other cultural traditions: Written Hebrew was traditionally considered as a ‘sacred language’ (Lashon ha-qodesh) and Eliezer Ben Yehuda, the ‘father’ of modern Hebrew in the beginning of the twentieth century had difficulties convincing other members of Jewish communities that it was necessary to speak in a form of Hebrew for lay purposes. A similar situation, even more similar to the modern Tibetic languages, is attested with the modern Arabic languages, because of dialectal diversity. One often opposes the spoken lay ‘dialects’ called darja with fusha i.e. Literary Arabic, the ‘sacred Qoranic language’. As in the case of the Tibetic languages, modern Arabic languages called ‘dialects’ do not allow a good mutual intelligibility and is some cases do not allow even basic communication.

\(^{25}\) The situation is similar to the speakers of Romance languages (Italian, French, Spanish, Portuguese, Romanian, etc.) who were obliged to learn Latin until the 1960s, 1970s. They felt the language was not relevant to daily life, as well as terribly difficult, and thus were often bored.
Within Tibet, under the Chinese administration, speakers of the main ‘languages’ (or groups of dialects, see Chapter 9) of the three traditional Provinces – Ü-Tsang, Amdo and Kham – who could have easily transcribed their language in a written form, have thus far resisted the idea. Following the long-established tradition, they have favored the use and the teaching of Literary Tibetan (in a modernized form) as the ‘common written language’ within Tibet (both in the TAR and TAPs).²⁶ The reason is that distinguishing written forms of Amdo, Kham and Central Tibetan would undermine the political and cultural unity of ethnic Tibetans living in the Tibet Autonomous Region and the Tibetan Autonomous Prefectures of Eastern Tibet.

Despite the existence of a common literary language (བོད་ཀྱི་ཡིག་སྐད་), the need of ཤི་སྐད་ ( şikä ‘Common (spoken) Tibetan’) remains a real issue since oral communication between Tibetans of various distant areas is still impossible or very difficult in many cases. The precise definition of ཤི་སྐད་ ( şikä in its spoken form (but also in its written form) has generated many debates (see e.g. DPAL, JAM 1999; DON, GRUB LHAR GYAL, ibid.). In the TAR and in the exile communities (see Chapter 9.6), a form of Common Tibetan, based on the language of the capital, has de facto emerged (see Chapter 9) and could easily spread to the regions of Eastern Tibet, but so far, it has not received any official support.

In the other Tibetic areas of India, Nepal, Bhutan and Pakistan the attitude towards the transcription of vernacular languages varies with each region.²⁷

However, generally speaking, the option to write the vernacular languages or གཏན་སྐད་ ( PHAL-SKAD ) phälkä has generated hot linguistic and political debates from officials intellectuals and monks in the whole Tibetic area for more than a century.

For example, the great scholar Gendün Chömphel stated:²⁸

²⁶. Even if some authors incorporate a few dialectal elements in their writing on occasion (see 6.7.2).
²⁷. We leave aside the situation in Myanmar where the number of Tibetic speakers is very small.
²⁸. In the review La tse (Fall 2005), p. 30 LADWAGSKYI AGBAR DANG DE’RGYAB DJONG an article by BA DRAG, TSHE RING, RANG RE NAMS JISRI STAN CHOS KYI THAS NYAD DANG CHOS SKAD CHA
“As long as we use Dharma language and the terms of the Buddhist doctrine, Tibetan language will remain united in the whole of Tibet. However, nowadays some people from the periphery [of the Tibetan area] have the negative habit of writing books and dictionaries in the vernacular language. One should really be very careful with such trends.”

During the last quarter of the twentieth century, new linguistic policies began to emerge which favored the promotion of the vernacular languages such as Dzongkha, Lhoke, Ladakhs, Sherpa and Balti and their transcription in a written form. Until now, out of these five languages only Dzongkha, and to lesser extent Lhoke and Ladakhs, have produced a significant number of publications.

Writing down these languages required linguistic expertise for standardization purposes as well as political backing and support. School books, dictionaries, history books, collections of proverbs, language manuals, and newspapers, etc. have been published in Dzongkha, Lhoke and Ladakhs, and some publications have also been issued in Sherpa and Balti.

Apart from the numerous newspapers published in Literary Tibetan, a few newspapers or magazines have been published in other Tibetan languages, the main ones being the Dzongkha newspaper Kuensel (KUN.GSAL), Ladags Melong ‘The Mirror of Ladakh’ (LADWAG.ME.LONG), the Sikhimese newspapers Sikkim Today (Sikkim Today) and the Sikkim Herald (Sikkim Herald).

29. Some attempts were made much earlier. In the case of Ladakhs, a written style had already been elaborated by Moravian missionaries in the nineteenth century.

30. With the help of native philologues or even foreign linguists such as G. van Driem for Dzongkha and S. Koshal for Ladakhs.
Many newspapers in Literary Tibetan are now published online (see some examples in the webography), but online newspapers in other Tibetic languages are very few. A notable exception is the Bhutanese newspaper Kuensel which is now online. Bhutan Broadcasting Service (BBS) launched its Dzongkha website in 2008.

In the Tibetic area of the southern and western Himalayas, along with the vernacular languages (Dzongkha, Lhoke, Ladaks and Sherpa) and the national or official languages (Hindi, Urdu, Nepali, English), Literary Tibetan has remained in use mainly for religious purposes in the Buddhist or Bönpo monasteries.

Even outside Tibet, the attitude towards the transcription of Tibetic vernacular languages has not always been positive and is still a hotly contested issue, for example in Ladakh. This attitude appears clearly in an article, “Ladakhi language change: progress or decay?” by Sonam Wangchuk published in the Ladags Melong “The Mirror of Ladakh” (2005: 18–22):

“For several months under the umbrella of the Ladakh Cultural forum and the Central Institute of Buddhist studies (CIBS) Choglamsar, have been running an aggressive campaign against what they call ‘the destruction of the old grammar’ by Ladags Melong. They have been issuing veiled and even open threats of mob violence against Ladags Melong for writing in Modern Ladakhi of the twenty-first century.”

One could argue as Zeisler (2006a): “that Literacy in phalskat, on the other hand, as well as an understanding of its grammar through adequate training in school, may well enhance the understanding of choskat, which after all is the younger cousin of Ladakhi and Balti phalskat.”

Even if a part of the Ladakhi clergy together with some Tibetan monks have been very critical in the past about the use of written Ladaks, the situation is now gradually changing. A few prominent members of the Ladakhi clergy have produced excellent translations of religious and historical texts from Classical Tibetan into Ladaks.

As reported by Zeisler (2006a), Bakula Rangdol Nima Rinpoche, an eminent Ladakhi lama “wrote a grammar for Ladakhi phalskat just because ultimately everybody should learn choskat.” The book called "LADWAGS,SI BRDA’,SPROD BZHUS,SO Ladakhi grammar" is composed in written Ladaks (in a style influenced by
The same author, Bakula Rangdol Nima, has also published in 2014 a translation in Ladakhs language (from CT) of the famous Nāgārjuna’s BSHES.PA’I SPRING.YIG ‘Letter to a Friend’ and in 2010, the translation of the Dhammapada CHOS.SI TSHIGS.BC. A well-known Khenpo, Konchok Phanday, has also published several books in written Ladakhs, including CHOS.BYUNG.RAB.BSUD.31 A Brief Modern Religious History of Tibet and Ladakh (2017). Despite its title (both in English and Tibetan), the book also deals with the secular history of Tibet and Ladakh. These authors rightly argue that if the Ladakhs language is not written down and taught in the schools, it will be replaced not by Tibetan (written or spoken) but by Hindi-Urdu or English and that will be the end of the Ladakhi culture.

In the case of Sherpa, the clergy has clearly supported the idea of writing down vernacular Tibetic languages. For instance, Ngawang Tenzin, the abbot of Tengpoche monastery in the Khumbu area of Nepal has advocated the use of written Sherpa in order to preserve the language, which is threatened by the spread of Nepali and English.32

The choice of a script

In general, the choice of Tibetan script to write down Ladakhs, Dzongkha and Lhoke has been natural and obvious.33 However, in some cases, there has been debates about the type of script used to transcribe the language, because of political, national or religious factors.

For example, Balti has been written down using Tibetan, Arabo-Persian and Latin alphabets.34 Some Balti scholars have been reluctant to use Tibetan script because they

31. The title literally means a Brief History of the Dharma. In the book, three languages are used: the preface is in Literary Tibetan, whereas the first part is in written Ladakhs and the second part is in English.

32. Ngawang Tenzin Rinpoche has written the preface for the Sherpa-English dictionary with Literary Tibetan and Nepali equivalents (Tournadre et al. 2009).

33. For example, in Central and Upper Ladakh, the strong cultural references to the Tibetan Buddhist culture have motivated the choice of the Tibetan alphabet. Due to the complex political and linguistic situation within the state of Jammu and Kashmir, Northern India, Ladakhi children have to learn two, three or four scripts: Tibetan alphabet, Devanāgarī alphabet, Arabo-Urdu alphabet and Latin alphabet. The choice of Tibetan or Arabo-Urdu is usually motivated by the religion of the parents.

34. The Quran has already been translated into Balti using the Urdu script.
automatically associate it with Vajrayāna Buddhism. Other scholars on the contrary want to reintroduce Tibetan script, which was dropped nearly five hundred years ago after the conversion to Shi’ah Islam.\(^{35}\)

However, as we have demonstrated before, automatically associating a particular script with a given religion is not always the correct assumption.

A good example is provided by the Bangladeshis, who, after independence and despite the fact the large majority were Muslims, chose to maintain the use of the Bengali Indian script rather than Arabo-Persian.

Similar debates and discussions about the choice of a script have also taken place in the Purik-speaking area of Ladakh and even in Leh, within the Muslim Ladaks-speaking community. Some prominent members of the Arghon Sunni community have declared that they are clearly in favor of the Tibetan script, but lament the fact that the majority of books in the Ladaks language still deal with the Buddhist religion; they wish there would be more books on secular and scientific subjects.\(^{36}\)

Sherpa has been written in three alphabets: Tibetan, Latin and Nāgarī. The alternative scripts, Latin and Nāgarī, were proposed because Sherpa children already learn these two scripts within the Nepalese school system. However, for cultural and linguistic reasons, the Sherpa elite usually favor the Tibetan script. In Sikkim, the school books to learn Sherpa have been using Tibetan script since their introduction.

**Introduction of new graphs**

In most cases, the phonology of these Tibetic languages could be easily transcribed in Tibetan script without creating new letters. There are, however, a few exceptions,

\(^{35}\) A Balti historian from Hardas, Mohammad Sadiq, said “the loss of the Tibetan script was the ‘biggest accident’ in the History of Baltistan” (Mohammad Sadiq, pers. comm. Hardas, 2019).

\(^{36}\) N. Tournadre interviewed in December 2018, Mohammed Shafi Lasu, a well-known lawyer from Leh, who declared that his grandfather, a renown member of the Arghon community, knew only the Tibetan script and unfortunately, given the school system of Ladakh, which imposes the choice between "Urdu (Arabo-Persian)" and "Tibetan scripts" he himself had taken Urdu at school and could not read the Tibetan script although his mother tongue is the Ladaks language. Until the independence of India, educated Muslims of Ladakh in Leh and Kargil (whether Balti or Khache / Arghon) would mainly write in written Tibetan.
particularly in Balti and Lhoke. In most cases, the new letters are created by adding a diacritic sign to already existing Tibetan letters.

In Balti, additional letters were proposed to render various phonemes present in Persian or Urdu but absent in the Tibetan alphabet. The main letters are \( /\text{q}/, /\text{x}/, /\text{y}/ \) which are now standardised Unicode graphemes. For example, the word ‘choice, election’ is written ཨིན་ཏེ་ཁཱ༹བ་ (a word borrowed from Persian and ultimately from Arabic) in Balti Tibetan script. For the /q/ and the /x/ alternative notations, respectively as ཨ/ and ཨ/ with a subscribed dot, have been proposed in the same way as the Urdu words are transcribed in Nagari script. Both notations for /x/ and /q/ are both rational and functional. The notation using reverse letters is classical in Tibetan (such as ཨ/ and ཨ/) and so is the crook which is used to distinguish ཨ/ and ཨ/. Other new graphs have also been introduced to render the sounds /ʕ/ (a dot under ཨ), /f/ (a dot under ཨ), /ɖ/ and /ʂ/, etc. These sounds are also used in Balti to transcribe Urdu or Persian loanwords.

Whatever the final political decision concerning the choice of the Arabo-Persian script or the Tibetan script, it is important to develop a system which would allow automatic conversion from one script into the other. Such systems exist, e.g. for the automatic conversion of Nagari Hindi script and arabo-persian Urdu as well as for the conversion of Tajik Cyrillic and arabo-persian Farsi.

Another Tibetic language, Lhoke, has also introduced four additional letters ཨ/ /py/, ཨ/ /p'y/, ཨ/ /by/ and ཨ/ /my/. These notations are now standardized and Unicode graphemes have been created by Thubten Rigezin (alias Sébastien Carrillo) in 2010 for these letters.

**Introduction of spaces between words**

In Dzongkha and Lhoke, unlike in Literary Tibetan, words or phrases are separated by spaces, just as European written languages. This facilitates reading for people who are not trained in Classical Tibetan. The notation of spaces has generally not been used in other languages, such as Ladaks and Sherpa. So far Modern Literary Tibetan has resisted the temptation of introducing spaces between words. Manuals of
Central Tibetan (Lhasa Tibetan), Kham and Amdo for foreigners have not used this strategy either.

New letter combinations

Among the important innovations related to the new Tibetic written systems, one should mention some specific combinations of consonants which are absent in the syllabic pattern of Classical Tibetan. They include mainly the following combinations of final consonants:

- སར -RS, སན -NS and - སས -DS (Ladaks)
- སན -NGMO, སན -NGM, སན -NM, ས WB -DW, ས WB -LW, ས WB -RM, etc. (Dzongkha)
- སན -NGMO, སན -NGM, སན -NM, ས ས ས -DB, ས WB -RM, etc. (Drânjong)

Let us illustrate these combinations by the following words:

• in Ladakhi: བཟེརས -ZERS/zers/’to tell’ (past), བཟེ -MDZADS/dzats/’to make’ (past, Hon), བཟེ -LTANS/ltans/’to show’ (past);

One of the main principles of the modern written languages such as Ladaks, Lhoke, Dzongkha and Sherpa has been to preserve as much as possible the traditional orthography of Classical Literary Tibetan, i.e., whenever the reading pronunciation of the Literary cognate is not too far of the actual pronunciation in the given language.

37. It seems written Balti has not followed this principle and has not tried to preserve the traditional Tibetan orthography. This may be due to the fact that Balti philologists are not well versed in Literary Tibetan, but also to the fact they generally write Balti in Urdu script.

38. This is reminiscent of the strategy used in the orthography of Romance languages, particularly French, which made certain compromises to preserve some traces of Latin orthography.
In other words, the modern orthography of these written Tibetic languages is a compromise between the reading pronunciation of Literary Tibetan and their vernacular pronunciation.

Thus, for example, the word 'iron' is written ལྕགས་ LCAGS in Ladaks, Dzongkha and Lhoke, as it is in Literary Tibetan, although, the initial L and the final S are not pronounced in Dzongkha and Lhoke.

Of course, this orthographic principle would not apply to some grammatical morphemes and lexical items when they are absent in Classical Tibetan or when the reflexes are not obvious.

After a little training, people who have a good knowledge of Literary Tibetan would be able to read Dzongkha, Ladaks, Lhoke, Sherpa or Balti without great difficulty. Conversely, people who have a good knowledge of one of those modern written Tibetic languages can easily learn Literary Tibetan.

5.9. Transliteration

Tibetan script can be easily transliterated in the Roman alphabet. The transliteration provides the orthography of Literary Tibetan (see the Conventions, p. 25-30) but does not give any precise information about the modern pronunciation (for this purpose, see the transcription presented in the Chap. 7). Since Western scholars first encountered Tibetan script, multiple ways of transliteration have been proposed. At present, a standardized transliteration system, known as the Wylie Transliteration (already mentioned in Chap. 1) is widely used by scholars around the world. Let us compare several ways of 'letter-to-letter' transliteration of the thirty radical letters39 (see Chart V.6.).

Transliteration allows the rendering of the exact spelling of the original Tibetan text. This is achieved by representing each and every character and diacritic of the Tibetan orthography by one – or occasionally two – character(s) of the Roman alphabet.

39. Some transliteration systems may use different roman scripts for certain Tibetan letters depending on radical or preradical letters. There are other systems such as the Pelliot transliteration and the US library of Congress systems.
Transliterations in the Roman alphabet write the Tibetan letters in a linear horizontal form even when they are vertically stacked. The superiority of the system of Wylie transliteration is that its font set is limited to ASCII characters, which enables us to input the characters using any language keyboards and softwares and to save a document in any standard formats, such as .txt, .rtf and .doc.

An ‘extended’ version of the Wylie system has been developed by the Tibetan and Himalayan Library of the University of Virginia.

Wylie transliteration is very convenient to write in Tibetan Unicode, which allows writing documents and emails in Tibetan script. Additionally, as mentioned in 5.1, Wylie Transliteration can be automatically converted into Tibetan script and vice versa. Thus, apart from the visual display, the two systems are equivalent.

### CHART V.6. – Various transliteration systems for the Tibetan script

<table>
<thead>
<tr>
<th>Tibetan script</th>
<th>Jäschke</th>
<th>Hill</th>
<th>Zeisler/Bielmeier</th>
<th>Chinese style&lt;sup&gt;41&lt;/sup&gt;</th>
<th>Wylie</th>
</tr>
</thead>
<tbody>
<tr>
<td>ལོ</td>
<td>ka</td>
<td>ka</td>
<td>ka</td>
<td>ka</td>
<td>ka</td>
</tr>
<tr>
<td>སོ</td>
<td>k’a</td>
<td>kha</td>
<td>kha</td>
<td>k’a/ kha</td>
<td>kha</td>
</tr>
<tr>
<td>ོོ</td>
<td>ga</td>
<td>ga</td>
<td>ka</td>
<td>ga</td>
<td>ga</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>na</td>
<td>ŋa</td>
<td>ŋa</td>
<td>ŋa</td>
<td>nga</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>ca</td>
<td>ca</td>
<td>ča / ča</td>
<td>ča</td>
<td>ca</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>č’a</td>
<td>cha</td>
<td>cha / čha</td>
<td>č’a</td>
<td>cha</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>ja</td>
<td>ja</td>
<td>ja / Ja</td>
<td>dza</td>
<td>ja</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>nya</td>
<td>ŋa</td>
<td>ŋa</td>
<td>ŋa</td>
<td>nya</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>ta</td>
<td>ta</td>
<td>ta</td>
<td>ta</td>
<td>ta</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>t’a</td>
<td>tha</td>
<td>tha</td>
<td>t’a / tha</td>
<td>tha</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>da</td>
<td>da</td>
<td>da</td>
<td>da</td>
<td>da</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>pa</td>
<td>pa</td>
<td>pa</td>
<td>pa</td>
<td>pa</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>p’a</td>
<td>pha</td>
<td>pha</td>
<td>p’a / pha</td>
<td>pha</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>ba</td>
<td>ba</td>
<td>ba</td>
<td>ba</td>
<td>ba</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>ma</td>
<td>ma</td>
<td>ma</td>
<td>ma</td>
<td>ma</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>tsa</td>
<td>tsa</td>
<td>tsa</td>
<td>tsa</td>
<td>tsa</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>t’sa</td>
<td>tsha</td>
<td>tsha</td>
<td>t’sa / tsha</td>
<td>tsha</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>dza</td>
<td>dza</td>
<td>dza</td>
<td>dza</td>
<td>dza</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>wa</td>
<td>wa</td>
<td>wa</td>
<td>wa</td>
<td>wa</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>ža</td>
<td>ža</td>
<td>ža</td>
<td>ža</td>
<td>ža</td>
</tr>
<tr>
<td>ོ ོ</td>
<td>za</td>
<td>za</td>
<td>za</td>
<td>za</td>
<td>za</td>
</tr>
</tbody>
</table>

<sup>41</sup> The so-called Chinese style is based on the Wylie transliteration but is influenced by the IPA.
The Tibetan Script

For example, a stance such as:

\[\text{dge slong dag gam mkhas rnams kyis/bsregs bcdar ba'i gser bzhin du/legs par brtag la nga yi bka/bling bar bya yi gus phyir min/}\]

can be automatically converted into Tibetan script, in one click:

\[
\begin{align*}
\text{དགེ་སློང་དག་གམ་མཁས་རྣམས་ཀྱིས།} \\
\text{།བསྲེགས་བཅད་བརྡར་བའི་གསེར་བཞིན་དུ།} \\
\text{།ལེགས་པར་བརྟག་ལ་ང་ཡི་བཀའ།} \\
\text{།} \\
\end{align*}
\]

In this book, we will use the Wylie ‘letter-to-letter’ transliteration with only one minor amendment. The letter ཨ is normally transliterated as A, but we will transliterate it as ʔA. Our choice is motivated by two reasons. First every Tibetan letter is rendered by a specific sign of the Latin alphabet except this letter ཨ. The A of the Wylie transliteration is ambiguous because it may refer to either the inherent vowel following a consonant (see the chart above) or to the ‘consonant’ ཨ. The second reason is motivated by a phonological argument. The letter ཨ was not a vowel since it was grouped together with consonants and it probably corresponded to a glottal stop /ʔ/. Thus, in

\[\begin{array}{|c|c|c|c|c|}
\hline
\text{Tibetan script} & \text{Jäschke} & \text{Hill} & \text{Zeisker/Bielmeier} & \text{Chinese style}^{42} & \text{Wylie} \\
\hline
\text{ག} & .a & ha & ḥa & va & 'a \\
\hline
\text{ད} & ya & ya & ya & ja & ya \\
\hline
\text{ར} & ra & ra & ra & ra & ra \\
\hline
\text{ལ} & la & la & la & la & la \\
\hline
\text{ཤ} & sa & sa & sa & ca & sha \\
\hline
\text{ས} & sa & sa & sa & sa & sa \\
\hline
\text{ཧ} & ha & ha & ha & ha & ha \\
\hline
\text{ཨ} & 'a & a & a & a /ʔa & a \\
\hline
\end{array}
\]

42. The so-called Chinese style is based on the Wylie transliteration but is influenced by the IPA.

43. We did not use here the small caps or italics for the Wylie transliteration because the THL converter is case sensitive.

44. This is similar to de Nebesky-Wojkowitz’s (1956) way of transliteration.
this book, we transliterate the letter ཨ’ as ʔ. In addition, all the transliterated letters are
to be in italics and small capitals in order to distinguish clearly the transliteration from
the body of text and other transcriptions.

Another small difference with the Wylie transliteration is the distinction between
the ‘prefixed letter’ ༌ (preradical) in front of a radical letter ཡ and the radical letter ༌
in front of a subscript ཡ (postradical). Tibetan spelling does not allow such ambiguities
in other cases. In fact, this ambiguity does not appear in the script itself because in the
first case, the letters are written on a horizontal line whereas in the second case, they
are stacked vertically. བ འི and བ ར འི are both made of the same letters ༌-༌-༌-༌. In
the Wylie transliteration, these two words are distinguished by inserting a period after the
prefixed letter ༌ བ ‘abyss’ བ ར ‘wall’. In our transliteration system, we
do not use a period because this sign is used for the intersyllabic dot (see below). So in
order to distinguish such combinations, we use a medium level dot after the
prefixed letter ༌ བ ‘GYANG’ versus བ ར ‘GYANG’.

Tibetan texts appear as strings of syllables separated by an intersyllabic dot without
any word segmentation. The Wylie system uses a space to render the intersyllabic dot.
Thus a Tibetan sentence in Wylie transliteration appears as:

In the above example, there is no word segmentation.

In descriptive linguistic studies, sentences are often presented with word
segmentation. We will present word segmentation in the following way; polysyllabic

45. Thus for example in the Wylie style compare བ བ བ བ kha ‘mouth’ and བ བ བ བ ‘Amdo province’. In the first word, the Wylie e notes the default vowel of the sign བ བ kha, while in the second
word, it notes the consonant བ བ that we propose to transliterate as ʔA-MDO. Of course the ambiguity is
resolved by the combination of letters since the ‘a’ corresponding to the consonant བ བ occurs only at the
initial of a syllable, while the default vowel ‘a’ occurs only after an initial consonant.

46. “First one should contemplate this precious human birth which is hard to obtain and easily
lost. One should make this life meaningful.”
words are indicated by a dot between the syllables\(^\text{47}\) and words are separated by a space (just as in European languages). Thus the above sentence may be displayed as:

\[\text{DANG.PO BSGOM-BYA DAL. 'BYOR RIN.CHEN 'D/ /THOB DKA' JIG S} \text{LA DA,RES DON YOD BYA/}.\]

Thanks to the spaces, one immediately visualizes the words, either monosyllabic such as \text{BSGOM}, \text{'D/}, \text{THOB}, etc. or polysyllabic such as \text{DANG.PO}, \text{DAL. 'BYOR}, \text{RIN.CHEN, DARES}.

For glossing rules, it is also useful to mark the grammatical affixes (suffixes and prefixes) or clitics. For this purpose, we use the hyphen (as opposed to the dot which marks the syllables within lexical words).

\[\text{GZHAN-LA} \quad \text{LTA-BA-'I} \quad \text{MIG} \quad \text{YOD-NA}/\]
other-DAT look-NMLZ-GEN eye ENV-if

\[\text{RANG-LA} \quad \text{LTA-BA-'I} \quad \text{ME.LONG} \quad \text{DGOS}//\]
self-DAT look-NMLZ-GEN mirror need

This convention will be used in Chapter 8.

There is another issue in the transliteration: the capitalization rule. The Wylie system does not decide on a rule of capitalization for the proper names, but according to Wylie’s own convention (1962), the first letter of a word should be capitalized, as ‘Lha sa’ and ‘Skal bzang’. Another capitalization rule also exists: the radical letter of the first

\[\text{RGYALNGA} \quad \text{may be segmented as RGYAL 'victorious' NGA 'me' or RGYA 'vast' LNGA 'live', LAGZHI > LA 'past' + EZHI 'four' or LAB 'to talk' + ZHI 'calm'.}\]
\[\text{BYASNA > BYA 'bird' + SNA 'nose' or BYAS 'to do (past)' + NA 'if'; PHAGRO > PHA' there' + GRO 'wheat' or PHAG 'pig' + RO 'corpse'.}\]

\[\text{47. The use of a dot similar to a period to mark the Tibetan intersyllabic dot has been applied by various authors (Denwood 1999; Vokurková 2008; Oisel 2013). It is necessary to indicate the syllabic border because syllables are often meaningful. RGYALNGA may be segmented as RGYAL 'victorious' NGA 'me' or RGYA 'vast' LNGA 'live', LAGZHI > LA 'past' + EZHI 'four' or LAB 'to talk' + ZHI 'calm'. BYASNA > BYA 'bird' + SNA 'nose' or BYAS 'to do (past)' + NA 'if'; PHAGRO > PHA' there' + GRO 'wheat' or PHAG 'pig' + RO 'corpse'.}\]
syllable of a word should be capitalized, as ‘lHa sa’ and ‘sKal bzang’. This book uses the radical letter capitalization rule to make references.\(^\text{48}\)

5.10. Derivation of romanization from transliteration

As we explained earlier (see Conventions) a standardized system of romanization is necessary to avoid multiple spellings of Tibetan toponyms and names. In order to achieve this goal, we start from the Classical orthography in Wylie transliteration, which is standardized. The romanization obtained by this method roughly corresponds to the reading pronunciation of Central Tibet.

We summarize here the rules of derivation from transliteration to romanization. Only the letters which are pronounced are preserved in the romanization. This system of romanization was introduced by Tournadre and Sangda Dorje (1998, in Appendix 7). We have proposed here some minor adaptations.

For example, in གཞིས་ཀ་རྩེ་ GZHIS.KA.RTSE; only the letters in bold GZHIX.KA.RTSE are pronounced, and thus the romanization is Zhikatse. Here are other examples: in the words དབུ་གྲིན་མོ་ DZAM.BU.GLING, དགེ་གྲུབ་ DGE.LUGS.PA and DGE.BSHE, if we simply delete the letters that are not pronounced we get respectively dzambuling, gelugpa and geshe. With this convention, དཔལ་ KHAMS should be transcribed as Kham because it reflects the pronunciation. Note that the word Kham also refers to a Tibeto-Burman language spoken in Nepal and not related to Tibetan (see Watters 2009). Thus, we strongly recommend for linguists to use the Wylie transliteration kham or to specify ‘Kham Tibetan’.

The letters used in the transliteration and in the romanization are identical with one exception: བ ར which is rendered as č. For example, compare the transliteration and the romanization in the two following words: ཆོས་ TSHAS > Čöntsha. If the diacritic is not available, one can simply write without it as ‘Čöntsha’.

\(^{48}\) In the Bibliography and citing the author names of the previous works in Tibetan, the proper name is transliterated with the Wylie method, and the radical letter will be in capital letter and this radical letter will be considered as the initial letter for the roman alphabetical order. This method has also been used by René de Nebesky Wojkowitz (1956).
Additional rules

Some combinations of letters yield specific sounds. To render better pronunciation of ST following the main principle, explained above, we add a few signs and modify some of the Wylie letters.

**Modifications of the pronunciation:**

1) the initial consonant clusters *PY, PHY, BY and SBY/BY* become affricate sounds: *PY > c, PHY > ch, BY, SBY, BY > j*

Thus: བྱིན་བདག *SBY/N.BDAG > jindak, གྱིིང་ཐང་*BY.ANG.THANG > Jangthang,
ཟླྱི་པ་*JAM.PA > Jampa, བྱམས་པ་*JAM.DPAL.DBYANGS > Jampälyang,
གཅན་ཚ་*GCAN.TSHA >Čäntsha

2) the initial consonant clusters *KR, KHR, GR and PR, PHR, BR, TR, DR (and other combinations of GR, DR or BR with a preradical) become retroflex sounds: KR/PR/TR > tr
KHR/PHR > thr
GR/DR SGR, DR, BGR, BSGR, SBR, etc. > dr

Thus: བློལ་མ་*SGROL.MA > Drölma, གྲོ་ལེགས་*RTEN.BREL > trendrel,
འབྲས་སྤུངས་*BRAS.SPUNGS > Dräpung, ལྷུགས་*PHYAG.PHRENG > Chagthreng.

3) The initial consonants *G, J, D, B, DZ* without preradical are normally voiceless and aspirated in Common Tibetan but this is not the case in some conservative languages. We ignore this distinction in the Romanization in this book, but if needed, one may indicate that the initial consonant does not have any preradical with an apostrophe (as proposed by van Driem in 1998 for Dzongkha). In many Tibetan languages, reflexes of a radical letter without preradical yields a devoiced consonant, sometimes associated with an aspiration or a breathy sound.

4) The letter ‘’, historically pronounced as [fi] when directly followed by a vowel, is preserved in the initial of a word but is deleted elsewhere:

\[\text{འོལ་མོ་ལུང་རིང་} > \text{Olmo lungring}\]

5) Vowels

The vowels E, I are not modified. But A, O and U are transformed into ä, ö, ü in front of D, N, L and S. The vowels ä, ö are pronounced as in German or Swedish and ü is pronounced as the German ü or the French ‘u’ in ‘tu’. Ex. ཐུབ་བསྟན་ > Thubtän, རིགས་པ་ > Milarapă,

6) Specific notation of some combinations:

\[\text{DB} \text{ (in front of A, E, I) } \rightarrow \text{w}; \text{DB} \text{ (in front of O, U) } \rightarrow \text{u}; \text{DBR} \rightarrow \text{r}; \text{DBY} \rightarrow \text{y}; \text{ZL} \rightarrow \text{d}; \text{MY} \rightarrow \text{ny}.\]

The letters BA and BO are pronounced /wa/ et /wo/ when they appear as initial of the second syllable of a word:

\[\text{ཐུབ་དབང་} > \text{Thubwang}, \text{དབུས་} > \text{Ü}, \text{ངོ་བོ་} > \text{ngo} \text{w}o, \text{ལྷ་ས་བ་} > \text{lhasawa}, \text{སྨྱུག་གུ་} > \text{nyuggu}.\]

7) The final consonants G, B, D and S

The final G and S are transformed respectively into ‘k’ and ‘p’. \[\text{DGE.LEGS} > \text{Gek}, \text{KHABTAGS} > \text{khatak}, \text{K.A.SHAG} > \text{kashak}, \text{DON.GRUB} > \text{Döndrup}.\] The pronunciation of ‘k’ in the final position is very light and often realized as a glottal stop. The finals D and S (whether at the end of a syllable or a word) are not pronounced and deleted in the romanization: \[\text{MLA.RAS.PA} > \text{Milarapă}; \text{BOD} > \text{Bö}.\]

Whenever the pronunciation of local place names is significantly different from the Standard one, it appears in square brackets \[\]. For example, \[\text{DPAL,SKYID} \text{ Pälkyi locally pronounced as [Pashi],} \text{ DP.A.RIS} \text{ Pari locally pronounced as [Xwari],} \text{ ZL.CHU} \text{ Dachu, locally pronounced as [Dzachu] or [Lachu],} \text{ BA.LUNG} \text{ Balung locally pronounced as [Melung], etc.}\]
The main advantages of the romanization proposed here is that it limits the variation in the spelling of Tibetan names because the romanization is directly derived from the Classical orthography, which is standardized to a large extent.

The romanization rule is applied only for the Tibetic languages and their related proper names. It may not be applied for the non-Tibetic languages, even though they are spoken in Tibetosphere. Hence, we can understand the difference between the Tibetic languages and non-Tibetic ones in a clearer way, e.g. Minyak (a variety of Tibetic) and Minyag (a language of Qiangic) for the same Tibetan orthography and transliteration སྲིབ་ཟོང་ MLNYAG, and Gyalrong (a variety of Tibetic) and rGyalrong (a language of rGyalrongic) for རྒྱལ་རོང་ RGYAL.RONG.

5.11. Other scripts of the Tibetic area

As we mentioned in the introduction of this chapter, before the introduction of the Tibetan script in the seventh century, there is no evidence of a script used in Tibet despite the fact that various scripts had long been attested in both China and India. Since the invention of a writing system on the basis of Indic scripts, Tibetans have been very faithful to their script and during the last 1,250 years have neither written their language with Chinese characters nor any other scripts. Even nowadays, Chinese characters are never borrowed to write in modern Literary Tibetan. The situation is thus very different from the one found in Japan, Vietnam and Korea where many Chinese characters have been borrowed in the course of history.

The Tibetan script has spread not only to all the traditional provinces of Tibet – Ü-Tsang, To-Ngri, Amdo, Kham, and rGyalrong, etc. – but also, as we have seen in 5.8, outside Tibet to the Bhoti regions of Ladakh, Dränjong (Sikkim), and Bhutan, etc. A few other scripts are historically attested at the periphery of the Tibetan Empire. They include some ancient scripts: 'Phagspa, Mongol bichig and Soyombo (Mongolian scripts), Tangut (the script of the Xixia Kingdom), the Yi logographic script, the To-mbu pictographic script and the Geba syllabary of the Naxi, etc. Note that all these scripts except Mongolian (Bichig) are no longer in use or restricted to liturgical practices. More recent writing systems, such as the Lepcha alphasyllabic script (in Sikkim, India) and Yi syllabary (in Yunnan, China) are also attested. However, all these scripts, with the
exception of 'Phagspa, were not conceived by Tibetans and are considered as foreign scripts which do not belong to the 'Tibetan world'. We will briefly present the 'Phagspa script, which had an unusual history.

5.12. The 'Phags-pa script

The 'Phags-pa script is usually called ཟྭ་ཧོར་ཡིག་གསར་པ་ Horyik sarpa in Tibetan which means 'New Mongolian script'. It was invented by a Tibetan lama of the Sakya school, འཕགས་པ་བློ་གྲོས་རྒྱལ་མཚན། (PHAGS.PA BLO.GROS RGYAL.MTSHAN) Phagpa Lodrö Gyaltshän during the thirteenth century on the model of the Tibetan script and also possibly the Khotanese script (see Shen, Zhongwei 2008). In European languages, this script is called by the first name of its inventor འཕགས་པ། Phagpa in the transliterated form PHAGS.PA. In Chinese this script is also called 八思巴字 basiba zi (< PHAGS.PA).

This script is written vertically from left to right like the traditional Mongolian script (unlike traditional Chinese which is written vertically but from right to left).

For at least four centuries, 'Phags-pa script played a very significant role in the transcription of Chinese. A phonological work produced in the second half of the fourteenth century, 蒙古字韵 Menggu Ziyun is written in Chinese and in the 'Phags-pa transcription system, making knowledge of 'Phags-pa script crucial for the reconstruction of Middle Chinese.

Several authors have suggested that 'Phags-pa script, together with Tibetan, had an influence on the shape of the letters of the Korean alphabet when it was invented in the fifteenth century (see also Tournadre 2014b). For more about 'Phags-pa script, see Andrew C. West’s site: /www.babelstone.co.uk/Phags-pa/. When looking carefully at the 'Phags-pa script below, we can recognize in the vertical lines (from left to right) the letters of the Tibetan alphabet written vertically (with an additional bar on every letter).


A new Unicode font in Horyik is available in the Qomolangma series.
6. Literary Tibetan and its evolution

6.1. The various stages of the literary language

It is possible to distinguish three main periods of written Tibetan or literary Tibetan: Old Tibetan (eighth to eleventh centuries), Classical Tibetan (twelfth to nineteenth centuries) and Modern Literary Tibetan (twentieth century to the present).^1^

Old Tibetan is defined by Bialek (2018b) as “[…] the language(s) of non-translatory Tibetan documents discovered in Central Asian oases (Dunhuang, Turfan, etc.) and of the inscriptions from Central Tibet.”^2^

Some authors (see Miller 1970; Qu 1996; Nishida 1970; Zeisler 2004: 215-220) have proposed a more detailed analysis than the three periods listed above and have distinguished up to six stages to account for the evolution of the language over more than a millennium.\(^2\)

For Old Tibetan alone, one often encounters the following periodisation: Early Old Tibetan (EOT), Middle Old Tibetan (MOT) and Late Old Tibetan (LOT). See e.g. Nishida 1970; Róna-Tas 1992; Takeuchi 2012; Bialek 2018b.\(^3\)

Some contemporary dialects, which have preserved the most archaic features, could be directly derived from the two early stages as proposed by Bialek (ibid.): “Proto-WAT [Proto Western Archaic Tibetan] descended from EOT, Proto-AT

---

^1^ There isn’t a complete consensus on the dates. Zeisler (2004) defines OT as: mid-eighth to tenth century A.D.

^2^ For example Nishida (1970: 172-174) proposes the following periods: Old Tibetan (? to seventh century), Middle Old Tibetan (seventh to beginning of ninth century), Late Old Tibetan (ninth to tenth century), Middle Tibetan (tenth to early seventeenth century), New Tibetan (seventeenth to nineteenth century), Modern Tibetan (twentieth century to present).

^3^ There is some ambiguity in the use of “Old Tibetan.” It may refer to Old Literary Tibetan or to the reconstructed “Old Spoken Tibetan” (Bialek, pers. comm. 2020). In the above mentioned articles, EOT, MOT and LOT refer to the stages of the spoken language. The evolutions of the Literary language and the spoken language(s) are not entirely parallel and the pace of change is different. Due to its conservatism, the written language always reacts with some delay to changes and may integrate them a long time after they occurred in the spoken language(s).
[Proto-Amdo Tibetan] from MOT, whereas the remaining dialect groups most probably from LOT.”

We will not consider here a detailed classification of the stages, but instead will stay with the terms of Old Tibetan, Classical Tibetan and Modern Literary Tibetan, which are used by most authors (see e.g. Zeisler, 2004).

In Tibetan, the first period is usually referred to as བོད་ཀྱི་ཡིག་རྙིང་ BOD-KYI YIG-RNYING (‘Old written Tibetan’), the second as རྒྱུན་སྲོལ་བོད་ཡིག་ RGYUN-SROL BOD-YIG (‘Traditional Written Tibetan’), and the third as དེང་རབས་བོད་ཡིག་ DENG-RABS BOD-YIG (‘contemporary Written Tibetan’).

The term བོད་ཡིག་ BOD-YIG (‘written Tibetan’) may be used for all written forms of Tibetan, including Classical or Modern, and even Old Tibetan.

Note that the term ཆོས་སྐད་ CHOS-SKAD ‘the language of Dharma’ is also frequently used to refer to Classical Tibetan in relation to religious or philosophical writings. However, we should distinguish the literature translated from Sanskrit from original Tibetan literature. The translations of Buddhist texts into Tibetan have specific characteristics not found elsewhere and the language has sometimes been referred to as ‘Old Church Tibetan’ (Miller 1970). Their vocabulary and syntax clearly show some influence of Sanskrit. (See e.g. Simonsson 1957; Verhagen 2001.)

It is not easy to establish precisely the transition period from Old to Classical Tibetan. The “Sakya aphorisms” (ས་སྐྱ་ལེགས་བཤད་ SASKYA LEGS.BSHAD) composed by ས་སྐྱ་པཎྡི་ཏ་ཀུན་དགའ་རྒྱལ་མཚན་ Sakya Pandita Kunga Gyantshen (1182–1251) can be considered one of first great works written in Classical Tibetan.

The transition from Classical to Modern Literary Tibetan is no easier to determine. During the throughout the twentieth century and up to the present, one finds texts that are written in a Classical style.

We will present in Chapters 6.5 and 6.6 the distinctive characteristics of Old and Classical Tibetan from a linguistic point of view.
6.2. The liturgical language of Vajrayāna and Bön

Classical Tibetan has been used as the written medium in Tibet for many centuries but it has also been more largely the liturgical language (with the exception of mantras, see 5.1) for Vajrayāna Buddhism and Bön in various countries.

Consequently, aside from Tibetans themselves, the priests or monks — lama, ngakpa, labön, tampa, etc. — of many ethnic groups read pecha (religious books) or specific ritual texts and prayers in Classical Tibetan. The mother tongues of these religious performers include languages that belong to various families such as Tibetic, Bodic, Qiangic, Mongolic, Turkic and even Tani or Naic.

With the exceptions of Balti, Purik and some minorities of Amdo and Central Tibet who are followers of Islam as well as a few other minorities, the Tibetic-speaking ethnic groups use Classical Tibetan as their liturgical language. This is also the case of the various non-Tibetic speaking ethnic groups in the following countries:

- in Russia: Kalmyk, Buriat, Tuva;
- in Mongolia: Mongols;
- in China: Yughur (or “yellow Yughur,” Turkic-speaking ethnic group), Tongt, Oirat, Monghul, Dongxiang [Mongguer], Bonan, Shira Yughur (Mongolic-speaking ethnic group), Erzu, Namuyi, Prinmi (partly), Shuhing, Qiang (partly), rGyalrong, nGochang (Guqiong), nDrapa, Choyu, Minyag, Naxi (partly), Idu (marginally), etc.;
- in Nepal: Tarlikam (Dölpo), Gurung (partly), Manang, Tamang (partly), Thakali (partly), Baramu-Thami, Chantel, Lepcha (partly);
- in Bhutan: Tshangla, Bumthang, Dzala, Dakpa, Chali, Kurtö, Kheng, etc.;
- in India: Kinnauri, Lahuli, Almora;
- in China, Taiwan and many western countries, newly converted adepts of Tibetan Buddhism read their prayers in Classical Tibetan. In many cases, the Tibetan script is accompanied by a phonetic transcription.
6.3. The earliest written sources

The earliest-known documents written in the Tibetan script are 1,250 years old. No older materials have been found to date. From historical sources, we can postulate that the written language was based on the dominant language spoken in the Tibetan Empire (seventh to the ninth century A.D., see Chapter 5). The capital of the ancient Tibetan Empire was first situated in the Yarlung valley and then moved to Lhasa. (See Ryavec 2015, as well as the map, 'Tibetan at the heart of Asia', in this volume.)

In the ninth century, the Tibetan Empire controlled most of the Tarim Basin (East Turkestan, in the present Xinjiang Uyghur Autonomous Region of China), including the cities of Khotan, Kucha, Aksu, and Kashgar, most of Pamirs, the Karakoram Range and the Hindu Kush as well as some areas on the southern flank of the Himalayas. Thus the Tibetan empire extended westward as far as the ancient kingdom of Tukharistan (which is located in the modern states of Afghanistan and Northern Pakistan) including Baltistan, Gilgit and Ferghana (near modern Tashkent) and eastward as far as the Hexi (Gansu) corridor. Towards the South, the Tibetan army once even reached the Bay of Bengal.

Thus one finds various epigraphs or manuscripts in Old Tibetan across the extent of the Tibetan Empire. Old Tibetan texts (see Bacot & Toussaint 1940; Richardson 1998; Li 1987, etc.) have been found carved on pillars and rocks or written in the form of manuscripts, on palm leaves, carved in wood tablets or cast in metal bells (see Bacot & Toussaint 1940; Richardson 1998; Li 1987; Takeuchi 1995; Chen 1984). The oldest extant document, a stone carving, is the text on the Zhol pillar in Lhasa, dating from 764.

---

4. As expressed by C. Beckwith (1993: 83), [in 715] the Arabs from the west, the Chinese from the east, and the Tibetans from the south – the three greatest expansionistic states of early medieval Asia – had converged.

5. LHAMCHOG.RGYAL (2011) published a paper about a recently discovered bell inscription from the temple DGYUL.DAN.BYIN.CHEN in Gansu province. This author argues that the inscription dates to the reign of KHROLL.DEN.GTSUG.BRTSAN (704-754).
It is worth noting that many of the documents in Old Tibetan were written not only by Tibetans but also by various other ethnic groups which include: Tanguts, Han Chinese, Uighurs, etc.

### 6.3.1. Stone pillars

Many stone pillars were erected between the mid-eighth and mid-ninth centuries. About fifteen of these pillars have been preserved and described. Early examples that no longer survive were located in Chang’an (nowdays Xi’an), in 706, 733, 762, 767, 822 and at the Sino-Tibetan border in 732, 787. Below is a list of the eleven main surviving pillars and the subjects of their inscriptions (*KHA'SGANG BKRASHIS TSHE-RING* 2001).

<table>
<thead>
<tr>
<th>Content</th>
<th>Period</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oath between Tibet and China</td>
<td>764. During the reign of Thri Rälpačän⁶</td>
<td>Zhöl, in front of the Potala in Lhasa, TAR</td>
</tr>
<tr>
<td>Commitment to the Buddhist religion</td>
<td>767 (or a few years later). During the reign of Thri Songdetsän</td>
<td>Samyā monastery, Tserhang county</td>
</tr>
<tr>
<td>Description of the meeting between the prince of Kongpo, Kongkar Mangpo, and Thri Desongtsän</td>
<td>798-812. During the reign of Thri Desongtsän</td>
<td>Kongpo, Dhemo Nyingthi county</td>
</tr>
<tr>
<td>Praise to the King’s Minister, Nyang Ting Zangpo</td>
<td>798-812. During the reign of Thri Desongtsän</td>
<td>Zha Lhakhang, near Drigung</td>
</tr>
<tr>
<td>Another praise to the Minister, Nyang Ting Zangpo</td>
<td>812. During the reign of Thri Desongtsän</td>
<td>Zha Lhakhang (severly damaged)</td>
</tr>
<tr>
<td>Oath between Tibet and China</td>
<td>821-823</td>
<td>in front of the Jokhang, Lhasa</td>
</tr>
<tr>
<td>Genealogy of King Nyattri Tsänpo</td>
<td>Uncertain</td>
<td>Zha Lhakhang (carved stone)</td>
</tr>
<tr>
<td>Prayer for the development and permanence of Buddhism</td>
<td>798-815. During the reign of Thri Desongtsän</td>
<td>Gyälde Karchung Lhakhang, near the Kyichhu river, Rama sgang</td>
</tr>
</tbody>
</table>

---

⁶ Thri Rälpačän = KHRI RAL.PA.CAN; Thri Desongtsän = KHRI LDE.SRGON.BTSAN; Thri Songdetsän = KHRI SRGON.LDE.BTSAN; Gyälde Karchung = RGYAL.LDE.DKAR.CHUNG.
Accomplishments of King Thri Desongtsän

<table>
<thead>
<tr>
<th>815-836. During the reign of Thri Rälpačän</th>
<th>in front of Thri Desongtsän's tomb in the Chonggyä valley</th>
</tr>
</thead>
</table>

History of the construction of monastery and its consecration by the King

<table>
<thead>
<tr>
<th>815-836. During the reign of Thri Rälpačän</th>
<th>Tshurphu monastery (Tönlung county)</th>
</tr>
</thead>
</table>

The ten Buddhist laws and commitments

<table>
<thead>
<tr>
<th>1012. Geshe Zhangnanam Dorje Wangchuk</th>
<th>Phänpo, Gyällughä Tsuglagkhang</th>
</tr>
</thead>
</table>

### 6.3.2. Epigraphs

Many epigraphs written on rock are found throughout the Tibetic linguistic area. Some ancient epigraphs are located on the periphery of the Tibetan empire. For example, an inscription exists on a huge rock in Skardo, the capital of Baltistan in northern Pakistan. Other inscriptions have even been found in the Hunza valley in Northern Pakistan.

### 6.3.3. Inscriptions on bells

A number of inscriptions have also been incised into the surfaces of bells located in monasteries. The content of these texts is essentially religious or philosophical. The bell inscriptions are found mainly in Samyä Gegyä Tsuglagkhang, Drak-Yerpa Monastery and Thrandruk Tsuglagkhang.

### 6.3.4. Manuscripts or scrolls

About the significance of Old Tibetan texts found in Central Asia, Takeuchi expresses the following opinion (1995: 1):

“Since their discovery at the beginning of the twentieth century in the Dunhuang caves and other sites along the silk road in East Turkestan, Old Tibetan documents have proven to be an invaluable source for understanding the language, history, and culture of Tibet as well as Central Asia in general.”

More than five thousand manuscripts were discovered in Dunhuang caves (Gansu) and in the ruins of Mirang Khar (nub-chung) in East Turkestan, near the Lobnor lake along the Silk Road, mainly by Sir Aurel Stein and Paul Pelliot in the years 1907 and 1908.

Along with the Tibetan texts, there are manuscripts in other languages such as Chinese, Uyghur, Syriac, Sanskrit (in Brähmi script), Tangut, Tocharian, Khotanese,
Sogdian (in Aramaic script), Mongol (both in PHAGSPA and Mongolian scripts), Zhangzhung, Kharoshti, and Hebrew.

It is interesting to note that the majority of Old Tibetan documents were found not in the center of the Empire but at its border, on territory that now corresponds to Chinese Turkestan (Xinjiang) and the Hexi corridor. As mentioned earlier, it is also assumed that some of these texts have probably been written by non-native speakers of Tibetan.

These Old Tibetan documents are now scattered in various collections around the world: the Pelliot collection in Paris, the Stein collection in London, the Kozlov and Petrovsky collections in St. Petersburg, the German collection in Berlin, and the Otani collection in Kyoto. (See Takeuchi 1995: 14.) Most of these texts are now available at the Old Tibetan Document Online (OTDO website).

6.3.5. Wooden tablets

The majority of these tablets were located in the ruin of Miran Fortress (Xinjiang), and Terlenkha (Qinghai). A total of 389 were discovered by the British scholar Aurel Stein. The Soviet scholar Orbiash Cheshatushusich found a further six tablets. The museum of the Uyghur Autonomous Region (Xinjiang) preserves 200 tablets, of which eighty are very clear from the beginning to the end.

6.3.6. Palm-leaf documents

Old religious texts from India were written on palm leaves called Tālapatrā in Sanskrit and ཞ་ལའི་ལོ་མ་TA la in Tibetan. A number of Sanskrit palm-leaf texts are preserved in monasteries such as Sakya, Narthang, and Zhalu. The custom of writing on palm leaves was also practiced by Tibetans living in India. Aside from religious materials, palm-leaf documents deal with other topics such as grammar and history.

6.4. The literary genres

The inscriptions and the manuscripts in Old Tibetan include the following topics: chronicles and royal annals (དབེ་ཐེརང DER.THER), Buddhist texts such as sūtra (མདོ MDO) or śāstra (བསྟན་བཅོས BSTAN.BCOS), divination texts, legends and epics (such as the Rāmāyana), legal texts and private contracts, military and economic documents, medical and astrological treatises. They include many translations from
Buddhist texts in Sanskrit, as well as translations from Chinese, such as those of Chinese annals (Chin: 尚书 shangshu) and military treatises (Chin: 战书 zhanshu).

Classical Tibetan encompasses the main bulk of the Tibetan literature and one the greatest literatures of Asia. It is traditionally divided into ten traditional sciences, following the Indian tradition.

The five major "sciences" (ཨེ་ལོག་པ་རིག་བ་ལྔ་ RIG.GNAS CHE.BA LNGA): "science of sound," linguistics (ཨེ་ལོག་པ་རིག་བ་ལྔ་ SGR.RIG.PA), logic (ཀི་ལོག་པ་རིག་བ་ལྔ་ GTAN.TSHIG.PA), medicine (ཀི་ལོག་པ་རིག་བ་ལྔ་ GTO.BA.RIG.PA), "sciences of forms," "general morphology," [painting, sculpture, architecture, handicrafts, etc.] (ཕན་པ་རིག་བ་ལྔ་ BZO.BA.RIG.PA), Buddhism (ཀི་ལོག་པ་རིག་བ་ལྔ་ NANG.DON.RIG.PA).

The five minor "sciences" (ཨེ་ལོག་པ་རིག་བ་ལྔ་ RIG.GNAS.CHUNG.BA.LNGA): astrology (དཀར་རྩིས་ DKAR.RTSIS), poetics (སྙན་ངག་ SNYAN.NGAG), metrics (སྡེབ་སྦྱོར་ SDEB.SBYOR), drama (ཟློས་གར་ ZLOS.GAR), lexicography (མངོན་བརྗོད་ MNGON.BRJOD).

However, Tibetan Classical Literature is not restricted to these fields. For a detailed presentation of the various genres listed here, see e.g. Cabezon and Jackson (1996: 30-31).

- The canonical texts of Buddhism and Bön (པོ་ལྡན་པའི་ དཔྱང་བསྐུར་ RAB.GN.AS, གཟིང་བརྟེན་ ZHABS.BRTEN), Sadhana (གྲུབ་ཐབས་ GRUB.THABS), long-life prayers (ཞབས་བརྟེན་ ZHABS.BRTEN), initiation (དབང་པོའིན་ DBANG.BSKUR), fasting rituals (གསུམ་པ་ MGNS.BYUN), etc.;
- Philosophical treatises such as the Abhidharma (མིང་པོ་དོན་ MNG.N.DZOD), the Prajñāparamitā (ཕྱིག་ཐེག་ PHAR.PHYIN), the Madhyamika (དབང་ཕྲལ་ DBUM.A), the Vinaya (འདུལ་བ་ DUL.BA) and the commentaries on logics (ཚད་མ་རྣམ་འགྲེལ་ TSHAD.M.RNAM.GREL);
- Instructional or soteriological texts and treatises on various practices such as the ‘Stages of the Path’ (ལམ་རིམ་ LAM.RIM), vows (སྐོ་རྒྱུད་ SDOM.PA), precepts and instructions (གདམ་ངག་ GDO.M.N.GA), Tantra (ལོངས་ལྷག་ GSN.SNG.GA), Dzogchen (རྫོགས་ཆེན་ RDZOG.SCHEN), Mahāmudrā (ཕྱག་ཆེན་་ ལོག་PHYAG.RGYA.CHEN.MO);
- Ritual texts such for consecration (རབ་གནས་ RAB.GN.AS), offering rites (ཀྲུང་པ་ MCHOD.PA), Sadhana (གྲུབ་ཐབས་ GRUB.THABS), long-life prayers (ཞབས་བརྟེན་ ZHABS.BRTEN), initiation (དབང་པོའིན་ DBANG.BSKUR), fasting rituals (གསུམ་པ་ MGNS.BYUN).
SMYUNG.GNAS), fire rituals (SBYIN.SREG), death rituals (e.g. BAR.DO THOX.GROL), Mandala construction (DKYIL.KHOR);

- Royal chronicles (RGYAL.RABS), histories of the Dharma (CHOK.BYUNG), hagiographies (RNAS.BNAM.THAR), stories of realizations (RTOGS.BRJOD), past lives of the Buddha (SKY.E.RABS, Skt. jātaka), etc.

They also include various artistic genres such as:

- the Gesar Epic (GESAR.SGRUNG), translations of the Rāmacandra, religious songs (MGUR) folk songs, poems (SNYAN.NGAG), tales, drama (ZLOS.GAR), music (ROL.MO).

Technical texts or administrative documents as well as guidebooks and encyclopaedic works such as:

- medicine and pharmacology (GSO.BA RIG.PA), astronomy/astrology (DKAR.RTIS, NAG.RTIS, BYUNG.RTIS), grammar (SGRA.RIG.PA), mathematics (RTIS.RIG), geography/cosmology (SAKHAMS.RIG.PA) arts and crafts (BZO.BA RIG.PA) [painting, sculpture, architecture, etc.];

- catalogues (DKAR.CHIG), pilgrimage guides (LAM.YIG or GNAS.YIG), encyclopedias (such as SHES.BYAKUN.KHYAB), etc.;

- official documents of the Tibetan government (GZHUNG.YIG), legal documents (KHRIMS.YIG).

Another characteristic feature of the Tibetan literary tradition is the existence of compilations and anthologies of various authors (PHYOGS.BSGRIGS, PHYOGS.BSDEBS) and the “collected works” of a single author (GSUM.BUM, BKA.’BUM).

Concerning philological studies, one should of course mention the canonical grammatical texts SUM.CU.PA and RTAGS.KYI’UG.PA (see e.g. Tournadre 2010: 121-125) and their numerous commentaries (RNAS.BNAM.BTAN.BCO) as well as grammars (BRDA.SPROD GZHUNG (or GSUM.RTAGS).
and various types of lexicographic documents, such as ཁིག་མཛོད་ TSHIG.MDZOD ‘dictionary’, གཞི་མཛོད་ MING.MDZOD ‘glossary’, དག་ཡིག་ DAG.YIG ‘spelling guide’. About modern lexicography, grammar and dialectology in Tibetan, see e.g. JUBSTAN.SKYONG 2018; SUM.BHA.DON.GRUB TSHERING 2011 2013).

Modern Literary Tibetan (MLT) appeared during the early twentieth century. It is the literary medium in all the Tibetan speaking areas within China (in the Tibet Autonomous Region and Prefectures). It is also used by Tibetan diaspora communites throughout the world.

In the 1980s, Modern Tibetan literature had a new development and many authors began to explore modern genres of literature found in the west (Robin 2003), particularly modern ‘free’ non-versified poetry called རང་མོས་སྙན་ངག་ rangmō nyāṃgak, short stories and even novels. Scientific publications (physics, chemistry, mathematics, biology, etc.), articles or books, are written in MLT. There has also been a development of school manuals, language dictionaries, periodicals, newspapers and academic journals.

6.5. Some remarks about Old Tibetan

Old Tibetan differs from the Classical language mainly in its spelling and its vocabulary. There are also some grammatical differences, but it is difficult to summarize them in the absence of reference grammar of OT.

Most Tibetan scholars well trained in CT are able to read Old documents to a large extent. It is beyond the scope of this book to give a detailed presentation of OT. The phonology of OT has been reconstructed by Hill (2010, 2011), and we will just provide here some remarks about orthography and lexicon.

6.5.1. Spelling characteristics
We will summarize below the orthographic peculiarities of Old Tibetan. One of the general characteristics of texts written in OT is the lack of consistency in spelling.

7. We basically follow the analysis and examples in BOD-KYI RDO.RING YI GE DANG DRIL.BU’ KHABYANG by Chen (1984: 10). We will also illustrate the analysis with examples from RNAM.RGYAL TSHERING’s dictionary (2001): BOD.YIG BRDA.RNYING TSHIG.MDZOD, [dictionary of archaic expressions].
(see below the section about the ‘orthographic reforms’). Thus the same word is sometimes written in various ways even within the same text. For example ‘(water) spring’ may be spelled as: ཚུ་དམྱིག་, ནའ་དམྱིག, དུ་མྱིག or མའ་དྲེ་དམྱིག.CHudder.

List of the main spelling differences between Old Tibetan and Classical Tibetan:

- The vowel I is often written in a reversed way: བ. However, this graphic variant does not seem to have any phonemix value (see Hill 2010a).
- There is confusion between aspirated and non-aspirated sounds. Aspirated sounds of CT are often written without aspiration while non-aspirated sounds of CT are written with aspirated consonants in OT. According to Hill (2007) aspiration “had begun to be phonemic” but these fluctuations could reflect the fact that aspiration had not acquired yet a fully phonemic status.

Whatever it may be, note that the orthography is not entirely systematic and the spelling of OT is sometimes equivalent to the CT spelling.

Example: ཚུ་དམྱིག་(OT) ‘one’ ↔ ནའ་དམྱིག (CT), བུ་མོ་ཆུང་ངུ་གཅིག་(OT) ‘a little girl’ ↔ བུ་མོ་ཆུང་ངུ་གཅིག (CT), བུ་མོ་མཐོང་ངུ་གཅིག (CT)

- Variation between voiced and voiceless sounds:

Example: ན་མིག (OT) ‘and’, མ་མིག (CT) ‘and’.

- The letters N and D are sometimes interchanged, when occuring as radicals or suffixes.

Ex: མདོ་CHED.PO (OT) ↔ མདོ་CHEN.PO (CT) 'big'.
▪ The second suffix ད་དྲག་ (DA) or DRAG is often used after N, R, L with verbs and more rarely with nouns.

Ex: གྱུརད་ (GYURD) ↔ གྱུར་ (GYUR) ‘to change’, སྩལད་ (STSALD) ↔ སྩལ་ (STSAL) ‘to grant, to bestow’, དབྱརད་ (DBYARD) ↔ དབྱར་ (DBYAR) ‘summer’ (CT).

▪ The labial M when followed by the vowels I or E becomes MYI and MYE.

Ex: མྲི་ (MI) ↔ མྲི་ (MYI) (CT) ‘human being’, མེད་ (MED) ↔ མེད་ (MYED) (CT), negation of ‘to exist’.

▪ The རི་ ‘r’ letter is often used after a final vowel.

Ex: རི་ (RI) (OT) ↔ རི་ (RI) (CT) ‘mountain’.

▪ Some combinations of graphs that were later abandoned are attested. They include the combination སྩ་ (STS), which is frequent in Old Tibetan and is simplified as S in Classical Tibetan.

Ex: ལ་སྩོགས་པ་ (LA) STSOGS PA (OT) ↔ ལ་སོགས་པ་ (LA) SOGS PA (CT) ‘and so on’, བསྩོད་ནམས་ (BSTSOD NAMS) (OT) ↔ བསོད་ནམས་ (BSOD NAMS) (CT) ‘merit’.

Other combinations include e.g. RHY and ZR.

▪ The prefixed letters G, D, B and superscribed letters R, L, S are often interchanged.

Ex: རྩང་པོ་ (RTSANG) ↔ གཙང་པོ་ (GTSANG) (CT) ‘river’; སྙི་ (SNYI) (OT) ↔ རྙི་ (RNYI) (CT) ‘trap’; དཔུར་ (DPUR) (OT) ↔ སྤུར་ (SPUR) (CT) ‘corpse’; ལུང་ (RNGUL) (OT) ↔ དངུལ་ (DNGUL) (CT) ‘silver’.

▪ In some cases, the prefixed letter D or G are added.

Ex: དམྱིག་ (DMYIG), གམྱིག་ (GMYIG) (OT) ↔ མིག་ (MIG) (CT) ‘eye’; བསྙིན་ (BSGNS) (OT) ↔ ནག་ (NGAG) (CT) ‘speech’.

▪ In some cases, the subscribed r is replaced by a y.

Ex: དི་ (GRI) (OT) ↔ དི་ (GRI) (CT) ‘knife’.
Modern reflexes of the glide y, which correspond to the archaic form, are still found in some modern languages such as Amdo, Thewo Tö or Čone.8

6.5.2. Lexical characteristics

There are grammatical differences between Old Literary Tibetan and Classical Literary Tibetan but the main discrepancies occur at the lexical level. Let us illustrate the lexical differences between the two periods with examples from the Lishi Gurkhang, a fifteenth century text and the Bod, Yig BRDA, RNying TSIG, MDZOD (‘dictionary of old terms’). It should be noted here that some of the words in this list were still marginally used in CT.

CHART VI.2. – Lexical differences between Old Tibetan and Classical Tibetan

<table>
<thead>
<tr>
<th>Meaning</th>
<th>Old Tibetan</th>
<th>Classical Tibetan</th>
</tr>
</thead>
</table>
| ‘to ask’         | བརྡ་རྙིང་   | རྒྱུན་གསོལ་བོད་ཡིག | 'to think' | བྲེལ་པ་ |  སྟབས་ |  བསྟེ། |  གྲན་པ་ |  རོག་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་པ་ |  འབྲེལ་pagination-error

8. Note however that all the dialects which have reflexes /c, ch, j/ or /tɕ, teh, da/ for the CT combinations KR/KHR/GR (e.g. 'gro' ‘to go’ or ‘gro’ ‘knife’) do not necessarily correspond to archaic forms. This is the case for example in Southern Kham (Gyalthang) in which these forms correspond to innovations and are derived from KR/KHR/GR (and not KY/KHY/GY). See Suzuki (2020).
There are also grammatical discrepancies between OT and CT but this topic still needs further research. The main differences lie in the verb forms. It seems that irregularities existing in OT were later systematized in CT (see section 6.6.3 below). The evidential system started to develop gradually in CT but was very limited in OT (see Shao 2016).

### 6.5.3. The three linguistic reforms

At the time of the Tibetan Empire and during the following centuries, Literary Tibetan underwent three important reforms, called བཀས་བཅད་རྣམ་པ་གསུམ་ (BKAS.BCAD RNAM.PA GSUM), lit. 'the three (linguistic and religious) decrees'. The first reform took place from the time of Thrisong Detsän to Thritsuk Detsän, also known as Rälpačän during the eighth and ninth centuries. Many grammarians and translators, including those known as the Great Nine Lotsawas 'translators' (ལོ་ཙ་བ་རབ་དགུ་ (LO.TSA.BA RAB DGU)), took part in the reform. Its main objective was the standardization of the terminology used to translate Buddhist texts and commentaries, which caused important lexical treatises to be composed during this period. Among these was the 'First volume (of the treatise) on the formation of words' (སྒྲ་སྦྱོར་བམ་པོ་དང་པོ་ (SGRA.SBYOR BAM.PO DANG.PO)).

The second reform happened at the time of Thritsuk Detsän’s reign during the ninth century. Many lotsawas were invited from India to Tibet to discuss translation issues. They composed a bilingual Sanskrit-Tibetan dictionary, called བྱེ་བྲག་ཏུ་རྟོག་བྱེད་ཆེན་པོ་ (BYE.BRAG TU.RTOG.BYED CHEN.PO) in Tibetan or ‘Mahāyutpatti’ in Sanskrit. The great translator Zhang Yeshede and the Indian Panditas Jinamitra and Danashila composed the treatise called སྒྲ་སྦྱོར་བམ་པོ་གཉིས་པ་ (SGRA.SBYOR BAM.PO GNYIS.PA) ‘The second volume (of the treatise) on the formation of words’, which was a compilation and correction of all the previous lexicographic works. The ‘second volume (of the treatise) on the formation of words’, which used the "second reform orthography," was later
included in the བསྟན་འགྱུར་ Tängyur, the famous collection of commentaries on Tibetan Buddhist canonical texts.

Three measures of standardization were decided:

- grammatical and lexical standardization of the transcription of Sanskrit;
- fidelity to the original text;
- clarity and simplicity of the translation.

There were four types of translation and transcription:

a) phonetic transcription (སྒྲ་སྒྱུར་ SGRA.SGYUR);

b) semantic translation (དོན་སྒྱུར་ DON.SGYUR);

c) calques from Sanskrit (ཐད་སྒྱུར་ THAD.SGYUR);

d) transformation of the original meaning (བཅུས་སྒྱུར་ BCUS.SGYUR).

There were three reasons for using a phonetic transcription and not a translation:
a) when the meaning of the original language was unclear, the transcription was added;
b) when there was some ambiguity in the Sanskrit meaning; c) when there was some ambiguity in the Tibetan meaning.

Four important points were given special attention: the main vocabulary of religion and philosophy of the various schools should be described very precisely, all Tibetans, regardless of their regional origins, should use the same terminology, all neologisms should be reported to the king and approved by the main lotsawas. The transcription of mantras should also receive official approval from the king.

During the ‘second reform’ བཀས་བཅད་བར་མ་ BKAS.BCAD.BAR.MA, unclear ancient terminology was abandoned and replaced by new terms. An important orthographic reform, which is described in the text མཁས་པའི་དགའ་སྟོན་ MKHAS.PA.I.DGA.STON, also took place. Several letters or combinations were abandoned. They include the glide ɣ subscribed to the M (ི M), the ‘second suffix’ Ո called DRAG, the final ’A CHUNG and the reversed vowel I was dropped.

The ‘third reform’ བཀས་བཅད་གསུམ་པ་ BKAS.BCAD.GSUM.PA, began in the tenth century, after the collapse of the Tibetan Empire, starting in Western Tibet during the reign of Lha lama Yeshe Ö. According to historiographic tradition, this reform was actually a gradual series of modifications to Tibetan orthography and lexicography,
which were proposed by a lineage of 170 lotsaws over a period of 470 years. The first lotsawa taking part was lotsawa Rinchen Zangpo and the last was Zhalu lotsawa Chökyong Zangpo, who lived in the fifteenth century (see e.g. Tuttle & Schaeffer 2013).

What is essential here is that the lexical and orthographic standardization promoted by the so-called “three reforms” extended over six or seven centuries! The process only achieved “complete unification” during the fifteenth century, as suggested by the dictionary LISH'I GUR.KHANG composed by Kyoktön Lotsawa Rinchen Trashi in 1476.

Finally, one ought to mention a striking characteristic of the translation strategy adopted by the Tibetan lotsawas. Instead of quoting the Sanskrit names of persons, deities and places phonetically, they chose to ‘translate’ them into Tibetan.

Thus for example, the towns Rajgir and Vaishali in Bihar, places often cited in Tibetan Buddhist texts, are called respectively རྒྱལ་པོའི་ཁབ་ RGYAL.POT.RI KHAB and ཡངས་པ་ཅན་ YANGS.PACAN. Buddha’s favorite retreat, Gilbrakṣa, is called བྱ་རྒོད་ཕུང་པོའི་རི་ BYARGOD PHUNG.POT.RI ‘Vulture peak’, etc. The names of Buddha, Ānanda, Nāgarjuna and Śāntideva (Sāntideva) are rendered respectively into Tibetan as: སངས་རྒྱས་ SANGS.RGYAS, དབང་ཕྱུག་ DBANG.PHYUG, ཁྱབ་འཇུག་ KHYAB.JUG and དཔལ་ལྡན་ལྷ་མོ་ DPAL.LDAN.LHAMO.

### 6.6 Essential morphological features of Classical Tibetan

#### 6.6.1 Nominal morphology

We will briefly present here the main characteristics of Classical Tibetan grammar. For more details, see e.g. Kesang Gyurmé’s རབ་བསལ་མེ་ལོང་ RAB.BSAL.ME.LONG (1981) or its French translation and commentary, Le Clair Miroir (1992) or its Chinese version (实用藏文文法) and Beyer’s Classical Tibetan (1992), Hahn (1996),

---

9. Of course ‘complete unification’ or ‘complete standardization’ is a relative notion. No language, even ‘major languages’ such as English, French or Japanese are completely standardized.

10. Many commentaries in italics were added by N. Tournadre in the 1992 and 1994 editions to clarify some issues related to the traditional terminology.
Schwieger (2006). The aim of the present description is to provide a synthesis of the Classical language’s morphology in order to compare it with the grammar of modern languages and dialects.

The ancient prefixes and suffixes found in Pre-Tibetic (see Chapter 5) are no longer syllabic in Old and Classical Tibetan. They correspond to single consonants, which may be called ‘formatives’ following Beyer (1992). For example, from the verb ניא (‘to be sick’), it is possible to derive the noun ניאד ‘disease’ through the formative ‘D’.

Both sound alternations and formative derivations have sometimes been called “inner derivation” because they occur within a single syllable. “Outer derivation” refers to affixes (prefixes or suffixes) that are attached to a syllabic stem, following Beyer (1992: 111).

“We will use the term inner derivation to refer to those processes of derivation that operate within the syllable, using such formatives as prefixed s (as in sring “be long”, sring “make long”), suffix d (as in dro “be warm” drod “warmth”), and in fixed voicing (as in than “emerge”, don “eject”). We will use the term outer derivation to refer to those processes of derivation that operate outside the syllable, using either such syllabic formatives as –pa “one having to do with” (as in mda’ “arrow”, mda’-pa “archer”) or reduplication of the syllable as a whole (as in rnyamb-po “bright” rnyam-rnyam “dazzling”).”

Inner alternations and formatives belong either to derivational or inflectional morphology. Derivational morphology deals with the lexicon (i.e. the creation of new words), while inflectional morphology conveys grammatical meanings such as tenses, cases, and so forth. Whether belonging to derivational or inflectional morphology, formatives and alternations found in CT reflect an archaic stage of the language that was no longer productive in OT.

The various modern languages and dialects are all derived from OT and thus usually share the same inner formatives. The outer affixation (only through suffixes), however, is probably a later phenomenon. Thus, the various modern Tibetic languages and dialects do not necessarily share all the same suffixes or prefixes (see Chapter 7).

6.6.1.1. The initial formatives
In OT and CT, one finds mainly the ’S’- “causative” formative, the ’S’ “animal” formative, the formative ’M’- “human body” and ’M’- “honorific.” These formatives are derived from the prefixes that existed in Proto-Tibetic (see Chapter 4).
Formative N:་བྲུག་ 'tiger'; རྒྱུ་ 'SPYANG.GU' 'wolf'; འོ་ 'SDOM' 'spider'; སྡིག་པ་ 'SKYIN' 'ibex'; ཡང ཤློ་ 'SRANG.BU' 'fly'; བུ་ 'SIRUL' 'snake'; ཐོ་ 'SKYIN' 'crab'.

Formative M: གཅོ་ 'MCHIN.PA' 'liver'; གཅོ་ 'MKHIS.PA' 'bile'; གཞོ་ 'MGQ' 'head'; གཅོ་ 'MGRIN.PA' 'neck'.

6.6.1.2. The final formatives

In CT just as in OT, nouns are often derived from the verb by adding a consonant 'D', 'N', or 'S', sometimes followed by a suffix.

Formative D: བ་ 'NA' 'to be sick' → བུ་ 'NAD' 'disease'; བུ་ 'LU' 'to cough' → བུ་ 'LUD.PA' 'spatum'; བུ་ 'TSHA' 'to be hot' → བུ་ 'TSHA.D.PA' 'hot temperature'; བུ་ 'RGQ' 'to become old' → བུ་ 'RGAD.PO' 'old person'; བུ་ 'RTSE' 'to play' → བུ་ 'RGAD.PO' 'game'; བུ་ 'NGU' 'cry' → བུ་ 'NGUG.D.MO' 'sob'; བུ་ 'LTA' 'to look at' → བུ་ 'LTA.D.MO' 'show, spectacle'; བུ་ 'DRO' 'warm' → བུ་ 'DRO.D.MO' 'heat'.

Formative N: བུ་ 'RKU' 'to steal' → བུ་ 'RKUN.MA' 'thief'; བུ་ 'ZIN' 'to eat' → བུ་ 'ZAN' 'food'; བུ་ 'GRO' 'to go' → བུ་ 'GRO.PA' 'guest, visitor'; བུ་ 'GDA' 'to be there' → བུ་ 'GDA.MO' 'seat'; བུ་ 'RGQ' 'to become old' → བུ་ 'RGAN.PA' 'old person'; བུ་ 'GCI' 'to urinate' → བུ་ 'GCI.PA' 'urine'; བུ་ 'BSHA' 'to slaughter' → བུ་ 'BSHA.MO' 'butcher'; བུ་ 'BSU' 'to welcome' → བུ་ 'BSUN.MA' 'lady who welcomes guests'; བུ་ 'RZU' 'to pretend' → བུ་ 'RDZUN' 'lie'; བུ་ 'RGYU' 'to run, move' → བུ་ 'RGYUN' 'flow, stream'.

Formative N: བུ་ 'ZIN' 'to eat' → བུ་ 'ZAN' 'food'; བུ་ 'LIA' 'to look at' → བུ་ 'LIA.D.MO' 'omen'; བུ་ 'KR' 'to bathe, wash' → བུ་ 'KRUS' 'bath'; བུ་ 'PHYUGS' 'to be rich' → བུ་ 'PHYUGS' 'cattle'; བུ་ 'BSAM' 'to think' → བུ་ 'BSAM.SEM' 'mind'; བུ་ 'SKYOB' 'to protect' → བུ་ 'SKYOB.SKYABS' 'protection'.

6.6.1.3. The nominal affixes

The main nominal suffixes found in CT (and in OT) are བུ་ 'PO, བུ་ 'BO, བུ་ 'MO, བུ་ 'PA, བུ་ 'BA, བུ་ 'MA. Other less frequent suffixes include བུ་ 'KA (and its variants བུ་ 'KHA and བུ་ 'GL), བུ་ 'CHA, བུ་ 'SO, as well as the diminutive བུ་ 'BU and its variants བུ་ 'GU and བུ་ 'U. Additionally a few words, which mostly correspond to kinship terms, may have a prefix: བུ་ 'PA.
Ex: རྒད་པོ RGAD.PO ‘old person’; རྟེ་ལེན་LRTA.MO ‘show, spectacle’; མཆིན་པ་ MCHE.MA ‘liver’; མཁལ་མ་ MKHA.MA ‘kidney’; སྤྱང་གུ SPYANG.GU ‘wolf’; མ་ ་ PAMA ‘mother’; ལེ་ ་ PAJO ‘elder brother’; མ་པ ་ PHA.MA ‘father’; བ ་ PAKHU ‘paternal uncle’; བ ་ PAZANG ‘maternal uncle’; བ ་ PANE ‘paternal aunt’, etc.

6.6.1.4. The case system

For more than a millennium, the Tibetan grammatical tradition has used categories largely based on the Sanskrit model to describe Classical Tibetan. This choice was motivated by both cultural and religious reasons; however, it has generated some problems for grammatical analysis since the two languages belong to different families – Indo-European and Sino-Tibetan – they cannot be described efficiently with the same linguistic categories. By choosing Sanskrit as their model, the Tibetan grammarians were ‘obliged’ to use concepts and categories that are not relevant for Tibetan and lacked categories that are fundamental to the description of their language. The case system is a good example of the problems that arise when using Sanskrit grammar as a model for the Tibetan language.

Traditionally, the Tibetan case system (རྣམ་དབྱེ་ RNAM.DBYE) is presented as having eight cases, which correspond to the Sanskrit cases (see Chart VI.3.).

An analysis exclusively based on Tibetan morphology and syntax yields ten cases for Classical Tibetan, see Tournadre (2010) and Hill (2011, 2012): absolutive (Ø); ergative11 (གྱིས GYIS and its variants: ཨི་ YIS, ཨི་ GIS, ཨི་ KYIS, ས ཀ S); dative (ལ་ LA); purposive also called “terminative” (དུ་ DU and its variants: སུ་ SU, ཙ R, སུ་ RU, ཡ TU); elative (ནས་ NAS); genitive (གྱི་ GYI and its variants: ཨི་ YI, ཨི་ GI, ཨི་ KYI); locative (ན་ NA); ablative (ལས་ LAS); associative (དང་ DANG); and comparative (བས BAS).

11. In CT, the “ergative” is also used for the instrumental functions.
12. In Tournadre (2010), the variant ཡ R was classified together with the dative ས LA, but it is generally grouped together with the purposive (or terminative) (see Hill 2011, 2012a). The fluctuations in some adverbial functions between ཡ R and ཙ RU as well as data from Purik (Zemp 2018) advocate for the grouping with the terminative case.
Traditional labels for the Sanskrit and Tibetan grammatical cases

<table>
<thead>
<tr>
<th>Sanskrit</th>
<th>Literary Tibetan</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kāraka कारक</td>
<td>RNAM DBYE རིན་མོ་དབྱེ</td>
<td></td>
</tr>
<tr>
<td>1) nominative</td>
<td>kartā कर्ता</td>
<td>NGO.BO.TSAM ཉགོ་བོ་ཙམ་</td>
</tr>
<tr>
<td>2) accusative</td>
<td>karma कarma</td>
<td>LAS.SU.BYA.BA བྲས་སུ་བྱ་བ་</td>
</tr>
<tr>
<td>3) instrumental</td>
<td>karāṇa कराण</td>
<td>BYED.SGRA བླེད་སྒྲ</td>
</tr>
<tr>
<td>4) dative</td>
<td>sampradāna सम्प्रदान</td>
<td>DGOS.CHED ཉགོས་ཆེད</td>
</tr>
<tr>
<td>5) ablative</td>
<td>apādāna अपादा</td>
<td>'BYUNG.KHUNGS དབྱུང་ཁུངས</td>
</tr>
<tr>
<td>6) genitive</td>
<td>sambandha संबन्ध</td>
<td>'BREL.SGRA བི་རེལ་སྒྲ</td>
</tr>
<tr>
<td>7) locative</td>
<td>adhikaraṇa अधिकरण</td>
<td>GNAS.GZHI གནས་གཞི</td>
</tr>
<tr>
<td>8) vocative</td>
<td>sambodhana सम्बोध</td>
<td>BOD.SGRA བོད་སྒྲ</td>
</tr>
</tbody>
</table>

The CT case markers are neither inflection like classical cases, e.g. in Sanskrit, Greek, or Russian, nor adpositions like those found in French or English (e.g. the prepositions ‘à’ in French or ‘to’ in English): They are clitics and attach at the end of a noun phrase (concerning clitics and affixes in TB languages, see Genetti 1993). The case markers never occur independently. Another difference directly related to their

13. I give here the Tibetan transliterations of Sanskrit cases according to the Tibetan tradition. The earlier Sanskrit tradition mentions only six kārakas (see Verhagen 2001). These Sanskrit transliterations are rarely mentioned in the commentaries and usually replaced by their Tibetan equivalents.

14. LAS.SU.BYA.BA can be translated literally "activity towards a work/for a work" (see Zeisler 2006b: 59).
The clitic nature is that Tibetan cases occur only once for each NP (Noun Phrase), unlike ‘classical’ case systems of Sanskrit, Latin, Greek or Russian, where a case, for example dative, is marked on each constituent of the NP whether nouns, adjectives, demonstratives, numerals, quantifiers or pronouns.

Another consequence of the clitic nature of the cases is that the various constituents of the NP never undergo any morphological variation. The only morphological variation is related to the clitic morpheme itself, which may undergo a variation depending on the final consonant or vowel of the preceding word.

Some case markers are clearly allomorphs and represent formal variations of a single morpheme in a certain environment. The variation is linked to an old morphophonological rule and does not reflect any difference in terms of grammatical semantics.

This is for example the case for GI, KYI, GYI, ʔ and YI, which are allomorphs of the same genitive case as well as GIS, KYIS, GYIS, ʔIS and YIS, which are allomorphs of the same ergative (or ‘agentive’) case.

Zero marking (Ø) should also be considered as a case marker although it is formally void. The reason is that the absolutive case, traditionally referred to asngo.bo.tsam, plays an essential role in the ergative constructions. The absolutive is used for both the unique participant of an intransitive construction and the patient of a transitive construction.

Concerning their syntax and semantics, the main characteristic of CT cases is that they are multifunctional, transcategorial and sometimes optional (LaPolla 1995; Tournadre 1997, 2010; DeLancey 2011a). They are multifunctional in the sense that every case has a wide array of functions.

The various cases of CT indicate grammatical roles, when occurring after a noun or a NP. They indicate its grammatical role or function such as Agent, Patient, Beneficiary, Instrument, Goal, Source, etc.

The case markers also have connective functions, when placed after a verb or a nominalized verb. They function as coordinators or subordinators. Apart from these main functions, one also encounters adverbial functions after nouns or adjectives and postpositional functions mainly after nouns.
6.6.2. Adjectival morphology

In both Old and Classical Tibetan, adjectives are often derived from the verb by adding an adjectival suffix བོ PO, བོ BO, མོ MO, མོ PA, བ མོ BA, བ མོ MA (a subclass of the nominal suffixes) and གུ NGU: གསལGSAL 'to be clear' → གསལཔོGSAL PO 'clear'; དིིམZHIM 'to be tasty' → དིིམཔོZHIM PO 'tasty'; རིངRING 'to be long' → རིངཔོRING PO 'long'; ོདཀརDKAR 'to be white' → ོདཀརཔོDKAR PO 'white'; ཚུངCHUNG 'to be small' → ཚུངངུCHUNG NGU 'small'; རྡུངTHUNG 'to be white', etc.

Sometimes, when the stem ends in a vowel, the suffix is preceded by the formative 'N added to the stem (as we have seen in 6.6.1.2):

ཧྲ RGA 'to become old' → ཨྲ RGANPA 'old'; བསམBSAM 'to raise' → བསམPA 'alive'; རྡོ RNO 'to be sharp' → རྡོPA 'sharp'; བློ SNGO 'to be blue' → བློ SNGONPA 'blue'; མོི MTHO 'to be high' → མོིPA 'high'; སྤྱོར TSHO 'to be fat' → སྤྱོརPA 'fat'; སྤྱོརCHEN 'to be big' → སྤྱོརPA 'big', etc. Note that the reduplication of adjectival roots rarely occurs in CT, however, it is frequent in Modern Literary Tibetan and in the modern languages (see 8.1.7).

From a morphosyntactic point of view, one also finds in Classical Tibetan compound adjectives such as: སྦྱོར་དུ་རྙེད་ཡིད་ཡོན་ཏན་YID.DU. 'attractive', 'handsome', 'charming' (lit. 'coming to' or 'fitting the mind'); སྐོབས་ཡུལ་འདས་པ་BSAM.YUL.BTAB.PA 'unthinkably' (lit. 'overcoming the mind'); ཐོབས་གྱིས་བ་པ་BSAM.GYIS.BTAB.PA 'inconceivably' (lit. 'not embraced by the mind'); བསྲེལ་འདུལ་རྙེད་ལྡན་པ་KHYAD.DU.PHAGS.PA 'superior, sublime' (lit. 'particularly noble'); སྷེ་བཅོ་སྐྱེལ་RNAB.MI.'GRO' 'unpleasant' (lit. 'not fitting the ear'), etc.

6.6.3. Verbal morphology

Morphological alternations of consonants or vowels (apophony) are found in Old and Classical Tibetan. These alternations can be compared to those found in Germanic languages. For example, in English sing, sang, sung or think, thought, thought correspond to present, past, past participle. In a similar way, although more complex, one encounters in Tibetan verbs that have four inflections འདེབས་DEBS, བཏབ་BTAB, གདེབས GDAB, འན་འབོད THOBS 'to plant', corresponding respectively to "present," "past," "future," and
“imperative” according to the traditional terminology. However, these forms do not only correspond to tenses but also to aspects and modalities (see Zeisler 2004).

The inflexions were probably originally due to verbal affixes (prefixes and suffixes) which progressively merged with the verb (see also Bialek 2020, 2021; Jacques 2021 for recent debates).

There are, in fact, two types of verbal inflexions. The first type, which is pervasive in all the Sino-Tibeto-Burman languages, indicates the opposition between causative (mostly transitive) and anticausative or resultative (mostly intransitive) verbs. For example, བཅད་ ‘to cut’, ཕེད་ ‘to be cut’, བསྒྱུར་ ‘to change, translate’, གྱུར་ ‘to be changed’, མོང་ ‘to boil’ vs. མོལ་ ‘to be boiled’, etc. (see the list of 200 verb pairs in Kesang Gyurmé 1992 or in Tournadre & Sangda Dorje 2003).

The second type of verbal inflexions (such as DEB, BTAB, GTAB, THOBS mentioned above) indicates tense, aspect and modality (TAM). The inflexions are irregular and unpredictable. While some verbs are invariable, many have up to four stems (or inflexions) to indicate the present, future, past and imperative. The verbal stems show variations both in vowels and consonants. The forms are often based on different stems, although etymologically related. The following chart shows some examples of stem variations.

**Chart VI.4. – Stem variations according to the tenses and modality**

<table>
<thead>
<tr>
<th></th>
<th>past</th>
<th>present</th>
<th>future</th>
<th>imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘to make, do’</td>
<td>བྱས་</td>
<td>བྱེད་</td>
<td>བྱ་</td>
<td>བྱོས་</td>
</tr>
<tr>
<td>‘to cut, decide’</td>
<td>བཅད་</td>
<td>བྲེད་</td>
<td>བྲེད་</td>
<td>བྲེད་</td>
</tr>
<tr>
<td>‘to plant’</td>
<td>བཏབ་</td>
<td>ཐེབས་</td>
<td>ཐེབས་</td>
<td>ཐེབས་</td>
</tr>
<tr>
<td>‘to give, offer’</td>
<td>བསྲལ་</td>
<td>བསྲལ་</td>
<td>བསྲལ་</td>
<td>བསྲལ་</td>
</tr>
<tr>
<td>‘to eat’</td>
<td>བཟའ་</td>
<td>བཟའ་</td>
<td>བཟའ་</td>
<td>བཟའ་</td>
</tr>
<tr>
<td>‘to lay down, to sleep’</td>
<td>ཉལ་</td>
<td>ཉལ་</td>
<td>ཉལ་</td>
<td>ཉོལ</td>
</tr>
<tr>
<td>‘to change’ (involuntary)</td>
<td>གྱུར་</td>
<td>གྱུར་</td>
<td>གྱུར་</td>
<td>གྱུར་</td>
</tr>
<tr>
<td>‘to demand’, ‘to offer’</td>
<td>གསོལ་</td>
<td>གསོལ་</td>
<td>གསོལ་</td>
<td>གསོལ</td>
</tr>
<tr>
<td>‘to see’</td>
<td>མཐོང་</td>
<td>མཐོང་</td>
<td>མཐོང་</td>
<td>མཐོང</td>
</tr>
</tbody>
</table>
Some verbs such as 'to make, to do', 'to plant', 'to cut' have four stems, some verbs such as 'to give, to offer' have three stems. Other verbs such as 'to sleep' and 'to change' have only two and 'to demand, to offer' is invariable.

Among the differences between CT and OT verb forms, one must mention the final second suffix D, which was still written in OT (for the past): བསྟིལ་ STSALD, དྱུརད་ GYURD, སྐྱེལ་ GSOLD (compare with the above chart).

The form BSTSAL (CT) with a B prefix in the past, instead of STSALD attested in OT was probably invented for the systematization of verb forms. Another clear case of this phenomenon is ར་ ZA 'eat' which had རེས་ ZOS (past) in OT while CT uses སྦྱར་ BZAS. Sometimes, the verbal inflexion used originally in Old Tibetan for the causative (transitive) / anticausative or resultative (intransitive) opposition have become part of the tense paradigm of the verb in Classical Tibetan. For example, in the Dunhuang documents, རིམ་ 'BUL and སྦེ་ PHUL correspond respectively to the causative and the anticausative (or resultative) forms of the verb 'to offer', while in Classical Tibetan 'BUL and PHUL correspond respectively to the present and the past stem.

Based on an article written by Li (1933), Coblin (1976) has shown that it was possible to reconstruct eight paradigms of verb forms for CT.

<table>
<thead>
<tr>
<th></th>
<th>Present (1)</th>
<th>Past (2)</th>
<th>Future (3)</th>
<th>Imperative (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td><code>-R</code></td>
<td>b-R-s</td>
<td>b-R</td>
<td>R-s °o</td>
</tr>
<tr>
<td>II</td>
<td><code>-R-d</code></td>
<td>b-R-s</td>
<td>b-R</td>
<td>R-s °o</td>
</tr>
<tr>
<td>III</td>
<td>R-d</td>
<td>b-R-s</td>
<td>b-R</td>
<td>R-s °o</td>
</tr>
<tr>
<td>IV</td>
<td>g-R</td>
<td>b-R-s</td>
<td>b-R</td>
<td>R-s °o</td>
</tr>
<tr>
<td>V</td>
<td>g-R</td>
<td>b-R</td>
<td>d-R</td>
<td>R-s °o</td>
</tr>
<tr>
<td>VI</td>
<td><code>-R-d</code></td>
<td>b-R</td>
<td>d-R</td>
<td>R-s °o</td>
</tr>
<tr>
<td>VII</td>
<td><code>-R</code></td>
<td>b-R-s</td>
<td>d-R</td>
<td>R-s °o</td>
</tr>
<tr>
<td>VIII</td>
<td><code>-R-d</code></td>
<td>b-R-s</td>
<td>d-R</td>
<td>R-s °o</td>
</tr>
</tbody>
</table>

For some verbs, it is necessary to postulate the existence of two stems.

15. For example in Takeuchi 1995, text 25, line 3 and 4: LHA RIS GYI RANG SGOR BRE PHUL MYI CHAD BAR / BUL / DER MA PHUL LAM GYI GYUS ZHIG TSHAL NA.
PART 2 – CHAP 6. Literary Tibetan and its evolution

འདེབས་’DEBS(1), བཏབ་’BTAB(2), གདབ’GDAB(3), ཐོབས’THOBS(4) ‘to plant’; stems: -dollar and -thab.


འགེགས་’GEGS(1), བཀག་’BKAG(2), དགག་’DGAG(3), ཁོགས་’KHOGS(4) ‘hinder, lock up’; stems: -gag and -khag.

The G prefix found in གདབ’GDAB is not found for the future in the reconstructed paradigm (in the chart above). Coblin (1976) proposed that the original future prefix G was D and that “the pre-initial D later dissimilated to G before the stem initials C(¬CH), NZ, T(¬TH), D, N, TS, (¬TSH), ZH, Z, Y, SH and S” (Coblin 1976: 56).

In Classical Tibetan, verbs with four stems are usually controllable (volitional)16 and transitive. Most of non-controllable (volitional) verbs have only two stems or are invariable.

In some exceptional cases, the verbal forms of Classical Tibetan are derived for entirely different verbs. This suppletive strategy may be illustrated by the following verbs:

▪ གྲོ་’GRO ‘to go’ (present, future), ཕོང་’SONG ‘to go’ (past and imperative) or ཐོན་’PHYIN ‘to go’ (past).

▪ ཀང་’ONG ‘to come’ (present, future), འོངས་’ONGS ‘to come’ (past), དོག་’SHOG ‘to come’ (imperative).17

As we will see (chap 8), in some modern Tibetic languages, the suppletive strategy to indicate tenses has been used with various other frequent verbs.

In CT, tenses, aspects and modalities are marked not only by the inflectional forms, which correspond to an archaic morphology, but also by a system of verbal auxiliaries and connectives.

16. Although DeLancey, Tournadre and others authors used previously the term “volitional,” the term “controllable” is preferable. See the comments in the section 8.3.4.2 on Controllability.

17. The latter is derived from the verb ཁོ་ ‘go, come’ which imperative ཁེ་ is sometimes found in early Classical. Such suppletive verbs are also found in modern dialects for verbs such as ‘go, come, give’, etc. see 8.3.10.
The verbal auxiliaries are essentially used in final clauses, while the connectives are used after subordinate (or non-final clauses).

In the final clause, the verb may occur alone but it is frequently followed by an auxiliary.

In the latter case, the verb is followed by an auxiliary which may itself be preceded by a relator (REL). The relators correspond either to a nominalizer (NMLZ) (sometimes associated to a nominal case) or to a connective (CO) (see Oisel 2013; Tournadre and Konchok Jiatso 2001).

Thus we find the following constructions:

a) (NEG)-V[flex]

b) (NEG)+ V[flex]+(REL)+(CASE)+AUX

c) V[flex]+(REL)+(CASE)+ (NEG)+AUX

The verb is the only compulsory element and may be inflected as noted in (a) by the parenthesis [flex]. As we have seen above, many verbs are invariable.

In the subordinate (or non-final) clause, the verb is usually followed by a connective. One often finds series of clauses linked together by connectives (see section below). Thus, the complex sentence structure may be represented in the following way:

(SN)+V[flex]-CO, (NP)+ V[flex]-CO, (NP)+V[flex]+(REL+Case)+AUX

Each non-final clause is followed by a connective. The tense which is often marked by the auxiliary (possibly in combination with the verbal inflection) usually appears in the final clause.

6.6.4. Auxiliary verbs

As we have seen in 6.6.3, in order to indicate tenses, aspects and modalities (TAM), Old and Classical Tibetan have developed a system of auxiliary verbs postponed to the main verb, in addition to the verbal inflexions. The auxiliaries in CT play a considerable role in the TAM marking. Evidentiality and epistemic modality are also marked by auxiliaries, but they play a rather marginal role in Classical Tibetan (see Oisel 2013; Hill 2013; Zeisler 2018b).
The main auxiliaries used in CT are:

- **Copulative verbs:** ཨིན་ ‘to be’, ལགས་ ‘to be (Hon)’, རེད་ ‘to be’.
- **Existential and location verbs:** འདུག་ ‘to sit’, ཕྱིན་ ‘to exist’, བདོག་ ‘to have, to stay’.
- **Motion verbs:** འོང་ ‘to come’ and their variants འོངས་, རྭོག་, བྱུང་, བྱེད་, མཛོད་, བགྱིད་.
- **Activity verbs:** བྱེད་ and its inflected forms: རེད་, བགྱིས་, བདོག་, མཛོད་, བགྱིད་.
- **Modal and other types:** འདྲ་ ‘be similar’, སྣང་ ‘to appear, to become manifest’, མོང་ ‘to experience’.

The above list includes both variable auxiliaries such as བྱེད་, བགྱིད་ and invariable auxiliaries such as འདུག་, པདྨེ་, འདུག་, etc.

Here is a list of the most common constructions to indicate TAM in final clauses. In some cases, the relator and the case may be dropped.

**Completed past**

V[past] + AUX.

V[past]-REL + AUX.

V[past]-REL + CASE + AUX.

**Perfect**

V[past]-REL + AUX.

---

18. *RED* ‘to be’ is rarely used in CT. It was spread only after the seventeenth century but remained marginal until the twentieth century. See also Shao (2016).
V[past]-REL (པར/ BAR) + AUX. (འདུག/ DUG/གདུག/ GDA/ཡོད/ YOD/མཆིས/ MCHIS/སྣང/ SNANG)

**Present or uncompleted past**

V[pres.]
V[pres.]-REL+CASE (པར/ PA+R) + AUX. (བྱེད/ BYED/མཛད/ MDZAD/བགྱིས/ BGYIS)
V[pres.]-REL+CASE (པར/ PA+R) + AUX. (འགྱུར/ GYUR)

**Progressive**

V[pres.]-REL (བཞིན/ BZHIN/ཅིང/ CING/གིན/ GIN) + AUX. (འདུག/ DUG/གདུག/ GDA/ཡོད/ YOD/མཆིས/ MCHIS)

**Future**

V[fut.]
V[fut.]-REL+CASE (པར/ PA+R) + AUX. (བྱ/ BYA/མཛད/ MDZAD/བགྱིས/ BGYIS)
V[fut.]-REL+CASE (པར/ PA+R) + AUX. (འགྱུར/ GYUR)
V[fut.]-REL (རྒྱུ/ RGYU/པ/ PA/གི/ GI) + AUX. (ཡིན/ YIN/རེད/ RED)
V[fut.]-REL (ནས/ NAS/ཏེ/ TE) + AUX. (འོང/ ONG)

### 6.6.5. Connectives

The category of connectives, just as auxiliaries, plays an important role in the grammar of CT. One must distinguish between various types of connectives linking nouns (or noun phrase), adjectives (or adjective verbs) and verbs.

We will briefly present here the main verb connectives which also function as clause coordinators or subordinators. Some of these connectives also function as nominal cases (ནས/ NAS, སང/ DANG, རང/ LAS, ས/ S, ཀ/ NA, རི/ GYI, ས/ DU) and as constituents of the auxiliary verbs (see above).


As we have seen earlier, each non-final clause is followed by a connective. In most cases, the non-final verb is not followed by an auxiliary.

The main noun connective is ཆང་ DANG, while adjective connectives are ཆིང་ CING (and its variants བཞིང་ ZHING and བཤིང་ SHING), and ལ་ LA.

6.6.6. Verb and clause nominalization

Among the characteristic features of CT (and OT), one finds the very frequent use of “nominalizers,” which serve as the main tool to nominalize a verb or even an entire clause. Their role is similar to the one of the infinitive in European languages, however nominalizers play a more essential paper in the grammar of CT. Nominalizers are used not only to nominalize verbs or entirely clauses but also to form relative clauses and to indicate, in combination with an auxiliary, the tense-aspect marking.

Various nominalizers are used in CT. They include ས་ PA, མཁན་ MKHAN, ས་ SA, ལྡེ་ RGYU, རྒྱུ་ STANGS, ཆིང་ SROL, སྐྱེད་ BYED, སྤྱ་ BYA, སྟོད་ YUL, ལྡེ་ PHRO. However, the universal nominalizer is ས་ PA/བ་ BA. It is plurifunctional and occurs more frequently than the other nominalizers.

6.7. The relation between Tibetic languages and Classical Tibetan

As mentioned in the introduction, all modern Tibetic languages are closely related to Classical Tibetan and, in some cases, archaic forms found in Old Tibetan. As we will show in the Historical and Comparative Tibetic Lexicon (see the HCTL, Chapter 12), more than 95% of the core vocabulary of modern Tibetic languages is related to a CT, or in some cases to an OT form (see Chapter 4). Some marginal languages at the periphery of the Tibetic linguistic area may have a lower rate of their vocabulary related to Classical Tibetan, due to borrowing from neighboring languages.

However when we compare the modern languages, we can see that the proportion of common vocabulary is much lower (see the HCTL in Chapter 12).
According to Qu Aitang (1996), the modern languages (or “modern dialects”) have only about 60% common vocabulary.19

As we will see (in Chapters 7 and 8), the phonology and grammar of modern Tibetic languages have a clear correspondence with Classical Tibetan. Not only are the modern lexical forms comparable to their literary counterparts, but, in addition, the modern languages exhibit regular sound correspondences with CT as well as fundamental grammatical characteristics. This systematic similarity suggests that all the modern Tibetic languages have been derived from a common ancestor, which was very close to the literary language.

However, depending on languages, we can also find features which are not related to Classical Tibetan. Thus, we consider the development of modern Tibetic languages not as a simple evolution of a single language, but as a result of complicated language contacts.

The split between the various Tibetic languages must have occurred after the development of Tibetan script, probably at the time of the Tibetan Empire (i.e. seventh to ninth century). As shown in Chapter 4, the dialectal diversification could only occur after various phonological processes, characteristic of the Proto-Tibetic period, had already taken place. These phonological features are not attested or, not systematically attested, in the neighboring Bodish languages.

6.7.1. Impact of the literary language on modern Tibetic languages

The relation between CT and the modern languages is not limited to the genetic affiliation. In most cases, just as with other literary languages of the world, CT has slowed down the natural evolution of the spoken languages that were using CT as their written form.

Many monks and lamas devote much of their activity to reading religious and philosophical texts in CT. This reading activity has had the effect of bridging the differences between oral and written forms. Some monks or lamas are able to recite

19. The situation is comparable in many ways to the relationship between modern Romance languages. Similarly, when we compare the modern languages with Latin, the rate is much higher.
entirely by heart some sutras or tantras. Within their colloquial speech and during ritual debates they often quote sentences in CT.

There are cases when a given dialect has changed the natural pronunciation of a word because of the reading style. For example, before 1959, the word for འབྲས་’BRAS ‘rice’ in Lhasa was commonly /’pä:/ following the regular sound change in this dialect (loss of the r after the labial, see Tournadre & Sangda Dorje 2003: 399), but as a result of the education policy in the Literary language, Lhasa people now pronounce this word according to the literary reading /’tä:/.

If we look at the situation of European languages, it is clear that the impact of literary languages on vernacular speech was sometimes significant. For example, in the case of French this impact is quite extraordinary. Many Old French words that had already undergone a significant evolution (such as the loss of a syllable) were ‘Latinised’ back into Middle French. For example, the word medicin ‘medicine’, leume (vs. légume) ‘vegetable’, rade (vs. rapide) ‘quick’, bénecison ‘benediction’ were under the ‘artificial’ influence of the Literary language written and pronounced subsequently as medecine, légume, rapide and benediction.

6.7.2. Impact of modern Tibetic languages on Literary Tibetan

In addition to the strong hypothesis of a common ancestor, the tight relationship between modern languages or dialects and the written language might be partially explained by the integration into the literary language of numerous dialectal words and expressions over the last 1,000 years.

In fact, one of the striking features of literary Tibetan is the existence of numerous quasi synonyms. For example the literary words JIGS, SKRAG, ZHED, and BRED (see the dictionary, part 3) correspond to the same meaning, ‘to fear’. It is thus quite possible that these words belong originally to various dialects and have been integrated into CT.

One should note that the orthography of the nominal and adjectival suffixes -BA and -BO instead of -PA and -PO is emblematic of Amdo and differs from the texts written in Kham and Central Tibet: ‘house’ གཤེགས་ KHANG,BA (Am) vs. གཤེགས་ KHANG,PA; ‘leg, foot’ ལྕགས་ RKANG,BA (Am) vs. ལྕགས་ RKANG,PA; ‘empty’ སྟོང་ སྟོང་, BA (Am) vs. སྟོང་, PA; ‘first’ དང་ཤེས་ DANG,BO (Am) vs. དང་, PO.
Generally, it seems that Tibetans have been very tolerant towards dialectal forms as opposed to what happened in France, for example. The Academy ‘purified’ the French language from all dialectal influences and strived to eliminate dialectal words and expressions. Of course, this intolerance concerning dialectal words is not confined only to France, but, rather, is frequent in the literary languages of the world. However, the degree of ‘dialectal tolerance’ is certainly higher in German and Italian.

A number of Classical Tibetan texts clearly show dialectal influences. Some writers inserted into their texts lexical items or even grammatical words respectively from Kham, Amdo or Tsang. It seems that, unlike some European traditions, Tibetans and other Bhoti groups never tried to eliminate dialectal expressions and words from their literary works. One possible explanation for this is linked with the oral traditions of Tibetan Buddhism and Bön. Tibetan culture has been transmitted both through the written tradition of *pecha* (བདེ་ཆ་ DPE.CH'A) texts and by means of oral traditions. Both are considered equally important. A written text, especially if it is a root text, *tsawa* (རྩ་བ་ RTSA.BA), receives oral commentaries and explanations, which often serve as practical instructions. The *tsawa* is essential for the theoretical approach, whereas the oral instructions (གདམས་ངག་ GDAMS.NG.K) are fundamental for their implementation.

The following texts are considered to be influenced by dialectal features (Dung-dkar Blo-bzang ’Phrin-las 1997: 316; Sumlba Don grub}, Tshe Ring 2011):

**Traces of Phünpo dialect**

- *བེའུ་བུམ་སྔོན་པོ་ BE’U BUM SNGON.PO* and *དཔེ་ཆེས་རིན་ཆེན་སྤུངས་པ་ DPE-CHOS RIN-CHEN SPUNGS.PA* composed during the twelfth century by DGE.BSHES.PO TO.BA.

**Traces of Tsang or Tö dialects**

- The famous Milarāpa’s biography *མི་ལ་རས་པའི་རྣམ་ཐེར་ MILLA RAS.PA’I RNAM.THAR*, composed during the fifteenth century by GTSANG.SMYON.HE RU.KA SANGS.RGYAS RGYAL.MTSHAN.

**Traces of Amdo dialect**

- *ཧོ་བདེ་བའི་ཕལ་སྐད་ཟབ་ཆོས་ GO BDE.BA’I PHAL.SKAD ZAB-CHOS* and *ཟློས་གར་གྱི་བསྟན་ BCOS ZLOS.GAR-GYIBS.TAN.BCOS*, composed in the Labrang Amdo dialect during
the nineteenth century by GUNG.THANG RIN.PO.CHE.DKON.MCHOG.BSTAN.PA'I SGRON ME; and BEL.GTAM, composed in the nineteenth century. • MDO.SMAD CHOS 'BYUNG, edited by BRAG.D贡.PA.DKON.MCHOG.BSTAN.PA.RAB.RGYAS in the nineteenth century.

Traces of Kham dialect
• GTAM.PAD.MA.TSHAL.GYI.ZLOS.GAR, composed in the nineteenth century by RDZA.DPAL.SPRUL.RIN.PO.CHE.

Further research is needed to determine dialectal influence in Tibetan Classical literature. Some of the above texts contain not only dialectal expressions but also some grammatical dialectisms. The contemporaneous literature of Amdo, which is still written in a style close to Classical Tibetan, shows even more dialectal influences in some texts. Among the famous works that manifest such influences are short stories composed by DON.GRUB.RGYAL, the famous Amdo writer and poet who committed suicide in 1980 (see e.g. Robin 2003).

It is also likely that lamas and lay authors in Ladakh, Bhutan and other areas outside Tibet also were influenced by their local language or dialect (as it is the case currently). Thus it is almost certain that over many centuries various Tibetic dialects have poured lexical items and expressions into the literary language.

Both the dialectal influence on literary language and the literary influence on dialects might have occurred during their millennium of “cohabitation.” However, the situation is probably much more complicated than this, and some fundamental research has yet to be done in this field. Several important issues are still to be clarified:

a) The language underwent some significant transformation in its vocabulary (see section 6.5 in this chapter) during the transition from OT to CT. What prompted these transformations?
b) Sanskrit had an impact on the literary language, particularly on the lexicon and grammar of the canonical texts which still needs further research. However, did Zhangzhung, Tangut, Chinese, or any ancient ST languages have any significant
impact on Old Tibetan and Classical Tibetan?

c) What influence did Bodic, Kiranti, Kinnauri, Qiangic or rGyalrongic languages and other Tibetospheric languages have on the Tibetic languages?

d) What influence did the various Tibetic languages have on the evolution of Classical Tibetan? To what extent did Classical Tibetan integrate dialectal words?

All these questions require further research, ideally involving interdisciplinary collaboration e.g. in the fields of linguistics, history and anthropology.

6.7.3. The written language and the reconstruction of protoforms

Few ‘compact language families’ in the world have long written traditions, which are helpful for the reconstruction of the proto-languages. In most cases, there aren’t any written traditions (more than 95% of world languages do not have a written tradition, or, if one exists, it has been in limited use for less than 100 years). Written records are, for example, not available in the case of Australian, Oceanic, most Amerindian (with a few exceptions such as Guarani, Quechua, Nahuatl, etc.) and most African languages (except for the Semitic languages and a few other exceptions). The only area with abundant and ancient written records is Eurasia. Even in this region of the world, some language families, such the Uralic family, lack records that would help to reconstruct the whole family. Modern languages or language families with written documents dating back more than a thousand years belong to a small “club,” which totals less than twenty members: Greek, Romance, Germanic, Armenian, Celtic, Persian, Indo-Aryan, Slavic (which are all Indo-European branches), Semitic, Dravidian, Austronesian, Japanese, Mon-Khmer, Kartvelian, Turkic, Sinitic and Tibetic.

20. By “compact family,” we mean here a language family whose languages are closely related and have a Proto-language that is related to an attested written language or can be easily reconstructed.
7. A phonological outline of the modern Tibetic languages

The diversity of Tibetic languages manifests itself in various linguistic fields, such as phonetics and phonology, morphology and lexicon; on the other hand, as mentioned earlier, they share many common characteristics. This section deals with the phonological aspect. The section is divided into two parts: one is a pandialectal overview from the synchronic, macroscopic viewpoint with a unified list of the phonemic components to be able to describe all the varieties of Tibetic languages (section 7.1 and 7.2) as well as a brief description on the suprasegmentals generally called “tones” (section 7.3); the other is a historical or diachronic overview of the sound development compared with CT forms (section 7.4).

7.1. Pandialectal phonetic description and its transcription

The two charts below present the symbols used in order to transcribe the consonants and vowels found in most of the Tibetic languages or dialects and include some rare sounds attested only in a few dialects.

The sounds that are common to most Tibetic languages are given in bold and a larger font size than the sounds which are more specific. The transcription of frequent sounds (in bold) is rather straightforward for English speakers and meant for non-linguists. It has been chosen for simplicity’s sake and in many cases corresponds to the English pronunciation.¹

For example the sounds /ts, dz, j, zh, ng, ny/ are similar to the pronunciation of the spellings ts, dz, j, zh, ng, ny respectively in the English words lots, adz, jazz, show, Brezhnev, king, canyon.

Linguistic phonetic transcriptions make use of the International Phonetic Alphabet (IPA) authorised by the International Phonetic Association. When our phonetic transcription differs from IPA, the symbols of IPA are given in a square bracket. Some phonetic symbols are not registered in IPA but are mainly employed in Chinese linguistic

---

¹ This is meant for the convenience of a general readership. Linguists can/should use a phonetic description (IPA and other symbols) when citing the data from our book if necessary. For criticism against forcing non-English words to follow the English convention, see Hill (2012b).
works. They are well defined in Zhu (2010); for the sake of providing a precise description, we adapt them in the following charts. Otherwise, the transcription letter is equivalent to the corresponding IPA symbol.

### Chart VII.1. – Pandialectal transcription of the consonants

<table>
<thead>
<tr>
<th>Plosive</th>
<th>Labial/ labio- dental</th>
<th>Denti- denteal</th>
<th>Retro- flex</th>
<th>Pre- palatal</th>
<th>Pre- velar</th>
<th>Velar</th>
<th>Uvular</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>sibil.</td>
<td>p</td>
<td>t</td>
<td>ʈ</td>
<td>s</td>
<td>c</td>
<td>ky [kʰ]</td>
<td>k</td>
<td>q</td>
</tr>
<tr>
<td>voiced</td>
<td>b</td>
<td>d</td>
<td>d</td>
<td>j</td>
<td>gy [gʰ]</td>
<td>g</td>
<td>α</td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td>ts</td>
<td>ʈs</td>
<td>ts’ [ʈʰs]</td>
<td>ts’ [ʈʰsʰ]</td>
<td>x’ [xʰ]</td>
<td>x</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>dz</td>
<td>dʒ</td>
<td>j [dz]</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>s</td>
<td>ʃ</td>
<td>ʒ</td>
<td>x</td>
<td>x’ [xʰ]</td>
<td>x</td>
<td>z’ [zʰ]</td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>s’ [sʰ]</td>
<td>ʃ’ [ʃʰ]</td>
<td>ʒ’ [ʒʰ]</td>
<td>x’ [xʰ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td>ʈ’ [ʈʰ]</td>
<td>ʈ’ [ʈʰ]</td>
<td>j</td>
<td>y’ [yʰ]</td>
<td>γ</td>
<td>υ</td>
<td>θ</td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>ʈ’ [ʈʰ]</td>
<td>ʈ’ [ʈʰ]</td>
<td>j</td>
<td>y’ [yʰ]</td>
<td>γ</td>
<td>υ</td>
<td>θ</td>
<td></td>
</tr>
<tr>
<td>Vibrant</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r’ [rʰ]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>r’ [rʰ]</td>
<td>r’ [rʰ]</td>
<td>r</td>
<td>r’ [rʰ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m’ [m̩]</td>
<td>n’ [n̩]</td>
<td>ny’ [ŋ̩]</td>
<td>n’ [ŋ̩]</td>
<td>ng’ [ŋ̩]</td>
<td>n’ [ŋ̩]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>m</td>
<td>n</td>
<td>ny’ [ŋ̩]</td>
<td>n’ [ŋ̩]</td>
<td>ng’ [ŋ̩]</td>
<td>ng’ [ŋ̩]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semi- vowel</td>
<td>w, v</td>
<td>y [j]</td>
<td>y [j]</td>
<td>y [j]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Some sounds have not been included in the chart VII.1. either because the accuracy of their phonetic description is debatable in the Tibetic languages, e.g. [ʃ], [ʒ], [ʃʰ], [ʒʰ], [dʒ]. We did not include either some very rare sounds, such as dental fricatives [s̪], [z̪], dental-postalveolar affricates [ʈs̪], [dz̪], pharyngeal fricatives [ɧ], [ʕ], and epiglottal fricative [ʢ].

2. The postalveolar articulation rarely has a phonemic status in the languages spoken in East Asia, and the prepalatal counterpart functions instead. Cf. Zhu (2010).

3. These sounds are well attested in Arabic languages.

4. This sound is well attested in Hebrew.
that [w] is a double-articulated sound of the bilabial and velar positions, [ի] is a double- or triple articulated sound of the (bilabial), palatal and velar positions. The phonetic symbol [ɾ] is defined as an “alveolar trill” in IPA, but for the simplicity’s sake, it designates here multiple kinds of “r-like sounds,” which are usually not distinguished by speakers of the Tibetic languages, including alveolar trill ([ɾ] in IPA), alveolar flap ([ɾ] in IPA), alveolar approximant [ɹ], retroflex flap [ɽ], and even voiced retroflex fricative [ʐ] (the last one appears only in the case of no phonemic contrast between /r/ and /ʐ/).

CHART VII.2. – Pandialectal transcription of the vocalic sounds

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>unrounded</td>
<td>rounded</td>
<td>unrounded</td>
</tr>
<tr>
<td>High</td>
<td>i</td>
<td>i</td>
<td>ü [u]</td>
</tr>
<tr>
<td></td>
<td>і</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid high</td>
<td>e</td>
<td>ə</td>
<td>ɵ [o]</td>
</tr>
<tr>
<td>Mid low</td>
<td>ɛ</td>
<td>ɜ</td>
<td>ʌ</td>
</tr>
<tr>
<td>Low</td>
<td>a</td>
<td>A</td>
<td></td>
</tr>
</tbody>
</table>

Oral consonantic vowels: 5 ɿ (ɿ), ɿ ʰ, ɿ ₇

Oral consonantic vowels are sounds with both the features of a semi-vowel (see the chart of consonants) and a vowel, functioning as a syllable core like a full vowel. They are attested only in a few Tibetic languages (see below 7.2.1.2.). From the typological point of view, one should note that these sounds are not so frequent in the World’s language.

Some of the vowels presented in the chart may be nasalised and have a phonemic length distinction. Nasalised vowels are noted with a tilde such as [ã, ǐ, ũ], etc. and long vowels are noted as [aː iː oː uː], etc.

More rarely, one can also find retroflex (˷), velarised (ˠ or ᶛ) or pharyngealised (ˤ) vowels.

5. The description of [ի] is slightly different from the IPA’s, but here we use the identical symbol. See Suzuki et al. (2019).

6. ɿ is realized as a syllabic labiodental approximant; ʰ is articulated as in American English, in “it” of bird and in Beijing Mandarin er ‘two’; ɿ resembles a frictionlike or humming [ɾ].
In the above chart, the vowels in bold are the basic vowels found in almost all the Tibetic languages; those not in bold are found in only some dialects.

**Central rounded vowels**

The existence of central rounded vowels (ü ö) plays an important role in most of the Central and Southern Tibetic languages (Ü, Tsang, Kham, Hor, Dzongkha, etc.). In addition, in Amdo, Eastern Tibetic languages, Kham and Hor, one generally finds a central vowel (ə) with phonemic status.

### 7.2. The main characteristics of pandialectal segmental phonology

We summarise below the phonological features on consonants and vowels of the Tibetic languages.

#### 7.2.1. Consonants

We will first deal with the modes of articulation and then present the places of articulation.

**7.2.1.1. Modes of articulation**

- **Aspirated sounds**

An important characteristic found in all the dialects is the phonological function of aspiration. It occurs only with voiceless sounds, unlike Hindi and some other Indic (Indo-Aryan) and Dravidian languages. Thus in all the languages we find a distinctive opposition between /p, t, tʃ, k, ts, č/ and their aspirated counterparts: /p', t', tʃ', k', ts', č'/.

Additionally, in many languages of Eastern Tibet (e.g. Kham and Amdo), the opposition extends to other phonemes such as /s/ vs. /s'/, /ʃ/ vs. /ʃ'/, and /ʃ/ vs. /ʃ'/.

---

7. Most of descriptions on the “central rounded vowels” here regard them as “front rounded vowels” i.e. [y] and [œ]. However, the sounds attested in almost all the dialects, the notation with front rounded vowels is phonetically inappropriate and we use here u and o with umlaut instead.

8. Aspiration is also an important characteristic in Hindi, Chinese, and many other languages of the area.
Voicing

Voiced plosive and affricate sounds such as /b, d, ɖ, j, g, dz/ are found in the entire area, though in Ü and Tsang of Central Tibet, they do not have a phonemic status.

In most languages one also finds the following fricative voiced sound as opposed to voiceless: /s/ vs. /z/, /sh/ vs. /zh/. This opposition is generally absent in Ü and Tsang, where one has only voiceless fricatives.

The opposition between voiceless velar fricative /x/ and voiced /ɣ/ is well attested in Eastern Tibet as well as in Zangskar and Balti. In some rare dialects of Eastern Tibet, such as Cone and Gyalthang, one also encounters a voicing opposition for retroflex sounds /ʂ/ vs. /ʐ/.

In some dialects of Central Tibet (e.g. Tsang) and southern Kham (Gyalthang), the vibrant voiced sound /ɾ/ is opposed to a voiceless /ɾ'/.

The opposition between denti-alveolar voiced lateral /l/ and voiceless lateral /l'/ is also ubiquitous with a few exceptions, notably Central Ladaks, Khöpokhok (Zitsadegu), Thewo-mä and Drugchu, which lack the latter.

In many languages spoken in Eastern Tibet and in some Tibetic languages of the southern Himalayas such as Lhoke, the opposition of voicing extends to the nasal sounds: /m/ vs. /m'/, /n/ vs. /n'/, /ny/ vs. /ny'/, /ŋ/ vs. /ŋ'/.

Ex.: མྱེ་། SNA 'nose': /ʰna/ vs. /na/ 'sick'; མྱེ་དི་/SNYING/ 'old'.

Nasal series

The existence of nasal /m/, /n/, /ŋ/, /ny/ is ubiquitous in all the Tibetic languages (only some dialects of Cone, Markham (in Tibet) and Dazundam (Myanmar) lack /ny, ny/). The phonetic realization of /ny/ has been disputed among scholars. This sound is realized as a prepalatal in almost all varieties of the Tibetic languages, but since the symbol /n/ is not an IPA convention, most scholars use the sign /ɲ/, which refers

---

9. In Tsang the voiceless /ɾ'/ is often realised as a retroflex fricative consonant [ʂ].
to a palatal position. This distinction is in fact important, because a few languages of Kham distinguish a palatal (/ɲ/), and a prepalatal nasal (/ȵ/; see Suzuki 2016).

As seen above, in many dialects of Kham, a series of voiceless nasal plosives is also found: /m/, /n/, /ny/, /ng/.

7.2.1.2. Place of articulation

The following commentaries will deal with specific features of the Tibetic languages and do not present an exhaustive list of the phonemes.

- **Labial obstruent series**

  The labial plosives /p/, /p/, /b/ are pervasive in all Tibetic languages, except for some dialects of Central Tibet where the completely voiced sound [b] is generally absent.

  Labial or labio-dental fricative phonemes /f/ and /v/, often realized as labial fricatives, respectively [ɸ] and [β], are found in Zanhar and in a few dialects spoken in Amdo. In Amdo, depending on dialects, /f/ is pronounced as [f] or as [ɸ] and they are not free variants.

- **Denti-alveolar obstruent and lateral series**

  The denti-alveolar plosives /t/, /t/, /d/ and africates /ts/ /ts/ and /dz/ are pervasive in all Tibetic languages.

  The denti-alveolar fricatives /s/, /z/ are found in most languages, although some dialects of Central Tibet (such as Ü and Tsang) lack a voiced counterpart /z/. A few dialects do have additionally interdental fricatives /θ, ð/ in Zanhar (see Hoshi & Tondup Tsering 1978), and some eastern dialects have /ɭ/ and /ɮ/ in mBalhag, Pashi and Čone (see Suzuki 2009, 2013).

- **Retroflex obstruent series**

  One of the characteristics of modern Tibetic languages is the existence of retroflex sounds: /ʈ/, /ʈ/ and /ɖ/ (plosive), /ʂ/ (fricative), /ɾ/ (vibrant or flap). Some very rare

11. The retroflex articulation may be considered as an areal feature, since it is also found in Hindi and Chinese. However, the phonetic articulations are different. In many Tibetic languages, the
dialects have a distinction between /ʈ/ and /ʈʂ/ (e.g. Gyälthang, Čone). Other rare varieties entirely lack retroflex sounds (e.g. Thewo-ма).

- **Prepalatal obstruent series**

  The affricate phonemes /ɕ/, /ɕ'/ and /j/ are found in almost all Tibetic languages. The fricative /ʃ/ and /ʒ/ are also found in most languages (although /zh/ is not found in Ü and Tsang). The plosives /t, t', d/ are attested in some very limited dialects of Kham.

- **Palatal and prevelar obstruent series**

  The palatal and prevelar obstruent series are limited to a few Tibetic languages. In Ü and Tsang, one finds a series of prevelars /kʰ/ and /gʰ/, while in some dialects of Amdo and Kham, one encounters palatal plosive or affricate realizations /c/, /ɟ/ /cç/, etc. In some dialects of Kham one can encounter palatal fricatives /ç, ʝ/.

- **Velar obstruent series**

  The existence of velar plosives /k/, /k'/, /g/ is pervasive in Tibetic languages. However, the voiced sound /g/ does not have a phonemic (or distinctive) value in Ü and Tsang.

  The existence of velar fricatives /ʃ/ and /ɬ/ is attested in both Eastern and Western Tibetic languages (Amdo, Kham, Sharkhok, Ladaks, Zanhar and Balti), but is not found in the languages of Central Tibet nor in the Tibetic languages of the southern Himalayas.¹²

- **Uvular obstruent series**

  Several dialects of Amdo, Kham, Hor and Eastern Tibetic languages, such as Pälkyi [Pashi] and Purik, have some uvular articulations, especially the voiceless uvular plosive /q/. The fricative /χ/ appears more frequently as a final. The dialects of Amdo, Kham and retroflex obstruents are made between a tongue tip and postalveolar. Chinese scholars often regard our “plosives” as affricates; however, contrary to denti-alveolar and prepalatal affricates, one cannot pronounce lengthened retroflex sounds. This means that retroflex sounds are primarily plosives, which are often with a weak friction.

¹² Concerning the phonetic value of the sign /ʃ/ and /ɬ/ and the opposition in Sharkhok, see Suzuki (2008).
Pälkyi [Pashi] have the uvular series as an initial. They also occur in Balti as final. Uvular sounds are not distinctive in the written system of Classical Tibetan.

- **Glottal obstruent series**

  The glottal stop /ʔ/ is found in many Tibetic languages. Some dialects of Amdo lack a glottal stop. The fricative voiceless /h/ and voiced /ɦ/ are also ubiquitous. Note that the /ɦ/ is sometimes described as breathy voice (see Suzuki 2013d & 2015a).

  Additionally, in many Central and Eastern languages as well as Zanhar, the glottal stop is frequently found in the final position (see section 9).

### 7.2.2. Vowels

- **Oral vowels**

  The oral vowels /a, i, u, e, o/ are common to all the dialects. They correspond to the diacritic vowels found in Classical Tibetan.

- **Nasalized vowels**

  Phonetically, one can hear nasalized vowels such as [ä, ů, ū], etc. in most of the languages and dialects. However, some phonemic distinctions are found in particular in Kham, Tö, Hor Nagehu, Sharkhok and the Nubra dialect of Ladakh. Scholars hold different views over the phonemic status of nasal vowels in Lhasa. Vowels do not undergo nasalisation in Amdo and Ladaks.

  For the languages that have only a phonetic realization, we will note the nasalization with an N as /aN, āN, eN, ēN, oN, ūN/ but when the nasalization has a phonemic status, for example in Kham and Čone, we will note it with a tilde /ä, ē, ī, ŏ, ū/ as in IPA.

  Ex. སོན་ ’medicine’, /sman/ (Ba, La), /‘man/ (Am), /‘men/ (Sh), /ˈmāN/ (Ū, Ts), /m’ē/ (Kh).

---

• Central Rounded vowels

Ü, Dzongkha, Kham, Hor and Sharkhok, etc. have developed the phonemic rounded vowels /ü, ü/. These vowels are generally not found in many varieties of Amdo, Ladaks, Purik or Balti dialects.

Ex. སྣྲུལ་ sbrul 'snake' /’ʈʊː/ (Kh), /’bʊ/ (Ho, Dz); བོད་ bod ‘Tibet’ /’p’ö:/ (Kh), /’b’ö/ (Dz).

• Oral consonantic vowels

This series of vowels such as /ɿ, ɾ, ʋ, ʊ/ is rarely found in the Tibetic languages, although /ɿ/ (ʅ) appear more frequently in some dialects of Kham and Amdo. Phonetically, [ɿ] and [ildo show a complementary distribution so that they can be analyzed as one phoneme. (See Suzuki 2014b.)

/v, ʋ/ are only found in a few dialects of Kham (Gyālthang). /ʊ, ʊ/ are found in Drugchu and some dialects of Amdo (see Suzuki 2013c and Tsering Samdrup & Suzuki 2019).

• Retroflex, velarized or pharyngealized vowels

The vowels with a secondary oral articulation do not frequently appear. However, some varieties from South-eastern Tibet have several vowels of this kind. Not all the vowels can take these secondary articulations. The retroflex vowels /əː/ is frequently attested.

7.2.3. Phonotactics

Phonotactics, i.e. syllable structure, is essential to understanding the phonetic and phonological features of Tibetic languages (see also the syllabic structure of CT in 5.2). From the synchronic viewpoint, a syllable can be divided into two parts: initial and rhyme. The former is either a single consonant, a cluster, or even zero; the latter can be sub-divided into syllable core and final. Here we will briefly explain the components of the syllable structure: initial consonant cluster, syllable core and final.14

14. A general phonotactics includes a description of suprasegmentals, but in Tibetic languages the suprasegmentals can be born by multiple syllables (see next subsection), they are thus treated separately.
• Initial consonant clusters

If we take into consideration phonotactics, several Tibetic languages such as Balti, Ladaks, Purik, Zanhar and Amdo have a significant number of consonant clusters. The clusters consist generally of a preinitial consonant and an initial consonant.

• Preinitials

The first position of the initial consonant cluster in a syllable is called a 'preinitial'. The preinitial elements may include obstruents, nasals and liquids of the labial (/p/, /b/, /w/), denti-alveolar (/s/, /z/, /l/, /n/), retroflex (/r/, /ʂ/), velar-uvular series (x,ɣ,χ,q,ʁ) and glottals (h,ɦ).

From a historical point of view, preinitials correspond in CT to the preradical letters G, D, B, M, R, S, L or to the ante-preradical letter B. In a more marginal way, they may also correspond in some languages to a radical letter PH, B followed by a glide R, Y. We propose to note the preinitial sounds with a small exponent letter. For example, depending on the dialect, the word རྟ་ RTA is pronounced as /ʁta/, /s따/, /ʰ따/, etc. or not pronounced at all /따/, and the superscripts ʰ, ɾ in front of the /따/ are the reflexes of the preradical R. Note that in some dictionaries, the preinitials are noted by a parenthesis. For example (s)ta 'horse'.

Preinitials play a very significant role in the phonology of Tibetic languages and have not received sufficient attention. They bear two fundamental characteristic features. First, the preinitial sounds may undergo many variations, even in the case of closely related languages or dialects. Sometimes, the variation may also depend to a certain extent on the speakers and various sociolinguistic parameters (age, formal speech, individual variation, etc.) may intervene in the realization of the preinitials. They may be pronounced in a different place or manner of articulation or in some cases entirely disappear:

For example, གདོང་ GDONG 'face' is pronounced /ʁdong/ (Ba: Skardo) vs. /sdong/ (Ba: Khapulu), /sdongpa/ (口中); ཨི碉碉 DPE,RA 'language, speech' (derived from CT ཨི碉碉 DPE,SGRA 'speech') /spera/ (Ba), /spera/ (La), /spera/ (Nyoma); བོར་ SKYAG,PA 'feces, excrement' /skaŋka/ (La) vs. /skakpa/(Ba), /skakpa/(口中); རྟ་ RTA 'horse'/ta/ (La, Nubra), /ta/ (Ba, Sham, Am: dr), /ʰta/ (Am: ro), /ʰta/ (口中); ང་ LTA 'to look at'/ta/ (La,
PART 2 – CH 7. A Phonological outline of the modern Tibetic languages

Ba, Pur), /ˈta/ (Am: dr), /ˈta/ (Am: ro), /ˈta/ (Ü, Ts, Kh, Dz); ཡོ་མཚོ། MTSHO 'lake' /ˈtsʰo/ (Am: dr), /ˈtsʰo/ (Am: ro, Kh, Ho), /ˈtsʰo/ (La, Ba, Pur), /ˈtsʰo/ (Ü, Ts, Dz); བརྗེད་ BRJED 'to forget' /ˈzhet/ (Ba) vs. /ˈzhet/ (Ba), /ˈjet/ (Am: rNgawa) vs. /ˈja/ (Am: Chabcha), /ˈje:/ (Ü); བརྒྱད་ BRGYAD 'eight' /ˈgyat/ (Ba) vs. /ˈgyat/ (Purik), /ˈjal/ (Am: Themchen) vs. /ˈjat/ (Am: rNgawa), /ˈgjä:/ (Ü), /ˈgä/ (Dz).

Note that when the preinitial disappears, the initial, in most cases, goes unchanged, but modifications of the initial may also occur. For example, ཕན། SGO 'door' /ˈgo/ (Sham), /ˈgo/ (Am), /ˈgo/ (Am), whereas in many languages (Ü, Ts, Kh, Dz, Sh) the preinitial is no longer present and the initial is left unchanged: /ˈgo/. In some dialects such as Zanhar and Leh, the preinitial is modified and the word becomes /ˈgo/.

Variations in preinitials across languages and dialects are sometimes due to the fact that CT has a combination of preradicals, such as BR in the case of བརྒྱད་ BRGYAD or བརྗེད་ BRJED, and that a given language or even dialect may choose the first preradical (B) or the second (R).

However, in most cases, the preinitial sounds are the reflexes of the same preradical sounds in CT (or OT). See above e.g. the words 'horse', 'face', 'lake', etc. The fluctuation of preinitial sounds across the languages is thus due to an inner evolution of the preinitial: for example, in the case of རྟ་ RTA, it was first an /r/, as attested in some Amdo pastoralist dialects and in Shamskat (Western Ladakh), but then underwent a transformation to /s/ or a lenification yielding a glottal fricative /h/ or even a shwa and finally the entire disappearance of the preinitial.

Another important point to be noted is that, in most cases, the reflexes of the preradical are regular, but in some words the reflexes do not follow the expected rule. For example, this is evident in the following Ladaks and Purik words: /ˈchin/ 'urine', /ˈdemo/ or /ˈdema/ 'beautiful' and /ˈdutpa/ or /ˈdutpa/ 'knot', which respectively correspond to the CT words: གཅིན་ GCIN, བདེ་མོ་ BDE.MO, མདུད་པ་ MDUD.PA. Whereas in Balti, the reflex of the preradical G yields a velar preinitial, in Ladakh, the velar has been deleted or replaced by a /l/ preinitial (particularly in front of an affricate /č/). This is probably due to an analogy and a convergence with many words which do have an initial cluster in /č/ such as /ˈcangma/ 'tree', /ˈcaks/ 'iron' and /ˈce/ 'tongue' derived


from CT བཀྲ་ཤིས LＣＡＧＳ and བྱེ་ＬＣＥ. Concerning the word ‘knot’, /dutpa/ or /r∅utpa/ (< མདུན་པ་ MDＵＮ.ＰＡ). Western languages (unlike eastern Tibetic) have not preserved prenasals and thus in most cases the words are left without a preradical: /dun/ ‘in front’ or /da/ ‘arrow’ derived from CT མདུན་ MＤＵＮ and མཛད་ MＤＡ. The presence of a preinitial in the word མདུན་པ་ MDＵＮ.ＰＡ /dutpa/ or /r∅utpa/ could not have been inherited and can only be explained by analogy with words which have similar clusters such as /doa/ ‘stone’ and /dak/ ‘to lick’ from CT རྡོ་བ་ RＤＯ.ＢＡ and ལྡག་ LＤＡＧ.

Second, the preinitials can be pronounced in the same way as the radical, but they are often pronounced with a weak sonority, essentially realized as a reduced volume. The sonority depends on the given language or even dialect. There are three types of phonetic variations of preinitial in the various languages.

1. Preinitials are always pronounced in the same way as the radical.
2. Preinitials are always pronounced with a weak sonority.
3. Preinitials are pronounced either with a strong or weak sonority.

In some dialects demonstrating the third type, the strong versus weak opposition may even have a phonological status.

The weak realization corresponds to a “secondary articulatory mode” well attested in Tibetic languages, which plays a significant role in phonological systems. This specific mode is frequent in Amdo, Kham and most eastern Tibetic languages but, from a typological point, it is extremely rare in world languages. A few Amdo pastoralist dialects have the first realization whereas the majority of Amdo pastoralist dialects use a secondary articulatory manner.

It is important to note that dialects that have lost preinitials at the beginning of a word may still have traces of the preinitials inside a word. For example, in Ü and Ladaks or Purik dialects, རྡུག DＵＧ /‘tuʔ/ ‘there is’ (Ü) vs རྡུག་ MI-DＵＧ /‘miʔtuʔ/ ‘there

15. When it is necessary to distinguish between a strong and a weak realization in a single dialect, we propose to note the strong articulation by underlining the preinitial. Thus, for example, we can distinguish between /som/ /‘sam/.
A Phonological outline of the modern Tibetic languages

isn’t’ (Ū); མོ/go/ (Pur), ‘head’ vs. མོ/go/ ‘beginning of the year’ (lit. ‘head of the year’) (Pur). See also Zeisler (2009).

Finally, an orthographic remark should be made. Since people are not always aware of the preinitial, it has some indirect impact on the orthography of words. It is particularly obvious with the nasal preinitial which is a reflex of མོ. For example, it is worth noting that some Tibetans, who don’t know the Classical orthography of the negative copula མི་འདུག་ often spell this negative copula: མིན་འདུག་. This spelling is used to render the nasalization /n/ that is noted by a small letter in exponent /mi-n/), pronounced [mintu:]. But many people are not aware that the nasalization comes in fact from the letter མོ which triggers nasalizations in many words: རྒྱུ་འབྲས་/‘karmic causality’, དགེ་འདུན་/‘Saṅgha’, དུས་འཁོར་/‘wheel of time’ (Kālacakra), དཔལ་འབྱོར་/‘economy’, ཤ་འབྲས་/‘meat and rice stew’. These words are all pronounced with a nasalization that comes from the letter མོ. The use of མིན་འདུག་ instead of མི་འདུག་ is problematic because the former spelling also exists in CT but with a different meaning. It corresponds to the negation of མི་འདུག་ ‘the sensory inferential’ corresponding in Common Tibetan to མི་འདུག་ ‘the sensory inferential’.

Nasal preinitials or “prenasalization”

The nasal preinitials /n/ and /m/ are usually pronounced as sounds that are weaker than that of the main consonant. Many clusters may include prenasalized sounds: /kn/, /ng/, /nd/, /nt/, /nt/, /dz/, /c/, /n/, /p/, /m/, etc. This type of secondary articulatory mode, which is called “prenasalization,” is very frequent. Prenasalized sounds are mostly found in Eastern Tibet, as well as Hor and Tö. They are generally absent in Western, Central and Southern Tibetic languages (Ladakhs, Balti, Ü, 16 Tsang, Sherpa, Dzongkha, Lhoke, Choča-ngača, etc.).

16 Some Japanese scholars, such as Kitamura (1977), Yukawa (1971), Hoshi (2003), and Kitamura & Nagano (1990), have described prenasalisations in a Lhasa variety. Kitamura and Yukawa have done pioneering work in describing this variety based on the speech of a woman from the Tsarong family who grew up in the Horkhang family since childhood.
Ø Glottal preinitials or "preaspiration"

The glottal preinitials /h/ and /ɦ/ are usually pronounced as weak sounds. This type of secondary articulatory mode, "preaspiration," is also frequent, although less than prenasalization. Let’s note that preaspiration is uncommon in World languages but is also present in some Scandinavian languages, such as Icelandic and Faroese, as well as Celtic languages, such as Scottish Gaelic.

In some Tibetic languages, preaspirated sounds are found before nearly all types of consonants. The combination of a different voicing is also allowed so that the preaspiration is written /h/ and /ɦ/ based on its voicing reality. There are obstruents with a preaspiration such as /ʰp, ʰt, ʰk, ʰts, ʰɛ, ʰʃ, ʰs, ʰb, ʰd, ʰɡ, ʰđ, ʰdz, ʰʃ, ʰz, ʰpb, ʰtt, ʰkt, ʰbʒ, ʰʃb, ʰm, ʰn̥̂, ʰŋa, ʰnya, ʰma, ʰnga, ʰnya, ʰma/, etc., and resonants such as /ʰn, ʰma, ʰŋa, ʰnya, ʰna, ʰma/, etc.

Ø Velar or uvular preinitials

Velar sounds, sometimes realized as uvular ones, are found especially before voiceless unaspirated and voiced consonants in Amdo, Balti and Purik. They may be realized as strong or weak ("secondary articulatory mode") depending on the languages and dialects.

In some dialects, such as Labrang and Rebgong, velar preinitials can freely alternate with preaspirated sounds, but in other dialects, such as rNgawa and Themchen, they are distinctive.

Ø Labial preinitials

In Dzongkha, Balti, many pastoralist dialects of Amdo, and some Kham dialects, one finds the following labial series /ʰp, ʰtʃ, ʰtʃ', ʰtʃ', ʰt/, /ʰp, ʰtʃ, ʰtʃ', ʰtʃ', ʰtʃ/, /ʰp, ʰtʃ, ʰtʃ', ʰtʃ', ʰtʃ', ʰtʃ/, etc.
Denti-alveolar preinitials

The preinitials /s, z, l/ frequently occur in Western Tibetic languages, such as Balti, Purik, Ladakhs and Zanhar. They originally derive from the preradical letters S, Z and L of CT found in various combinations such as SK, SG, ST, SD, LT, LD, SP, SB, LK, LCH, ZL, SL and SR. Denti-alveolar preinitials are usually pronounced as strong sounds and not as “secondary articulatory modes.” Frequent clusters include /sk, zg, st, zd, lt, ld, ls, sp, zb, lz, lts/.

Ex. སྐད་ 'sound, language': /skat/ (Ba, La), སྐྲ་ 'hair': /skra/ (Purik), སྒོ་ 'door': /sgo (Ba), སྟོད་ 'upper part, upper valley': /stot/ (La, Ba), སྒུག་ 'sorrow, grief': /sgug/ (Ba, Purik), སྤྱིི། 'frog': /sbal/ (Ba), སྲིག་ 'iron': /skr/ (La), /srl/ (Ba); སྤྱིི། 'to repeat, say': /srau/ 'conversation' (Za), སྲིག་ 'ZL 'OD 'moon' (litt. 'moon-light'): /srlt/ (Ba), སྲིག་འོད་ 'moon': /srlad/ (Za).

Other rare combinations include the following /lk, lng/. They are attested in rNgawa.

Ex. སྲིག་ 'mute': /lk/ (Am: rNgawa), སྲིག་ 'key': /lde/ (Am: rNgawa), སྲིག་ 'five': /lnga/ (Am: rNgawa).

Retroflex preinitials

The retroflex preinitials /r, ʂ/ are frequent in many Amdo dialects, but also occur in Western Tibetic languages such as Ladakhi and Balti.

Ex. སྐད་ 'language': /rket/ (Am: rNgawa), ʂ 'RTA 'horse': /rta/ (Am: rNgawa), ʂ 'RDO(BA)' 'stone': /rdoa/ (Am), /rdoa/ (Ba, Purik).

Initial sounds

From a historical diachronic point of view, initial sounds generally correspond in CT to the radical letter. They differ from the preinitial in the sense that they are more stable across the various Tibetic languages and are always realized with a “strong”
articulation unlike the preinitials which are often "weakly" pronounced. Initial sounds are usually followed by a vowel but they may be followed by a glide preceding the vowel.

- Glides

The glides /y/ and /w/ are pervasive in all Tibetic languages. One also encounters the glides /r/ and /l/. The former is found in some Tibetic languages such as Balti, Kyirong, Lhoke (in a marginal way), Cho-nga and in some dialects of Amdo (in contact with rGyalrongic languages) and one dialect of Kham (sProsgang), spoken in Rongdrak. The glide /l/ is essentially found in some Balti dialects.

Glides occur between the initial consonant and the vowel. They correspond to the letters R, Y and W which are subscribed to the radical letter in CT. In some cases (see below 'book', and 'pig'), they may also correspond to some specific innovations.

Ex: བྲས་ 'BRAS 'rice': /brä:/ (Kyirong), /blas/ (Western Balti); བྲུག་ 'BRUG 'dragon': /brug/ (Kyirong) 'dragon', /bluq/ (Western Balti), ཆུང་ 'DPa.Bu 'hero' (Am): /xawo/, སྤྲེ་ 'DPe.CHa 'book' (Am): /pweča/, སྤྲུ་ 'rtau 'grass' /hswa/ (Kh: Gyalthang), ཉ བྲུལ་ 'bird, poultry' /bya/ (Ba, Purik, Cho); བྲུག་ 'pig': /p'ye;// (Kh: dGudzong).

One should note that in CT, the L occurs as a LA.BTAGES: it is subscribed to the 'radical letter' in a similar way as R, Y and W in words such as བྲུས་ 'lama', ཁྲ 'mind', གླང་ 'ox', སླ་ 'thin, easy', 2) 'moon', ZLA 'moon', བྲངས་ (past) 'to take' which indicates that it is treated as a glide in the orthographic tradition. However, there is various evidence that points toward the interpretation that the L functions as an initial sound and it is preceded by preinitials S, G, B, etc. The strongest evidence is that the preradicals often disappear, which is not the case of the radical letters. In the above examples we often find in various languages and dialects reflexes of /b)la-ma/, /b)lo/, /b)lang/, /s)la/, /b)lang but never /b)la-ma/, /b)lo/, /g)lang/, /s)la/, /z)la/, /b)lang. For example, we find /lama/, /lo/, /lang/ but never: /lama/, /lo/, /lang/.

18. The fact that this verb has a present stem LEN is an additional proof that L is the radical as noted by Zeisler (pers. comm. 2020).
- **Syllable core**

The syllable core generally consists of a vowel, or a consonantic vowel. Vowels with any secondary articulations (including a length) can occupy the position “syllable core.” Additionally, some specific dialects can take a nasal as a syllable core, such as the Sogpho dialect of Kham (Rongdra). This case is quite exceptional. Ex.: *NGA-GI*’my’ /ng ce/ (Suzuki 2011c). In a few dialects, a number of diphthongs have appeared (see Qu 1987). However, the existence of diphthongs depends on how a syllable structure is defined. Hence there are accordingly differences among various phonological analyses. Our system generally does not allow diphthongs and note core vowel + glides.

Nasalized vowels are often analyzed as “vowel (syllable core) + neutralised nasal final” from the phonological point of view (see Kitamura 1977). However, in our description, we consider nasalised vowels as single units appearing in the syllable core.

- **Finals**

Classical Tibetan had ten distinctive simple sounds in the final position *g, ng, d, b, r, l, s, n, m,*’. Thus the set of final consonants is much more limited than the set of initial consonants (30).

Some modern Tibetic languages still distinguish ten final sounds. The final sounds are best preserved in Ladaks, Purik and Balti. Additionally, Ladaks, Purik and some dialects of Balti have preserved final clusters /-ks, -ngs, -ps, -ms/. These clusters have disappeared in all the other Tibetic languages. Furthermore these languages have developed innovative clusters such as /-ts, -rs, -ns, -ls/.

Ex. བོད་*LCAGS*’iron’: /łčaks/ (La), ལྷ་གས་*SNGAGS*’mantra, incantation’: /ngaks/ (Purik), /ngaks/ (La), གོ་ཁོ་*KHRIMS*’law’: /ʈims/ (La). In some cases, these clusters correspond to innovations as in Ladaks: མོ་*BTANG*’to send (past)’ /tangs/.

On the other hand, some eastern Tibetic languages, such as Baima, Zhongu and Drugchu as well as a few dialects spoken in marginal areas of Kham have only open syllables and entirely lack final consonants.

The plosives /k, p, ʔ, q, t/ are very frequent in the Tibetic language area. Final plosives are always voiceless and non-aspirated.
The final sounds /k/ and /p/ are particularly frequent in the western, southern and central areas but they are also found in southern Amdo (rNgawa). /k/ and /p/ correspond respectively to G and B in CT.

The glottal stop /ʔ/ is distributed in central and eastern areas as well as in some languages of the southern Himalayas such as Sherpa or Lhoke. The glottal stop often corresponds to G, D, B and S in CT. Ex. འིག་ G˟ YAG 'yak' /`yaʔ/ (Ü, Ts), /ˈyaʔ/ (Kh), བོད་ BOD 'Tibet, Tibetan' /ˈpʰeʔ/ (Kh), /ˈpʰɛʔ/ (Sh), ཀྲ མ ཚ KHAB 'needle' /ˈkʰaʔ/ (Kh). In many dialects of Kham the only final consonant is a glottal stop. Some eastern Tibetic languages such as Sharkhog and Pashi are also similar to Kham, but they have one more specific final, an epiglottal fricative /q/. Ex. བྱག་ LCAGS 'iron': /ʰcaʔ/ (Sharkhok).

The final /t/ is a little less frequent. It is attested in the western Tibetic languages (Ladaks, Balti, Zanhar, Spiti) as well as in some southern Himalayas (Jirel), Choča-ngāča and in some southern Amdo dialects (rNgawa, mGolok). Ex. དབྱུག་ BRUG 'dragon': /bluq/ (Ba), སློད་ SNOD 'pot, container': /not/ (Jirel), འིན་ STOD 'upper part' /not/ (La), བྲག་ SKAD 'language' /ket/ (Am: rNgawa), བྲག་ BRAG 'rock': /blaq/ 'rocky mountain' (Ba), ལྱིོན་ PHOG 'salary': /hoq/ (Am: rNgawa), སྲིག་ SNGAGS 'mantra, incantation': /ʰngaq/ (Am: rNgawa).

Final fricatives /χ, ʁ, ɣ, ɸ, β/ are very frequent in the northern areas (both in the west and in the east: Balti, Ladaks, Zanhar, Amdo, etc.).

The final /l/ is more frequent than the final /t/. It is essentially found in the western and southern Tibetan areas (Balti, Ladaks, Zanhar, Spiti, Tö, Sherpa, Jirel, Choča-ngāča), however, it is also attested in some Amdo pastoralist dialects. Historically, the final /l/ is nearly always the reflex of L in CT, but in some cases, such as some Amdo dialects, it may be derived from D.

Ex. བལ། BAL 'wool' /pal/ (Sherpa, Tö), /ˈpal/ (Cho), སྲིག་ SBRUL 'snake': /rul/ (La), /ˈrul/ (Sh), /ˈʔil/ (Spiti), པདྨ། DNGUL 'silver' /smul/ (Ba), /ˈmul/ (La), /ˈngul/
The distribution of the final /s/ is restricted to the western region of the Tibetic area, namely Ladaks, Purik, Balti and Zanhar. Ex. བརྒྱད་/BRGYAD 'eight' /'gyäl/ (Am), འབྲོད་/BOD 'Tibet' /'wol/ (Am), སྟོད་/STOD 'upper part': /'tol/ (Am).

Finally, one should mention the very rare occurrence of final /ɦ/ which corresponds to the reflex of s in CT. Ex. རས་/RAS 'cloth': /шедш/ (Hor: Bachen).

The final nasals /ng, n, m/ are ubiquitous, except in the eastern area (Thewo, Drugchu, Khöpokhok and many Kham dialects) where many dialects often lack final nasals.

Ex. ལམ་/LAM 'road': /lum/ (Am, Ba, La), /十六条/ (Tö, Sh, Dzongkha, etc.), བཤམ་/SMAN 'medicine': / słm/ (La, Ba), /mən/ (Am), /mən/ (Cho), / menstrual/ (Sp, Dz), etc.

7.2.4. Vowel harmony

A type of vowel harmony or vowel assimilation is frequently attested in Tibetic languages. Unlike Uralic, Mongolic and Turkic languages, which possesses a systematic vowel harmony attested in various morphological aspects, the phenomenon in Tibetic languages is generally limited to vowel changes within a word when forming compounds.

This phenomenon has been documented for Central Tibetan (see Tournadre & Sanda Dorje 1998; Haller 2000), for Kyirong (Huber 2005), Zeisler (2004) or Ladaks (Norman 2019). Concerning vowel harmony in the dialects of Ladakh:

"Many Ladakhi [Ladaks] words are compounds of two syllables, each of which is a word root. However, the vowel of the first syllable often changes under the influence of the vowel in the second syllable. Most speakers of Ladakhi don’t realize they are doing this. Although vowel harmony happens in all dialects of Ladakh, it is strongest in Nubra and Durbuk, and less strong in Leh, Nyoma, Kenhat and Sham." (Norman 2019: xxix)

Here are some examples provided by Norman (ibid.).
The syllable /ts’e/ becomes /ts’i/ when followed by a syllable containing a /i/ but remains unchanged in front of a syllable with an /a/ vowel:

\[
\begin{align*}
\text{ཚེ་རིང་} & \quad \text{TSHE.RING} /\text{tshe+ring}/ > /\text{ts’iring}/ \text{PSN} \\
\text{ཚེ་དབང་} & \quad \text{TSHE.DRANG} /\text{tshe+wang}/ > /\text{ts’ewang}/ \text{PSN}
\end{align*}
\]

The syllable /ton/ becomes /tun/ when followed by a syllable containing a /u/ but remains unchanged in front of a syllable with an /o/ vowel:

\[
\begin{align*}
\text{དོན་གྲུབ་} & \quad \text{DON.GRUB} /\text{ton+dup}/ > /\text{tun+dup}/ \text{PSN} \\
\text{དོན་ཡོད་} & \quad \text{DON.YOD} /\text{ton+yot}/ > /\text{ton-yot}/ \text{PSN}
\end{align*}
\]

The syllable /de/ becomes /di/ when followed by a syllable containing a /i/ but remains unchanged in front of a syllable with an /e/ vowel:

\[
\begin{align*}
\text{བདེ་སྐྱིད་} & \quad \text{BDE.SKYID} /\text{de+skyit}/ > /\text{diskit}/ \text{PSN} \\
\text{བདེ་ཆེན་} & \quad \text{BDE.CHEN} /\text{de+čen}/ > /\text{dečen}/ \text{PSN} \\
\text{ནོར་བུ་} & \quad \text{NOR.BU} /\text{nor+bu}/ > /\text{nurbu}/ \text{PSN} \\
\text{ནོར་བཟང་} & \quad \text{NOR.BZANG} /\text{nor+zang}/ > /\text{norzang}/ \text{PSN} \\
\text{ཆོས་སྐྱིད་} & \quad \text{CHOS.SKYID} /\text{čos+skit}/ > /\text{čuskit}/ \text{PSN} \\
\text{ཆོས་སྒྲོལ་} & \quad \text{CHOS.SGROL} /\text{čos+dol}/ > /\text{čosdol}/ \text{PSN}
\end{align*}
\]

The words for numbers are also affected by this phenomenon, which was already attested in Old Tibetan:

\[
\begin{align*}
\text{བཅུ་གཉིས་} & \quad \text{BCU.GNYIS} /\text{čugnyis}/ \text{’twelve’}, \quad \text{བཅུ་བདུན་} \quad \text{BCU.BDUN} /\text{čubdun}/ \text{’seventeen’}, \quad \text{བཅོ་ལྔ་} \quad \text{BCO.LNGA} /\text{čo-nga}/ \text{’fifteen’}, \quad \text{བཅོ་བརྒྱད་} \quad \text{BCO.BRGYAD} /\text{čobgyat}/ \text{’eighteen’}.
\end{align*}
\]

As noted by Norman (2019) “the numbers are even spelled this way in Tibetan, indicating that vowel harmony has been in effect for many centuries,” as we can see with the spelling of ‘ten’ in CT: BCU and BCO.

### 7.3. Suprasegmentals and tonogenesis of the Tibetic languages

In various Tibetic languages, suprasegmentals function as phonologically distinctive features. Among them the tones play an important role in the phonology of many
Tibetic languages. However, a few western and eastern Tibetic languages such as Amdo, Ladaks,19 Purik, Zanhar and Balti do not have any distinctive tones.

Scholars generally consider that Proto-Tibetic had no suprasegmental distinction because Tibetan script does not reflect any feature of suprasegmentals, unlike Burmese script. This hypothesis posits that the suprasegmental distinction emerged after the creation of Tibetan script, and many scholars have challenged the explanation for the tonogenesis,20 i.e. the process of the generation of suprasegmentals.

Since there is little consensus about the analysis of tones in Tibetic languages, scholars often use very different notation systems.21 In the linguistic literature, the case of the Lhasa dialect has been well discussed,22 but there is not a complete consensus about its tonal system in spite of the abundance of discussions.

Whatever opinion one holds about the Lhasa tone system, it can’t be used as a common model for the tonogenesis of Tibetic languages.23 We believe that the tonogenesis of Tibetic languages is not based on a single model but, instead, has multiple origins. In Tibetic languages, there are at least three types of suprasegmental distinction: 1) pitch tones (or simply “tone”), 2) register (phonation), and 3) stress (accent).

### 7.3.1. Pitch tones

Among the Tibetic languages with a suprasegmental distinction, the pitch tone system is the most widespread and the word “tone” often refers to this type. It is
attested mainly in Ü, Tsang, Tö, Spiti, Ladakh Jangthang, Kham, Hor, Lhoke (Sikkim), Sherpa (and other Tibetic languages of Nepal), Dzongkha, Choča-ngača (and other Tibetic languages of Bhutan) as well as some languages belonging to the eastern section, such as Čone and Baima.

It is worth noting that there are two contact zones where tonal and non-tonal languages are genetically closely related and in some cases even allow some mutual intelligibility. The two areas of contact between tonal and non-tonal are Ladakh (India) and Amdo (TAP, China). In Ladakh, the only territory where tonal dialects are found is the Jangthang areas of Nyoma and Durbuk, and along the gorges or rong of the upper Indus. These Jangkat and Rongkat dialects are in contact with the non-tonal dialects of Central Ladakh. The other area of contact is Amdo. In eastern Amdo, the tonal dialects of Čone (Eastern section) are in contact with the non-tonal dialects of Luchu. In Amdo, the second area of contact is located in southwestern Qinghai, where the non-tonal dialects of the Golok areas are neighboring the Kham tonal dialects of Yūlshûl Prefecture.

Such cases of contacts between non-tonal and tonal languages, which are closely related and form a geolinguistic continuum, are very rare among World languages. Moreover, when Old Tibetan was written down, it was a non-tonal language and two linguistic sections out of eight (the Northeast and Northwest sections, see Chapter 10) still have non-tonal languages. From a typological point of view, this situation is unique since, on the one hand, we have a very old written language that is non-tonal, and on the other hand, the development of various types of tonogenesis in the modern languages (see below). This rare combination is vital to a better understanding of tonogenesis in general.24

Let us examine now the Lhasa pitch system, which has been well described in the literature. The pitch in Lhasa Tibetan is associated with the Tibetan script in a quite

---

24. Indeed only very few compact families in the world have both tonal and non-tonal languages. This is the case, for example, of Bantu languages, mostly tonal, but with some exceptions such as Swahili (non-tonal). The same is true for Khmer, which is non-tonal but is surrounded by other Austro-Asiatic tonal languages.
regular and straightforward way. For simplicity’s sake, we divide the pitch system into two cases: high (H) and low (L) at the word-beginning position and at the word-final position. The H and L at the word-beginning position is determined by the nature of the "radical letter (MING, GZHI)." When the radical letter is voiced (G, NG, J, NY, D, N, B, M, DZ, W, 'ZH, Z, Y, L, R), then the pitch begins with a L tone, whereas when non-voiced, then the pitch begins with a H tone. If there are some preinitials (MGO, CAN and SNGON, JUG) preceding resonant letters (NG, NY, N, M, Y, L, R), the pitch will be H.

With low tones, there is another pattern RL which represents rising-falling tone.

The H and L at the word-final position is determined by the nature of the final letters (RJES, JUG and YANG, JUG): when the final letter is G, D, B, S, then the pitch is falling (L), whereas if the final letter is none or NG, N, M, 'L, R, then the pitch remains high (HH). Note that the second final (YANG, JUG) D, which only existed in Old Tibetan, makes a falling pitch in Lhasa Tibetan. For example the verb གྱིན་ PHYIN is pronounced in Lhasa with a falling tone, which is probably a trace of the old second suffix D roofing PHYIND /'e' in/.

The analysis and the notation of tones that we present below were elaborated by Kitamura (1977).

HH 'lo: གློ་ BA' lungs'  "ka: འབ་ K.A.B.A' pillar'
LH 'lo: བོར་ OR' paper money'  'ka: བསྒར་ B.S.GAR 'to install, fix'
HL 'lo: མོག་ SLOG' to give back'  'ka: བཀག་ B.KAG 'to stop'
RL 'lo: པོན་ LOG' to return'  'ka: དགག་ 'GAG' to be blocked'

Kitamura’s analysis parallels Tournadre and Sangda Dorje’s analysis (2003). However, the former approach is more phonetic whereas the latter is more phonemic. For example, the phoneme /-k/ does disappear entirely in the final position and is realized as a lengthening with a HL tone as indicated above: ['ka:] དགག 'B.KAG' to stop', but in a reading style, a glottal stop can clearly be heard [‘ka?] Moreover, inside a word, the final -k does not disappear and may be realized as [-k] or [y]. For example:

LCAGS 'iron' ['ca:] vs. LCAGS.PAR 'typewriter' ['cak-par]
RKANG.LAG 'limbs' ['kang-la:] vs. LAG.PA 'hand' ['lak-pa]
DGA'.ROGS'boy/girl friend' ['ka-ro:] vs. ROGS.PA'companion' ['rok-pa]

RKANG.LAG'limbs' [kang-la:] vs. LAG.LEN'hand' [lay-le]

DMAG'war' ['ma:] vs. DMAG.MI ['ma-mi]'soldier'.

For all these reasons, we will maintain the notation of a ? even in final position: BKAG' ka?/

Another small difference with Kitamura’s notation is the notation of the phoneme /-r/. Although it is true that the final /-r/ is often deleted and generates a lengthening of the preceding vowel (see above the example LOR 'paper money'), the variant [i] or even [r] is often heard. Moreover it is always pronounced in formal and reading styles:

MAR 'butter' ['ma:] or ['ma] or ['mar].

For some words, there is a minimal pair that may oppose the reflexes of L and R:

GSER 'gold' ['se:] vs. SEL 'eliminate' ['se:]

It is thus important to note the contrast. For this reason, we will maintain the notation of a phonemic /-r/. This position is also adopted by Goldstein (2001) in his New Tibetan-English dictionary of modern Tibetan.

Lhasa Tibetan suprasegmentals function as a word-tone system. The pitch pattern is determined for each word as a unit, not for each syllable unlike Mandarin Chinese or Vietnamese. As Sun (1997: 489) noted: "One of the most important generalization on Tibetan tone [...] is that the primary register [i.e. in our terminology “tone pitch”] is realized only on the initial syllable of the phonological word; all the other syllables are predictably high-registered [i.e. high pitch]." 26

When a word has more than two syllables, the pitch pattern applies only to the first two syllables and the subsequent syllables have a neutralized pitch. In normal

25. The word may include affixes and corresponds to a phonological word. See Kitamura & Nagano (1990) for detail.

26. However, the statement that "other syllables are predictably high-registered" does not apply to all the Tibetic languages. Even in Lhasa, the suffixes are usually pronounced with a neutralized (relatively low) pitch.
conversation, the most important feature of pitch pattern appears on the first syllable of a word. Thus the word-tone system can be more simple than the syllable-tone system.

The total number of pitch patterns in Lhasa is five: high-level (HH), rising (LH), falling (HL), rising-falling (LHL) and atonal. Thus, in this word tone system, it is sufficient to note one tone for a word even when it has two syllables (or more).

Example of LH: བོ་ MO 'she' /ˈmo/, བུ་ MO 'girl' /ˈp’umo/;
Example of LHL: ཆེ་བརྒྱད་ BRGYAD 'eight' /ˈkyä/;
Example of HH: བྱེ་མ་ལེབ་ PHYE-MEA-LEB 'butterfly' /ˈemalep/;
Example of HL: སྤྱན་རས་ཟིགས་ SPYAN-RAS-GZIGS 'Avalokiteśvara' /ˈčänrä:si/.

The fifth “atonal” pattern corresponds to words with a grammatical suffix: གང་NGA-LA /ˈnga-la/.

Among the basic rules of the tonogenesis of Lhasa Tibetan mentioned above, the rule regarding the “radical letter” is also applicable to most of the Tibetic languages with a tonal distinction. Some specific dialects have a different tonogenesis, e.g. in the Minyak Rabgang dialect group of Kham Tibetan, the resonants without a preinitial can be realized as H (high). In addition, Čone has a pitch tone system synchronically, but the tonogenesis is not similar to the Lhasa system. It could be associated with the register (phonation) type instead (see below).

There is great diversity in the suprasegmental realizations of closely related dialects. For example, even in the Tö dialects of Ngari, some words are pronounced with different tones depending on the given dialect or variety (see Qu & Tan 1983). Dzongkha for example has a basic opposition between high and low register and additionally some tone contour distinctions have been reported. However, “the contour distinction does not exist in all dialects of Dzongkha” (Tshering & van Driem 2018).

27. The word PHYE-MEA-LEB is perceived as literary in Lhasa dialect. The common word for ‘butterfly’ is གེམ་ཅེམ་མ་ /čemčemma/.

28. Further research about the tonal variation is needed to allow a better understanding of the tonogenesis and more generally tone systems.
Given this extremely high diversity, it is impossible in the present book to note down very precisely the tone for each single dialect or variety. However, to illustrate words in the main tonal languages, we will follow Kitamura’s method: ‘S (high register, level), ’S (high register-falling), ‘S (low register, slightly rising), ’S (low register, rising-falling).

7.3.2. Register

The distinction with a register difference is attested mainly in the eastern section, such as Sharkhog, Khodpokhol, and Pālkyi [Pashi]. A general definition of the term “register” in this context is provided by Zhu (2010): “The register is related to various phonation types which are divided into three zones: high, mid and low.” High register corresponds mainly to “falsetto”, mid register is associated with voiceless sounds, low register to voiced sounds. Zhu (2010: 76) considers that the register and the pitch height are independent variables, though both of them are strongly related.

In languages with a register system, the main oppositions are related to phonation types and the pitch differences are not distinctive. In previous studies, register systems have sometimes been analyzed as pitch tone systems.

The phonation itself includes many phonetic phenomena such as voicedness, aspiration and creakiness, among which some features are normally regarded as segmental (i.e. consonants and vowels). At least, the register distinction should not be confused with a pitch difference.

In the Tibetic languages with a register opposition, two registers are normally distinctive: mid and low. Suzuki (2008, 2009b) has mentioned a register distinction in several Tibetic languages of the eastern section, such as Sharkhok, Khöpokhok, Babzo and Zhungwa. In these languages, the main opposition is between a high register which is “marked” (with ’ before a word) and often characterized by a preaspiration, a creaky voice and usually high pitch, whereas in the low register these characteristics are absent.

Ex. མ་ LCAGS /’ca/ ‘iron’ vs. ལ་ GRO /’co/ ‘wheat’; རད་ GNAM /’nã/ ‘sky’ vs. ར་ RNA /na/ ‘nose’ (Sharkhok); བེ DE /’da/ ‘he’ vs. དོམ་ MDOG /’do/ ‘color’; ར NgA /nga/ ‘I’ vs. ད NgA /nga/ ‘five’ (Babzo)
Another Tibetic language of the eastern section, namely Chosrje, has been described as using breathy voice as a distinctive feature (Sun 2003b; see also Suzuki 2015a) which may correspond to a low register.

In synchrony, it seems that the difference of a register does not have a good correspondence with the Tibetan script. However, from a historical point of view, the relation between the register and the script was straightforward. The Proto-Tibetic phonetic system was divided into voiced (obstruent: G, D, B, DZ, ‘, Z, ZH and resonants: W, Y, R, L, NG, NY, N, M) and non-voiced (K, KH, C, CH, T, TH, P, PH, TS, TSH, S, SH, H, ‘) categories of sounds, which directly corresponds to the low and mid registers respectively.

The breathy phonation (noted by two dots under a vowel: ā) is attested in some dialects of Kham Tibetan, such as Yulshul and Khyungpo. The Khyungpo dialect group has a two-way suprasegmental system, and distinguishes both pitches and phonations. The breathy voice plays an important role, particularly in the verbal morphology (see for example the Khromtshang dialect, Suzuki 2010). This is an innovative development in this group, which cannot be easily related to the Tibetan script. Ex. GYEN /chē/ ‘uphill’, BYOS/‘chē/ ‘do (imperative)’.

7.3.3. Stress

According to Caplow (2009), “stress has played an important role in the development of tone in Tibetan.” The distinction of stress is attested mainly in Balti. Caplow (2016) and some other scholars believe that Amdo or even Lhasa also have a stress opposition but it may play a secondary role in the system. The distinction of stress only occurs in polysyllabic words.

Ex. in Balti: ཇུར་བ་ KHUR.BA ‘bread’ /kur̩ba/ vs. ཀྲུང་ KHUR.BA ‘to carry’ /k’ur̩ba/; ཀྲིམ་KLAD.PA ‘brain’ /datpa/ vs. ཀྲིམ་GLAD.PA ‘to be tired’ /datpa/; ཀྲིམ་ THAG.PA ‘rope’ /thakpa/ vs. ཀྲིམ་THAG.PA /thakpa/ ‘to grind’.

Ex. in Lhasa Tibetan: བསམ་པ་ BSAM.PA ‘to think’ /sampa/; བསླབ་ BSAM.PA ‘to teach’ /sampa/; བསྲལ་ SLAB.PA ‘instruction’ vs. བསྲལས་BSLAB.PA /lappə/ ‘to count’; བསྲལ་ RTNIS.PA ‘astrologist’ /tsipa/ vs. བསྲལ་RTNIS-P.A /tsipa/ ‘to count’. (A CD recording of these minimal pairs is available in Tournadre & Sangda Dorje 2003, 2009: 445.)
Sun (2003: 779) gives a minimal pair on the position of the stress in Zhongu as: /mé-rə/ ‘ideophone mimicking moving currents’ vs. དམར་བ་ DMAR-BA ‘to be red’ /me-rə/.

Some dialects of Kham spoken mainly in Yunnan, which possess a pitch-tone system synchronically, may also have a stress-like system of tonogenesis such as Tormarong (Dongwang) and mBalhag because they show a prominent stress as a phonetic status.

Ex. in mBalhag, ཕོལ་བ་ SOL.BA ‘coal’ /ɬeːja/ (no stress) vs. སྲིང་མོ་ SRING.MO ‘younger sister’ /ɬiwu/.

It seems however that the stress is a secondary feature, at least in some dialects such as Lhasa and mBalhag, and that the pitch pattern is fundamental. In the above examples of Lhasa Tibetan, the difference of stress may be explained by the fact that the second syllable of the verb is an atonal suffix.

### 7.3.4. Prosodic patterns

Every language has its own prosodic pattern and this is also true for the Tibetic languages. However, because the prosody does not function as a part of phonology, it may be easily influenced by other languages and often changes in the languages or dialects spoken at the periphery of this linguistic area.

Let us look at the example of the Lhasa dialect. The basic prosodic pattern of Lhasa is a trochaic meter (strong-weak) for the first two syllables. Even if one word has more than two syllables, the syllable(s) after the third do(es) not bear a tone. This trochaic type is basically attested in most dialects with a tonal distinction, such as Kham, Dzongkha and Sherpa, etc. It is also the case of the dialects with a register opposition: the most important distinction of register is quite always attested on the first syllable and its following syllables will not bear a register distinction.

In some Amdo dialects, prosody has been influenced by Mongolian. Similarly, other Tibetic languages have been influenced by the prosody of their neighbors. This is the case of Spiti or Balti both under the respective influences of Hindi and Urdu or Persian. In Central Tibet, the media have adopted a prosody which shows some influence of Putonghua Chinese.
Some dialects of Southern Kham have an iambic prosodic pattern (weak-strong) for the first two syllables (Suzuki 2013b) which resembles neighboring languages with a sesquisyllable (one-and-half syllable word), such as T’rung and Jingpho.

7.4. Historical phonology

In order to establish a classification of Tibetic languages, one must study the historical phonology and examine the sound developments attested between Old Tibetan and the modern languages. However, we should be careful when considering the sound correspondences. They do not indicate a direct shift from Old Tibetan to the modern languages, but imply a complex process of sound developments.

From the viewpoint of general historical linguistics, the existence of regular sound developments is one of the striking features of compact language families. This is also the case of the Tibetan language family.

To put it simply, when a word in a modern Tibetic language exhibits a sound change compared to its Classical form, theoretically all the other words of this language with identical sounds will undergo the same sound change (for details and examples, see Chapter 10.5). The regularity of sound changes has been shown in other language families, particularly the Indo-European family, which has become a canonical example. Given the instability of some linguistic phenomena, it is one of the astonishing features of World languages.

7.4.1. Regular reflexes of Old and Classical Tibetan

We will present here the basic reflexes of Classical Tibetan found in the main languages. In the case of exceptional reflexes found only in some specific examples of one dialect, we will not mention them in the following paragraphs. Some of the cognate words listed below may have a slightly different meaning in the modern languages compared to their classical sense. Rare innovations that are specific to some regions or even to some dialects are marked with the pound sign (#).

- **Simple consonants**

  The reflexes of simple consonants will be listed according to the following groups:

  1. obstruents: KA, KHA, GA, CA, CHA, JA, TA, THA, DA, PA, PHA, BA, TSA, TSHA, DZA, T, SA, ZA, SHA, ZHA, HA, ’A
  2. nasals: NGA, NYA, NA, MA
  3. non nasal resonants: LA, RA, YA, WA
<table>
<thead>
<tr>
<th>CT</th>
<th>Pronunciations</th>
<th>Examples of words/morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>KA</td>
<td>k, #g, #q</td>
<td>ཀཀ་ KAKHA 'alphabet', ཀཀ་ KABA 'pillar', ཀོ་ KO.BA 'hide'</td>
</tr>
<tr>
<td>KHA</td>
<td>k', #s, #q'</td>
<td>མི་ KHA 'mouth', མི་ KHO '3SG', མི་ KHANG.PA 'house', བོན་ KHAR 'stone'</td>
</tr>
<tr>
<td>GA</td>
<td>k, k', # &amp; #q</td>
<td>བག་ GANG 'one', བག་ GALLE 'slowly', བག་ GO 'hear', བག་ GOS 'cloth'</td>
</tr>
<tr>
<td>CA</td>
<td>Չ, Չ, # &amp; #s, #t</td>
<td>ԿԱ LAAG 'thing', Չ 'what'</td>
</tr>
<tr>
<td>CHA</td>
<td>Չ', #g', #sh', #l', #ts'</td>
<td>ՉՇU 'water', ՉՇU 'pair', ՉՇ 'big', ՉՇ LAHNG 'Tibetan alcohol'</td>
</tr>
<tr>
<td>JA</td>
<td>Չ, j, #sh, #ts, #t</td>
<td>ՉԱ tea, ԺՈ 'lord', 旄 'robber'</td>
</tr>
<tr>
<td>TA</td>
<td>t</td>
<td>ՏIL 'sesame', ՏILS 'Mount Kailash'</td>
</tr>
<tr>
<td>THA</td>
<td>t'</td>
<td>ՏHUN 'short', ՏHOS 'roof', ՏHANG 'plain'</td>
</tr>
<tr>
<td>DA</td>
<td>t, t', d, s, #q</td>
<td>ԴԱ LAITA 'now', ԴՕM 'bear', ԴՈU 'poison', ԴՈU 'conch'</td>
</tr>
<tr>
<td>PA</td>
<td>p, #w, #w</td>
<td>ՊԱ 'print', ՊԱ 'lap', ՊԱ 'LOTUS'</td>
</tr>
<tr>
<td>PH</td>
<td>p', #h, #f, #ʃ</td>
<td>ՓԱ 'pig', ՓԱ 'parents', ՓԱ ' THOS 'useful', ՓԱ 'PHOR.PA 'wooden bowl'</td>
</tr>
<tr>
<td>BA</td>
<td>p, p', b, #w</td>
<td>ԲԱ 'wood', ԲՈ 'son', ԲԱ 'cow', ԲΟD 'Tibet'</td>
</tr>
<tr>
<td>TS</td>
<td>ts, #s</td>
<td>ՏՍ-DAN 'sandal wood', ՏՍ 'onion'</td>
</tr>
<tr>
<td>TSHA</td>
<td>ts', #s'</td>
<td>ՏՇԱ 'hot', ՏՇԵ 'life span', ՏՇՈNG.PA 'merchant'</td>
</tr>
<tr>
<td>DZA</td>
<td>ts, dz, ndz</td>
<td>ՏՍ-DZ ARG 'urgent', ՏՍ 'shameful'</td>
</tr>
</tbody>
</table>
| ZA | ?, O | ԶԱ-MAN 'mother', ԶԱ-MAN 'mother', ԶԱ-MAN 'liquor', ԶԱ-MAN 'mother', ԶԱ-MAN 'mother',
### Chart VII.4. – Reflexes of the nasals

<table>
<thead>
<tr>
<th>CT</th>
<th>Pronunciation</th>
<th>Examples of words/morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>NG</code></td>
<td>ng, #ma</td>
<td>ག་‘me, I’, གུ་‘to cry’, གོ་‘face’</td>
</tr>
<tr>
<td><code>NY</code></td>
<td>ny, #ng, #n</td>
<td>ནའ‘fish’, ནགའ‘to lay down, sleep’, ནོ‘to buy’, ནུ‘sun’</td>
</tr>
<tr>
<td><code>N</code></td>
<td>n</td>
<td>ཉའ‘to be sick’, ཉག‘black’, ཉན‘barley’, ཉུ‘breast’</td>
</tr>
<tr>
<td><code>M</code></td>
<td>m</td>
<td>བར‘butter’, བོ‘many’, བོ‘not to be’</td>
</tr>
</tbody>
</table>

### Chart VII.5. – Non nasal resonants

<table>
<thead>
<tr>
<th>CT</th>
<th>Pronunciation</th>
<th>Examples of words/morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>L</code></td>
<td>l, #y</td>
<td>བ‘mountain pass’, བས‘work’, བོ‘year’, བོ‘road’</td>
</tr>
<tr>
<td><code>R</code></td>
<td>r, #h</td>
<td>ར‘mountain’, རུ‘goat’, རུ‘bone’, རོ‘deep valley’, ‘gorge’ (by extension ‘cultivated land’)</td>
</tr>
<tr>
<td><code>Y</code></td>
<td>y, #zh, #sh</td>
<td>ཡི་‘up; upward’, ཡི་‘place, village’, ཡི་‘letter, syllable’, ཡོ‘to have’</td>
</tr>
<tr>
<td><code>W</code></td>
<td>w, Ø, #v, #h, #u</td>
<td>རོ‘fox’</td>
</tr>
</tbody>
</table>
- **The complex initials**

  The reflexes of the complex initials will be listed according to the following groups:

  1. radical+postradicals: LA, RA, YA, WA
  2. preradicals+radical (+postradicals)
     2.1. obstruent radicals
        2.1.1. preradical GA, DA, BA, RA, LA, SA
        2.1.2. preradical MA, 'A
     2.2. resonant radicals

| Chart VII.6. – Radical+postradicals |
|-----------------------------|-----------------------------|
| **CT** | **Pronunciation** | **Examples of words/morphemes** |
| WA | Ø, W | 'horn', 'hat', 'shwa' |
| YA | c, c', j ky, ky', gy ts', dz sh, sh' | 'you', '(mud) wall', 'house' |
| WA | sh, sh', zh pe, pe', bj pts, pts', bdz py, p'y, by ts', dz s, s', z sh, sh' | 'rich', 'sand', 'bird', 'flour' |
| MYA | ny mny #n | 'experience', 'fire' (OT), 'to experience, to taste', 'sprout', 'person' (OT), 'hell' |

29. The wasur is not pronounced in most languages. Even in the languages which do have a reflex of the /w/ (e.g. Gyälthang, Jol), it is not pronounced in some words such as TSHWA 'salt' or GRWA 'monk'.

## Phonological outline of the modern Tibetic languages

<table>
<thead>
<tr>
<th>CT</th>
<th>Pronunciation</th>
<th>Examples of words/morphemes</th>
</tr>
</thead>
</table>
| མ་KR.A | kr, kʼ, gr | མ་KHHAG ‘blood’,  
| མ་KHRA | t, tʼ, d,  
| མ་GRA | ʈ, ʈʼ, ɖ | མ་GRI ‘knife’,  
| མ་DRA | t, tʼ | མ་DRIS ‘ask’,  
| མ་PHA.RA | pr, pʼ, br | མ་DRU ‘child’,  
| མ་BRA | t, tʼ, d,  
| མ་PHA.RA | ʈ, ʈʼ, ɖ | མ་BRAG ‘rock/cliff’,  
| མ་SRA | sh, #/#/ | མ་SRAG ‘hard/solid’,  
| མ་SRA | s, hs | མ་SRUNG ‘protect’,  
| མ་SRA | ʈ, str | མ་SRAG ‘life’,  
<p>| མ་SRA | མ་SRAN.MA ‘bean’ |</p>
<table>
<thead>
<tr>
<th>LA</th>
<th>KLA</th>
<th>GLA</th>
<th>BLA</th>
<th>RL.A</th>
</tr>
</thead>
<tbody>
<tr>
<td>ཛལ་</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>གལྱ</td>
<td>ʰl</td>
<td>ʰl</td>
<td>ʰl</td>
<td>ʰl</td>
</tr>
<tr>
<td>ངལ་</td>
<td>y</td>
<td>y</td>
<td>y</td>
<td>y</td>
</tr>
<tr>
<td>རལ་</td>
<td>#l</td>
<td>#l</td>
<td>#l</td>
<td>#l</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ZLA</th>
<th>Ցླ་</th>
<th>Ցློ་</th>
<th>Ցློས་</th>
</tr>
</thead>
<tbody>
<tr>
<td>཭ྲལ་</td>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>ཝླྲ</td>
<td>ʰd</td>
<td>ʰd</td>
<td>ʰd</td>
</tr>
<tr>
<td>དྲླ</td>
<td>dz</td>
<td>dz</td>
<td>dz</td>
</tr>
<tr>
<td>དྲ</td>
<td>dz</td>
<td>dz</td>
<td>dz</td>
</tr>
<tr>
<td>བ</td>
<td>l</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td>མ</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>རླ</td>
<td>#y</td>
<td>#y</td>
<td>#y</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SLA</th>
<th>སླ་</th>
<th>སློབ་</th>
<th>སློང་</th>
</tr>
</thead>
<tbody>
<tr>
<td>ངླ་</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ཛྷླྲ</td>
<td>ʰl</td>
<td>ʰl</td>
<td>ʰl</td>
</tr>
<tr>
<td>མླྲ</td>
<td>ts</td>
<td>ts</td>
<td>ts</td>
</tr>
<tr>
<td>བ</td>
<td>s</td>
<td>s</td>
<td>s</td>
</tr>
</tbody>
</table>
### Chart VII.7. – Preradicals + radicals: obstruent radicals

<table>
<thead>
<tr>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHART VII.7. – Preradicals+radicals: obstruent radicals</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cтан.t</th>
<th>Pronunciation</th>
<th>Examples of words/morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>C (no trace)31</td>
<td>དཀར་ ‘white’, ཡུད་ ‘nine’, བཀའ་ ‘difficult’, བཀའ་ ‘to like, rejoice’</td>
</tr>
<tr>
<td>DA</td>
<td>C (no trace)</td>
<td>དེར་ ‘main’, བཅུ་ ‘ten’, དཀའ་ ‘to kill’, དཀའ་ ‘to explain, to talk’, བཀའ་ ‘order’ (H), བཀའ་ ‘to share’, བཀའ་ ‘excellent’, བཏས་ ‘lady, queen’</td>
</tr>
<tr>
<td>BA</td>
<td>C (no trace)</td>
<td>དཀར་ ‘white’, ཡུད་ ‘nine’, བཀའ་ ‘difficult’, བཀའ་ ‘to like, rejoice’</td>
</tr>
<tr>
<td>RA</td>
<td>C (no trace)</td>
<td>བཞི་ ‘four’, བདུན་ ‘seven’, བཅུ་ ‘ten’, དཀོན་ ‘to kill’, དཀོན་ ‘to explain, to talk’, བཀའ་ ‘order’ (H), བཀའ་ ‘to share’, བཀའ་ ‘excellent’, བཏས་ ‘lady, queen’</td>
</tr>
</tbody>
</table>

30. In this column, ‘C’ designates the radical consonant. Allophones of this consonant are neglected for the simplicity’s sake. Some peculiar reflexes are mentioned at the end of this tabular.

31. When the predical yields no reflex, it may however modify the pronunciation of the radical. In languages which have an opposition between voiced and voiceless (Dz, Lho) radical consonants, the preradicals are associated with voiced consonants whereas the plain radical is devoiced (sometimes with breathy voice). In tonal languages (Ü, T’s, Lho) with an opposition between aspirated (low tone)/ and non-aspirated (low tone), the reflex of the preradical triggers a lack of aspiration whereas the plain radical is aspirated (or with breathy voice).
In some dialects such as Zanhar, the preinitial S triggers a fricative initial such as /x, f, θ, ð/, etc.
### Chart VII.8. – Preradicals+radicals: resonant radicals

<table>
<thead>
<tr>
<th>CT</th>
<th>Pronunciation</th>
<th>Examples of words/morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ཞ</td>
<td>(no trace)</td>
<td>ཆིག་ 'two', ཊེགས་ 'sky', ཁག་ 'yak'</td>
</tr>
<tr>
<td>ཝ</td>
<td>(no trace)</td>
<td>ཆིད་ 'silver, money', ལྡེར་ 'red'</td>
</tr>
<tr>
<td>ཞ</td>
<td>(no trace)</td>
<td>ཁིམ་ 'sweet', ངོའ་ 'together', ཁགས་ 'bride, daughter in law'</td>
</tr>
<tr>
<td>ཞ</td>
<td>(no trace)</td>
<td>ཇེ་ 'drum', ཇཱ་ 'old', ླྀྲ་ 'ear', ཞྱ་ 'injury'</td>
</tr>
<tr>
<td>ཞ</td>
<td>(no trace)</td>
<td>ཁྱ་ 'five'</td>
</tr>
<tr>
<td>ཞ</td>
<td>(no trace)</td>
<td>མགོ་ 'early', ཁླི་ 'heart', གདམ་ 'nose', གནས་ 'medicine'</td>
</tr>
</tbody>
</table>

In many languages, some combinations have very specific reflexes. They include: ང་ སྲ་. The combination ང་ སྲ་ corresponds to /w/, /Ɂ/, /ʁ/, or /b/; སྲ་ corresponds to /l'/ or /l/, and even to /h/ in some dialects.

Some languages such as Amdo have specific reflexes on ང་ སྲ་: ང་ སྲ་ corresponds to /χw/, ང་ སྲ་ to either /ʃ/ or /ʃ/, ང་ སྲ་ to either /w/, /v/, /w/, or /v/. The reflexes of ང་ སྲ་ and ང་ སྲ་ are also very diverse in the modern languages and dialects.

33. In tonal languages, it triggers a difference in pitch. This is also true for the letters D, M, R, L. below.
Rhymes
We will first examine rhymes with a vowel (i.e. open syllable) and then rhymes with a vowel+final consonant (i.e. closed syllable):

**Chart VII.9. – Open rhymes V**

<table>
<thead>
<tr>
<th>CT</th>
<th>Pronunciation</th>
<th>Examples of words/morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>བོ་lhA</td>
<td>ཐག་</td>
<td>བོ་དམག་བཞི་, བོ་དམག་བཞི་'salt'</td>
</tr>
<tr>
<td>གྲ་ lhI</td>
<td>ཉི་</td>
<td>གཟི་'agate', གཟི་'person', གཟི་'dog'</td>
</tr>
<tr>
<td>གྲས་ lhU</td>
<td>ཉི་</td>
<td>ག्यུ་'turquoise', ག्यུ་'taste, corpse', ག্যུ་'sharp'</td>
</tr>
<tr>
<td>སྣང་ lhE</td>
<td>ཉི་</td>
<td>སྣང་'brass', སྣང་མིག་'eye', སྣང་'taste, corpse', སྣང་'sharp'</td>
</tr>
<tr>
<td>བོ་ lhO</td>
<td>ཉི་</td>
<td>བོ་'fire', བོ་'each', བོ་'top, summit, point'</td>
</tr>
</tbody>
</table>

**Chart VII.10. – Closed rhymes VC**

<table>
<thead>
<tr>
<th>CT final</th>
<th>Pronunciation</th>
<th>Examples of words/morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>བོ་ lhA</td>
<td>བོ་</td>
<td>བོ་'fire', བོ་'each', བོ་'top, summit, point'</td>
</tr>
</tbody>
</table>

34. In the following column, ‘V’ designates varieties of the vocalic quality, which are neglected for simplicity’s sake. Ɂ indicates a nasalised vowel and V a long vowel.

35. It is worth noting that the phonetic variants presented here (fricative velar [x] and uvular [χ] as well as plosive velar [k] and uvular [q]) do not have a phonemic value when they are attested in a single dialect.

36. In tonal languages, it sometimes triggers a difference in tone contour. This is also true for the final suffixes: -GS, -NG, -NGS, -D, -N, -R, -BS, -MS, -M, -S, -E, -X.
<table>
<thead>
<tr>
<th>प</th>
<th>GA + स</th>
<th>नाग्स 'forest', ग्जिग्स 'to see, look', ड्रुग्स 'breath', र्ग्त्सेग्स 'to pile up', रोग्स 'companion'</th>
</tr>
</thead>
<tbody>
<tr>
<td>न्ग</td>
<td>NGA</td>
<td>नाङ्ग‘in, within, house’, ग्लिंग‘island, mansion’, चुंग‘small’, त्सेंग‘over’, सोंग‘to go (past, imp.)’</td>
</tr>
<tr>
<td>न्ग</td>
<td>NGA + स</td>
<td>झाङ्ग्स ‘copper’, खुंग्स ‘source’, थेंग्स ‘time’, ल्जूंग्स ‘region’</td>
</tr>
<tr>
<td>ढ</td>
<td>DA</td>
<td>मेद‘negation of existential verb’, ग्न्यिंद‘fall asleep’, थुड‘cheese, cake’, शेंद‘strength’, स्तोद‘upper part’</td>
</tr>
<tr>
<td>न्स</td>
<td>NA</td>
<td>म्खांग‘agent’, सों‘seed’, झ्रिंग‘cloud’, र्जुंं‘lie’, स्तों-मो‘fingernail’</td>
</tr>
</tbody>
</table>

37. When they are attested in a single dialect, the variants [p] and [f] do not have a phonemic value.
### CT final

<table>
<thead>
<tr>
<th>CT final</th>
<th>Pronunciation</th>
<th>Examples of words/morphemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ོ་ལེ</td>
<td>Vp</td>
<td>བབས་ 'means', དབྱིབས་ 'shape', དོིའིཛ 'covering, case', ཁེ་ཞེ་ 'to come/go (H)', རོ་ཙྦས 'ability, strength'</td>
</tr>
<tr>
<td>ོ་བལ</td>
<td>Vf</td>
<td>བོམ་ 'orter', དོིཾ 'to sink, to be absorbed', མི 'one hundred thousand', མི་ 'to sew', རོཾ 'to have the time'</td>
</tr>
<tr>
<td>ོ་བལ + ོ་སའ</td>
<td>Vm</td>
<td>བཙམ་ 'Kham region', དེརིམ་ 'law', བཙམ་ 'hand/leg contract', མེས་ 'mind', བཙམ་ 'habit'</td>
</tr>
<tr>
<td>ོ་ལེ</td>
<td>Vr</td>
<td>བན་ 'rent', དོིཾ 'change', འཇོད་ 'gold', ཚོར་ 'hoe'</td>
</tr>
<tr>
<td>ོ་ལེ</td>
<td>Vl</td>
<td>བཤལ་ 'clear', བིམ་ 'cool', ངི། 'glass', ཁོལ་ 'boil', བཤིི 'snake'</td>
</tr>
<tr>
<td>ོ་སའ</td>
<td>Vs</td>
<td>བཤབས་ 'work, activity, karma', དིུས་ 'calculate', ཁེ་ཞེས 'day', བཙམ་ 'incense'</td>
</tr>
</tbody>
</table>

38. In the following column, 'V' designates varieties of the vocalic quality, which are neglected for simplicity’s sake. Ṽ indicates a nasalised vowel and Ṽː a long vowel.
At the period of OT, another second ད་ finals existed. In the modern languages, this feature has still remained as a tonal development in some limited dialects such as Lhasa.

7.4.2. Summary of the main differences

As we can see from the above charts, for simple initials, the variation in modern Tibetic languages is relatively limited. This is true particularly for the following initial sounds: ཆ་ ka, ར་ kha, ག་ ga, བ་ nga, ལ་ མ cha, ལ་ མ ja, ཡ་ nya, ཤ་ ta, ལ་ tha, བ་ da, བ་ na, ན་ tsa, ན་ tsha, ན་ dz, ན་ ma, ན་ sa.

From the chart, we see that the main variations are between voicedness (k, t, p, s, z versus g, d, b, z, zh) and aspiration (k, č, t, p, sh, s versus k', č', t', p', sh', s'). There are also suprasegmental variations (distinct tones or absence of tone). Apart from these frequent variations, a few sounds have undergone a specific evolution. These specific pronunciations are listed below. They concern only initial consonants:

Obstruents

- བ་ JA is realized as a simple fricative /sh/ in some Hor dialects (Amdo County) and as a retroflex affricate /ʈʂ/ in some Gyālthang Kham.
- བ་ DA is in some rare cases pronounced as a fricative. It is realized as a denti-alveolar fricative /s/ in à Hualong Amdo and as a retroflex fricative /ʂ/ in Pāmbar Kham (Pekar xiang). These pronunciations are extremely marginal across the Tibetic area.
- བ་ PA is pronounced as /w/ or /χw/ in some Amdo dialects.
- བ་ PHA is pronounced as /h/ or /χ/ in Amdo.
- བ་ BA is pronounced as /w/ in Amdo.
- བ་ ZHA is realized as a velar /x, χ/ or even as retroflex /ʂ/ in Kham dialects.
- བ་ SHA is pronounced as /x'/ in Amdo and many Kham dialects, or /χ'/ in some southern Kham dialects.
- བ་ SA is pronounced as an aspirated /s'/ in Amdo dialects.
- བ་ HA is realized as /χ'/ in some southern Kham dialects.
A was probably pronounced in OT as a fricative glottal ɦ. Apart from glottal realizations (ɦ, h, ʔ), modern languages of Eastern Tibet have also developed velar and uvular reflexes: /ŋ, ʁ/.

Nasals
- ʰʰNGA is exceptionally realized /ma/ in Western Tibetic languages such as Balti, Ladaks and Spiti.
- ʰʰNYA is pronounced /ng/ or even /n/ in some dialects of southern Kham. These pronunciations are extremely marginal across the entire Tibetic area.

Non nasal resonants
- ʰʰLA is realized as /y/ in a number of southern Kham dialects.
- ʰʰYA is realized as a fricative sound /s/, /ʃ/ or /zh/ in some southern Kham dialects. These pronunciations are very marginal across the Tibetic region.
- WA is realized as /wa/ or Ø in most western, central and southern regions. However in eastern Tibet, velar and uvular realizations /ŋ, ʁ, ɦʁ/ are regularly attested.

From the above list, we can easily see that the main phonetic innovations concerning initial simple consonants are essentially located in Eastern Tibet. Other regions such as Central and Western Tibet as well as the Tibetic-speaking of the south and western Himalayas have not developed such radical phonetic innovations.

For complex initials, there are many specific evolutions and phonetic diversity is quite high as we can see from the above chart. The phonetic variation is remarkable for the combinations involving postradicals. This is particularly true for the following combinations:

- ʰʰKHYA and ʰʰGYA
- ʰʰPHYA and ʰʰBYA
- ʰʰKHRA and ʰʰGRA
- ʰʰPHRA and ʰʰBRA
- ʰʰSRA
7.4.3. Classical Tibetan forms corresponding to modern sounds

Here we will show how modern sounds demonstrated in the two charts of pandialectal phonetic description (7.1) correspond to CT forms. This is the reverse approach of section 7.4.1: one observes the modern forms and look at their classical correspondences. In this section, we will not mention the obvious correspondences: for examples, the modern sounds /p, p’, b/ or /k, k’, g/ respectively comes from CT P, PH, B and K, KH, G and we will deal only with peculiar forms and will restrain our remarks to the sounds which have a phonemic status.

We will notably examine sounds which did not exist in OT or CT as well as sounds which did exist in OT and CT but were obtained in the modern languages by specific reflexes. The sounds are listed according to the following order:

a) obstruents (divided in articulatory position; from labials to laryngeals)

b) resonants
The labials

The fricative sounds /f/ and /v/ may derive respectively from the initial combinations SP and SB in some Amdo dialects. For example སྤུན་/spuN/ (Am), ‘brother’; སྦལ་པ་/sbal.ba/ (Am), ‘frog’. In Zanhar, the above combinations as well as DP may also yield a fricative labial sound: སྤེར་/dper.a/ ‘talk’.

They may also be derived from the final B in CT. སྤྱ་/sny/ ‘west’.

The denti-alveolars

From a diachronic point of view, the sounds /ts/ and /dz/ are the reflexes of Literary Tibetan TS and DZ but in some languages of Eastern Tibet, they also result from the combination of velars K, KH, G or labials P, PH, B when followed by R or Y. For example མཁྱི/khyi/ ‘dog’; /tsa/, རྒྱ/rgya/ ‘Chinese’: /hΔza/, བྱ/bya/ ‘bird, chicken’: /p/tsa/, གྱེ/gye/ ‘open’: /p/ts'i/, etc.

In northern Kham, Hor and Amdo, the sounds /ts/ and /dz/ may also result from the reflexes of SLA and ZLA. For example སླ/sla/ ‘easy’: /tsa/ (Ho), /htsa/ (Am, Kh), འཟླ/zla/ ‘month’: /hΔza/ (Ho), /hΔza/ (Am, Kh).

In Zanhar, the fricatives /θ/, /ð/ result from the combination of CT präradicals G, S, R, with dental T, D. For example རྟ/rta/ ‘horse’, བདེ/bde/ ‘beautiful’, བདོ/bdong/ ‘face’.

Historically, the aspirated fricatives /s'/ are the reflexes of fricatives S (without präradicals) while their non-aspirated counterparts are derived from Z (without präradical) or clusters involving S preceded by präradicals (GS-, BS-, etc.).

For example in Amdo: ས/sa/ ‘earth’ vs. ཁ/za/ ‘to eat’.

The sounds /s/, /z/ are usually the reflexes of Literary Tibetan S and Z. However, in some languages (Hor, southern Kham), they may be related to other sources such as the combinations of P, PH, B followed by a glide y. Other combinations such as SKYA may also yield /s/ in a few southern Kham dialects (Chakthreng).

For example མཉ/nya/ ‘mother’; བོ/bo/ ‘to eat’.
The retroflexes

From a historical point of view, Old Tibetan did not have retroflex plosive and they are not noted in the thirty consonants of the basic Tibetan alphabet. The retroflex sounds t̚, t̚', d̚ mainly appeared later as a combination of velar K, KH, G, dental D or labial P, PH, B with R.

For example ཁྲག  KHRAQ 'blood': / t̚aʔ/ (Ū, Kh), /t̚ak/ (La). ལྨ  BRAG 'rock': /t̚aʔ/ (Ū, Kh), /t̚ak/ (La); ོྲ DRUG 'six': /t̚uʔ/ (Kh), /t̚uk/ (La).

The retroflex /ʂ/ is essentially a reflex of the combination śr found in Classical Tibetan. This is the case in Ladaks, Zanhar, Sherpa, Tsang, many Tö dialects and most of Amdo dialects.

For ex: འྲྲན  SRAN  'pea': / s̚elma/ (Sh.), / s̚änma/ (Ts, Tö), / s̚anma/ (Am), ོྲུང  SRUNG 'to guard': / s̚oŋ/ (Am), / s̚un/ (La), / s̚un/ (Ts); ཉྲེ  SRAB-(MO/PO) 'thin': / s̚eme/ (Sh), / s̚awo/ (Am), / s̚appo/ (Ts, Tö), / s̚ablo/ (Hor); འཛེ  SREG 'to burn': / s̚aʔ/ (Ts), / s̚ak/ (Ho), / s̚ak/ (La), /s̚ag/ (Am).

The prepalatals

The aspirated /sh'/ are the reflexes of SH or the cluster PHY, whereas /sh/ are derived from ZH (without preradical) or clusters such as BY, SPY, SKY, etc.

For ex: ཡི་SHA /sh'a/ 'meat'.

From a diachronic point of view, /ɕ/, /ʑ/ and /j/ are the reflexes of Literary Tibetan C, CH and J but in many languages and dialects throughout the entire area, they can also be the result of various combinations, namely labials P, PH, B or velar K, KH, G followed by Y or even R. For ex: བྲྲི  KHYOD 'you': / s̚eʔ/ (Kh), འཛེ SKRA 'hair': / s̚a/, འཛེ SKYID.PO 'pleasant': / s̚hipo/, etc.

In many Western, Central and Southern Tibetic languages, the fricative /sh/ and /zh/ correspond to the reflex of Lit. Tib ʂ  SH and ʐ  ZH, but in most Eastern languages, they correspond to labials P, PH, B when followed by R or Y or even the reflexes of combinations such as SKR, SKY or SPY. For example དཀ་ BYLA 'mouse': /s̚awa/, ར བྲYA 'bird', 'chicken': /s̚a/, འཛེ SKRA 'hair': /s̚a/, འཛེ SKYID.PO 'pleasant': /s̚hipo/, etc.
The palatals and prevelars

From a diachronic point of view, these sounds are the reflexes of various combinations, such as $K, KH, G$ followed by $Y$ or even $R$. In some limited dialects they correspond to the reflexes of $C, CH, J$ and $P, PH, B$ followed by $R$. For example རྒྱུ་ RGYU /-gyu/ ‘nominalizer’ (Ú, Ts), /-ŋa/ ‘nominalizer’ (Am); རྒྱོ་ GRO 'wheat' /co/ (Am); རྒྱུ་ CHU ‘water’ /cq’u/ (Kh: Chamdo, Sangdam, Drugchu), རྒྱུ་ BRI 'female yak' /’ŋa/ (Kh). The fricative sounds /çi, či, ’j/ are the reflexes of various combinations, respectively: $PR, PHR, BR$ in Kham (Gyälthang). Additionally, the sound $çi$ also comes from the combination $SL, LH$.

For ex: རྒྱུ་ BRAG ‘cliff’ /’ça/ (Kh: Gyälthang, etc.).

The fricative $ŋ$ and $x$ are derived respectively from $ZH$ and $SH$.

For ex: རྒྱུ་ SHA /’xa/ ‘meat’ (Kh: Khyungpo).

The velars

From a historical point of view, the frivatives /x/, /x’/ and /ɣ/ are derived from $SH$ and $ZH, G$ in the final position as well as from some combinations such as $SK, RK, RG$.

For ex: རྒྱུ་ SHA ‘meat’: /’xa/ (Kh), རྒྱུ་ BSHAD ‘to tell’: /’xa/ (Kh: Derge, Bathang), རྒྱུ་ ZHWA ‘hat’: /’xa/ (Kh: Derge, Bathang). In Zanhar, glottal may also correspond to the combination of a preinitial and a radical such as $SK$ or $RK$: རྒྱུ་ SKANG.PA ‘leg’: /xangpa/ (Za), རྒྱུ་ SKAD ‘language’: /xat/ (Za); རྒྱུ་ RGAD.PO ‘old man’ /’xatpo/ (Za).

The uvulars

Historically speaking, uvular sounds are essentially the reflexes of Classical Tibetan $KH, G, DP, DB, W$’. The last four cases are found in Amdo as well as apart of the Eastern Tibetec languages. Ex. རྒྱུ་ KHATA ‘crow’: /q'ata/ (Am), རྒྱུ་ BROG.PA ‘/droqwa/ (Am) ‘pastoralist’, རྒྱུ་ STAG ‘tiger’: /’tay/ (Am), རྒྱུ་ LCAGS ‘iron’: /’tay/ (Ba); རྒྱུ་

39. Gyälthang, etc.
40. The voiceless fricative is sometimes pronounced as a glottal /h/ in Zanhar.
DPE 'example': /χwe/ (Am), ལྷ་པ་ 'power': /kang/ or even /wang/ (e.g. rNgawa), ང་ 'fox'> /wa/, ངོ་ 'light' /kt/, etc.

There is another origin of uvulars, which is merely attested in one vernacular of Kham (Myigzur): སྒོ་ 'fox'> /ɦʁa/, འོད་ 'light' /ʁɔt/, etc.

There is another origin of uvulars, which is merely attested in one vernacular of Kham (Myigzur): སྒོ་ 'fox'> /ɦʁa/, འོད་ 'light' /ʁɔt/, etc.

The glottals

Historically, the glottal stop mainly corresponds to the letter ང་ʔ, and the glottal fricatives to ཅ་ and ག་ respectively, with some exceptions.

For ex: ང་མ་ 'mother' /ʔama/ (La, Am), /ʔama/ (U, Kh, etc.), ང་: HORM 'Hor (proper noun)' /hor/ (U, Ts, Kh), ང་: HAGO 'to understand' /ха ‘k’o/ (U, Ts), / ‘h’o/ (Kh); ང་: be surprised' /hala/ (U), /hala/ (La); ང་: /hampa/ 'brutal' (U), /hamba/ 'courageous' (Ba), ང་: 'milk' /ɦoma/ (Ü, Kh) /ɦom/ (Sh), /ʔoma/ (Cone).

In some languages such as Sharkhok and Khöpokhok, /ɦ/ corresponds to H, as in ང་: HOPA 'Horwa (proper name)' /ɦo:wa/.

In some languages, mainly in Amdo and some Eastern languages, /h/ can correspond to PH, LH. In some rare cases (in Cone and Drugchu), bothʔ and ‘can correspond to /ʔ/.

For ex: ང་: PHUD 'to expel, to take off' /həl/ (Am), ང་: PHAN,THOGS 'to be useful' /hant’o/ (Am), ང་: LHAM 'shoe' /ham/ (U, Am), ང་: LHOD,LHOD 'relaxed' /’höhöʔ/ (Kh).

In Dzongkha, /h/ is also derived from CT resonant radicals with a preradical ང་.

For ex: ང་: SNA,PÅ 'nose' /haba/, usually spelled as ང་: LHA,PÅ (which is based on the sound and not the etymology), ང་: SNYING 'heart' /’hing/, ང་: SNGON,MO 'blue' /’hóm/.

41. See Suzuki (2014c). There is another report about the existence of uvular series originated from the same CT series in Chengzhang dialect of Kham (YE:SHES ’OD,GSAL 2008).
The resonants

The voiceless resonants (m', n', ny', ng', l', r') attested in multiple languages are generally derived from CT preradical S e.g. in Kham ཨན 'medicine' /m'/, ཨན 'nose' /n'a/, གོན བམ 'bean' /r'änma/, གེ སློ 'study' /h'lo/. The combination LH also corresponds to /l'/: མལ་ 'deity' /l'a/.

The vowels

In southern Kham (Gyälthang), /ɿ/ (ʅ) corresponds to several examples of a R final of CT, whereas /v̩, v̩/ corresponds to a simple vowel U and E respectively. For ex: གསེར 'gold' /h's/, མགོ 'head' /n'g/.

Historically, three secondary articulations (retroflex, velarized, pharyngealized) have a common origin, which is the CLT R except for those in the preinitial position. For ex: ར 'goat' > /ɾa/ (Kh, mThachu), སྤྲིན 'cloud' > /pə/ (Kh, mThachu), ལྷ 'deity' /l'a/.

Preinitial sounds

The preinitial sounds derive historically from the preradical letters G, D, B, M, 'R, L, S found in CT, which have often disappeared in many languages such as Ü, Tsang, Dzongkha, Sherpa, Jirel, Choča-ngača, etc. However, they are well maintained in Purik, Balti, Ladaks, Amdo, Kham, etc.

Prenasalizations mainly correspond to the reflexes of the preradical letters M and ' found in CT. Ex. གོང བམ 'arrow': /m'da/ (Am), /m'da/ (Kh, Tö), གེ སྲིན 'gold': /m'ts'o/ or /m'ts'o/ (Ü, Ts, Dz); བྲིན 'forget' (in several Amdo dialects, the labialised prenasal can be strongly pronounced in the manner of a consonantal cluster), བྲིན 'crawling insect and worm': /'bo/ (Am), /'bue/ (Kh), བྲིན 'human being': /'nye/ (Am), བྲིན 'fire' /'nye/ (Am), བྲིན 'high': /'thongbo/ (Am pastoralist dialect), /'tʰo/ (Khyungpo), /'tʰo/ (Kh).

Preaspirated sounds mainly correspond to the reflexes of the preradical letters G, D, B, R, L, S in CT. The existence of preaspirated series is well attested in Eastern
Tibet (Kham, Amdo, Sharkhok, Drugchu, etc.). The combination of a different voicing between the preaspiration and the main initial is attested especially in Eastern Tibetic languages such as Pashi, Khodpokhog, Drugchu, etc.

Ex. $^{5}$ RTA 'horse': /ʰtʰa/, $^{6}$ SNA 'nose': /ʰna/, $^{7}$ RDO 'stone': /ʰdo/.

One should note that both voiceless nasals and preaspirations are historically derived from the combination of s and a nasal ($N$, $NY$, $NG$, $M$) in CT. Both strategies are usually not encountered in the same language. Thus, for example, we find either $^{6}$ SNA 'nose': /ʰna/ or /n'a/

- Historically, the velar and uvular preinitial sounds correspond to the reflexes of the preradical letter G in CT, but they may also be derived from the letter D which was realized as /t/ at an early stage (see below 7.4.6).

Ex. གཅིག་ GCIG 'one': /tʃi/; གཞུ་ GZHU 'bow': /tʃo/, དཔྱིད་ DPYID (KHA) 'spring time' /ʃik'a/ (Am), /pit/ (Ba).

- From a diachronic point of view, the labial preinitial sounds have two different origins. They are either the reflex of the preradical B or the reflex of the radical $P, PH, B$ followed by a glide $Y$.

Ex. བརྗེད་ BRJED /vjel/ 'to forget' (Am), /jet/ (Ba); བྲག་ BRAG /pʈa/ 'rock' (Am), 'cliff' /pʃa:/ (Dz), བྱེ་མ་ BYE MA 'sand' /ʃema/ (Am), གླེ་ PHYLBA 'marmot': /ʃʰa/ (Am).

7.4.4. Pace of sound changes

It is quite difficult to have a precise idea of the pace of sound changes and the evolution for each language over the many centuries. However, some historical documents essentially written in Tibetan or Chinese provide very useful information about the pronunciation of Tibetic languages, particularly spoken in Central Tibet. Three texts are important to reconstruct the spoken forms of Ü and Tsang in the twelfth to thirteenth centuries, as was shown by Huang Bufan (1983):

$$YLGE\ BKLGS.THABS\ BYTS.PA$$

$$BDE.BLAG-TU\ 'JUG-PA'\ (KHON\ BSOD.NAMS\ RTSE.MO\ 1142–1182);$$

$$BYTS.PA\ BDE.BLAG-TU\ 'JUG-PA'$$
These texts, as well as other documents, clearly show that the Tibetan spoken in Central Tibet during the twelfth to thirteenth centuries (and in some cases even much earlier) had already undergone major changes:

1. the loss of preradicals (S, R, B, D, G) particularly in Ü region;
2. the predicals M and ' had already merged into a prenasalization (both Ü and Tsang);
3. the preradical L was pronounced as a prenasalization (by some speakers of Ü);
4. the preradicals (R, D, G) in Tsang region were probably realized as /r/; other documents (Chinese documents, chiefly Yuan dynasty annals) allow reconstructing the pronunciation of some consonant clusters in Tsang area;
5. the preradical S was still preserved as /s/;
6. the postradical R (at least for the combination GR) was still preserved as /r/ in Tsang.

However, the two last points are the subject of some controversy since there is no certainty about the methodology used by the Chinese for their transcriptions, particularly whether these transcriptions were based on spoken or reading pronunciation or were influenced by Tibetan orthography.

Finally, Huang (1983) and Qu (1996) hypothesized that the tones could have appeared as soon as the twelfth or thirteenth centuries. This hypothesis may be correct, but there is no evidence supporting it.

All the above features concern sound changes that probably took place in Ü and Tsang in the early medieval period. However, these changes did not occur in all Tibetic languages or may have occurred at a later period in some languages. For example, the clusters GR and BR have been maintained respectively as /kr/ and /pr/ until now in some Balti, Purik, Kyirong, Rongdrak or Choča-ngača dialects.
7.4.5. Types of sound changes and their geographic distribution

Concerning the realization of preradicals, it is possible to distinguish six types of changes in modern Tibetic languages:

The reflexes are manifested:

a) only as segmental features and prenasals: Amdo (Gyalrong and surrounding area);

b) only as segmental features without prenasal: Balti, Purik, Sham, Leh, Nubra, Western Zanhar;\(^{42}\)

c) as segmental features, preaspiration and prenasals: Amdo, Sharkhok;

d) as preaspiration and prenasals: Kham, Drugchu, Khöpokhok, Čone, Thewo;

e) as prenasals: Northern Kham, Hor, Tō, Spiti, Ü;\(^{43}\)

f) no trace of preradicals: Tsang, Dzongkha, Lhoke, Choča-ngača.

As we can see in the list above, the languages which exhibit segmental features are only found in some Amdo dialects and in the northwestern region of Ladakh and Baltistan. Preaspiration is only found in the eastern regions. Prenasalization is mainly found in the eastern regions of Amdo, Kham and Hor but also occurs in western Tibet.

7.4.6. Archaic phonological features in Tibetic languages and dialects

The qualification of a language as “conservative” or “archaic” is often an oversimplification for several reasons.

First, it is quite frequent that a language is preservative in some fields but innovative in others. For example, Balti and Ladaks certainly have a lot of archaic features in phonology but they have lost all the prenasals and their verb stem morphology, which is rather innovative.

\(^{42}\) The absence of prenasals is only true for initial syllables. Within a word, prenasals generate nasal reflexes.

\(^{43}\) In the Lhasa dialect, some speakers use a prenasalisation while others do not, but in other central dialects, the prenasalisation is clearly present. (See also: KAL.BZANG’GYUR.MED & KAL.BZANG DBYANGS.CAN 2002.)
Second, within a single “language” or “group of dialects,” some dialects can be quite “innovative” while others are “archaic.” This is, for example, the case with dialects in Amdo.

All the world languages undergo an evolution in the various linguistic fields (phonology, morphology, syntax, semantics, lexicon, pragmatics), but the pace of change differs from one language to another and even from one dialect to another and, as we have just said, the pace also depends on the various linguistic fields.

However, if we were to look only at the phonology of Tibetic languages, would it be possible to draw general conclusions about the degree of archaicty of the phonology in a given language or dialect? And if so, what would that tell us about the given language or dialect?

Traditionally, or theoretically, linguistic classifications are among other criteria made on the basis of shared innovations (see Chapter 9) and never on the basis of preservative features. However, archaicty in phonology may reflect a general socio-linguistic trend, which shows that isolate languages or languages spoken at the periphery of a linguistic area often preserve archaic phonological features, as has been shown in many works on dialectology and geolinguistics (or linguistic geography). Is that the case for Tibetic languages?

In order to clarify this question, let’s consider the main phonological reflexes listed in 7.4.1 and examine briefly their degree of archaicty.

- **Place and manner of articulation of the simple radical**

  They have generally been preserved in most languages. Only some languages have very specific phonological innovations for the labial sounds $P, PH, B$ (mainly in Amdo) and for the fricatives $SH, ZH$ (mainly in Kham). See 7.4.1.

- **Voicedness of initial simple plosives and fricatives**

  Voicedness of the initial simple plosives are *rarely* preserved in modern languages. The only exceptions are found among some dialects of Balti, Purik, Sham, Nubra, Garzha, Choça-ngaça, Pālki [Pashi], and Gyālrong surrounding Amdo.
Place and manner of articulation of the combination radical+wā

Reflexes of the postradical Wā, traditionally called wazur, are only found in some dialects of southern Kham and Choča-ngača.

Place and manner of articulation of the combinations KYA GYA

The preservation of prevelars for the combinations KYA and GYA as implied in CT forms is only found in the dialects of Central region as Ü, Tsang and Western regions such Ladakh and Baltistan. The other languages have an innovation as prepalatal affricates.

Place and manner of articulation of the combinations PHYA BYA

The preservation of labials with a glide for the combinations PHYA and BYA as implied in CT forms is only found (as least with the vowels A and U) in the dialects of Balti, Purik, Ngari, Tsamang Choča-ngača and Dränjong (Lhoke). For example, in these dialects, the words PHYUG 'rich' and BYA ‘bird’ are still pronounced as /bya/ and /p’yukpo/. The other languages have various innovations as presented in 7.4.1.

Place and manner of articulation of the combination MYA

The pronunciation /my/ is not reported in any of the modern languages. The more archaic pronunciation /"my/ is found in pastoralist Amdo dialects. It is for example attested in the words: DMYA-BA ‘hell’, MIY ‘person’, MYE ‘fire’.

Place and manner of articulation of the combinations KRA KHRA GRA

The preservation of velars with a glide for the combinations KRA, KHRA and GRA as implied in CT forms is only found in Balti, Purik, Gyalrong surrounding dialects of Amdo, and a few dialects from Rongbrag and Balung (Kham). For example, in these dialects, the segments /kr/ and /gr/ are still present in the words KHR ‘blood’ and GRU ‘knife’.

Place and manner of articulation of the combinations PHRA BRA

The preservation of labials with a glide for the combinations PHRA and BRA as implied in CT forms is only found in Balti, Purik, Kyirong, Choča-ngača, Gyalrong surrounding dialects of Amdo, and only two Kham dialect, sProsnang (belonging to
the Rongdrak group) and Phongpa (belonging to the Semkyi-Nyida group). For example, in the words གཉིས་ PHRU 'child', སྲག་ BRAG 'rock/cliff' or སྲིས་ BRIS 'write', the clusters /p’r/ and /pr/ are well preserved.

- **Place and manner of articulation of the combinations S ཙ R A**

  The preservation of /sr/ for the combination ཙ R A is preserved as implied in CT forms is only found in Balti, Purik and Sham. However, Balti and Purik also have an innovative form for this combination, such as /str/. For example, in the word རྫུང་ SRUNG 'protect', རློག་ SROG 'life' and སྲན་མ་ SRAN MA 'bean', the initial cluster is pronounced as /s(t)r/.

- **Segmental preinitials for non-nasal preradicals (G, D, B, R, L, S)**

  The preservation as a segmental preinitial for non-nasal preradicals (G, D, B, R, L, S) is only found in Balti, Purik, Sham, Nubra and some Amdo dialects and to a lesser extent in Zanhar (the preradical L). Note that the preradical D has not been preserved anywhere as a /t/-preinitial. In most cases, it is treated as the preradical R. For example, in the words གསུམ་ GSUM 'three', དགུ་ DGU 'nine', བཞི་ BZHI 'four', རྒད་པོ་ RGAD PO 'old man', ལྟ་ LTA 'to watch', སྒོ་ SGO 'door' the preinitial sounds (in bold) are clearly heard.

- **Prenasals for preradicals (’, m)**

  The preservation of a prenasal for preradicals (’, M) is found in Spiti, Tö, Hor, Kham, Amdo, Čone, Thewo-tö, Thewo-mä, Drugchu, Sharkhok, Khöpkhok and marginally in Ü: མཚོ་ MTSHO 'lake', མགོ་ MGO 'head', མདའ་ MDA 'arrow'.

- **Vowel quality in the open syllable (A, I, U, E, O)**

  Preservation of the vocalic quality as implied in CT forms (A, I, U, E, O) in an open syllable is only found in Ü, Tsang, Tö, Dzongkha, Choča-ngača, Spiti, Ladaks, Zanhar, Purik and Balti.

- **Segmental consonants for finals (G, NG, D, N, B, M, R, L, S)**

  Complete preservation of the segmental consonant finals as implied in CT forms (G, NG, D, N, B, M, R, L, S) is only found in Balti, Purik, Sham and Leh (Central Ladakh). The preservation of CT finals except for S is attested also in Choča-ngača, most of pastoralist dialects of Amdo, and some dialects from Khyungpo, Kham.
**Segmental consonants for final “second suffix” (s)**

Preservation of the segmental consonant finals as implied in CT forms \((GS, NGS, BS, MS)\) is found solely in Purik, Balti, Sham and Leh (Central Ladakh).

**Concluding remarks**

As shown above, Balti, Purik, Sham, Nubra, Leh (Central Ladakh), Amdo and Choča-ngača have preserved multiple archaic features. From the phonological point of view, they can be considered the most “conservative Tibetic languages.”

Purik and Sham (Ladakh) as well as Balti have the highest number of archaic features since they have preserved both preinitial (except of the nasals), postradical glides and final consonant clusters.

We can also note that these languages and dialects are located at the periphery of the Tibetic-linguistic area, i.e. the westernmost (Ladakh, Baltistan) and easternmost regions (Amdo); in addition, some languages, such as Choča-ngača, containing a certain level of archaicty, are spoken at the southernmost region of the Tibetic linguistic area.

Even if there isn’t a single language that has entirely preserved the phonology found in Classical Tibetan, the existence of the above mentioned “conservative languages” suffices to show that their phonology is directly derived from languages very closely related to Classical Tibetan.\(^44\)

---

\(^{44}\) In ergative constructions, the agent of a transitive verb is marked by a special case called “ergative” and in some cases “agentive.”
8. Grammatical outline of the Tibetic languages

This chapter aims to provide the essential grammatical features shared by the Tibetic languages. As we will see some of these characteristics are rather rare from a typological point of view in the world languages.

The fundamental morphological, syntactic, semantic, and lexical features of the Tibetic languages may be summarized in the following way:

(a) The verb normally occupies the final position in the sentence.\(^1\)
(b) Nominalizers as well as verb auxiliaries always come after the verb.
(c) Tense, aspect, modality and evidentiality (TAME) are marked by verb auxiliaries and/or suffixes.
(d) Modal verbs always come after the lexical verb.
(e) The negation marker is prefixed either to the lexical verb or to the auxiliary verb.
(f) Transitive verbs may trigger nominal ergative constructions, but ergativity differs in the degree of optionality and in its functions.
(g) Grammatical cases are marked by enclitics.
(h) Numerals and quantifiers are postponed to the noun phrase.
(i) There are only postpositions (no preposition).
(j) Light verb constructions are a major strategy among the lexical verbs.
(k) Lexical composition is a major morphological device.
(l) Ideophones constitute an important lexical category.

The position of adjectives and demonstratives depends on the given languages. They usually come after the noun but, in some languages, are placed before the noun.

Concerning grammatical and lexical semantics, we have the following characteristics:

---

\(^1\) There are some exceptions in the spoken languages, in the case of afterthoughts, antitopics or emphasis, particularly in informal registers. See Meunier et al. (1991).
(a) All the languages\(^2\) have developed rich evidential and epistemic systems.

(b) Intentionality is often grammaticalized and is marked by auxiliary verbs.

(c) There is a lexical distinction between controllable and non-controllable verbs.

(d) There is no grammatical gender.

(e) Number is a marginal category. It never co-occurs with numerals.

(f) Some languages have developed rich honorific registers for nouns, adjectives and verbs.

8.1. **Noun phrase**

The basic structure of noun phrases in most Tibetic languages is displayed as follows:

\[
\text{MOD} \text{NOUN} \text{MOD} (\text{QNT}/\text{NUM}-\text{DEM}/\text{DFM}-\text{COL}) \text{-case} (\text{-TOP}/\text{-ADM})
\]

About noun phrase structures, see also Garrett & Hill (2015).

The **head noun** may be preceded or followed by a modifier (MOD) which corresponds to an adjective or a relative clause. Then the noun may be followed by a sequence of optional suffixes or clitics\(^3\) which occur always in the same order: quantifier\(^4\) (QNT) or numeral (NUM), demonstrative (DEM) or a definiteness marker (DFM), followed by an optional collective marker (COL) which corresponds to a plural marker in European languages (see Kojima 2012). The next element is the **grammatical case** which usually signals the end of the noun phrase. All the cases are marked by an overt suffix except the absolutive case which remains unmarked.\(^5\) The

---

2. With the notable exception of Balti which has only a limited evidential-epistemic system.

3. “A clitic is a surface element part-way between a word and an affix in properties” (Dixon 2010: 221).

4. Quantifier corresponds to a type of determiner such as “some”, “each” in English.

5. It refers to the P (grammatical patient) function or S (single argument), but may also refer to periphrastic functions such as a locative. In some cases, the zero merely indicates the absence of grammatical case. Some spatial and temporal words such as ‘today’, ‘tomorrow’ are not marked by the locative case.
case may optionally be followed by a topic marker (TOP) or an adjunctive marker (ADM) meaning ‘also’ as in CT.

Here is an example from Ü and Common Tibetan:

(1) ཁྱི་ཆུང་ཆུང་དེ་ཚོ་ལ།
     KHYI CHUNG.CHUNG DE-TSHO-LA
dog small that-COL-DAT
‘(to) those small dogs.’ (Ü, ComTib)

The noun phrase begins with the head noun ‘dog’ (KHYI) which is followed by a modifier (an adjective), a demonstrative, a collective (or ‘plural’) marker and ends with a grammatical case (LA).

The modifiers may occur before or after the head noun depending on the type of modifier (adjective, relative clause, genitive noun phrase).

In most Tibetic languages, attributive adjectives are postponed to the head noun (see e.g. van Driem 1998; Zeisler 2004; Tournadre & Sangda Dorje 1998; Haller 2000, 2007; Häsel 1999, etc.). However, the reverse order is also attested in Balti and Purik (Bielmeier 2000; Zemp 2018, see also section 8.1.6.).

In the Tibetic languages, the relative clause usually precedes the head noun but there are also head-internal (e.g. in Ü and Tsang, Kh: Rongdrak), postnominal and headless relative clauses. Sometimes, the four types are attested within a single language (see Huber 2002).

In the Tibetic languages, the numerals are always postponed to the head noun. Thus one says:

(2) མི་གསུམ་ MI GSUM ‘three persons’ (Lit. ‘person three’)
     and not *གསུམ་མི་ GSUM MI

(3) སློབ་མ་ལྔ་ SLOB.MALNGA ‘five students’ (Lit. ‘student five’)
     and not *ལྔ་སློབ་མ LNGA SLOB MA

---

6. Ü designates the dialect of Central Tibet spoken in the capital Lhasa and in the neighboring towns and villages (see 9.6). It does not include Phänpo or Tsang dialects. Common Tibetan is the koine spoken in the TAR and in the diaspora. It is based on the Ü dialect. For details, see chapter 9.
Additionally, note the following general characteristics:

- The minimal structure of the noun phrase is made of a single noun or a pronoun, normally followed by a case (which may be zero-marked): NOUN/PRO-CASE.
- The quantifiers, numerals, definiteness markers, collective markers and grammatical cases always follow the head noun.
- Collective markers do not co-occur with numerals.
- The genitive noun always precedes the head noun.
- Classifiers are not used in the Tibetic languages. Some rare classifiers postponed to the noun occur with measurements, but they play a marginal role in the system.

8.1.1. Noun
The Tibetic nominal morphology essentially consists of four types of formation:

(a) Noun roots are usually monosyllabic.
(b) A derivational morphology characterized by nominal suffixes and in some rare cases nominal prefixes.
(c) An archaic derivational morphology characterized by initial and final formatives (see chapter 6).
(d) Compound nouns. This type of constructional morphology is a general device and found in many other language families of the world.

Noun suffixes
The most salient feature of the noun morphology is the presence of nominal suffixes (type b).

Some monosyllabic words such as "ME’ fire’, SHA’ meat’, SMAN’ medicine’, etc., and their correspondences in modern languages do not need a suffix but most monosyllabic stems are followed by a nominal suffix. The suffixes do not generally appear with compound word (In the examples below, bold underlines the suffix).
For example ལག་པ་ "LAG.PA'arm, hand' and རྗེ་པ་ "RKANG.PA 'leg, foot' form the compound word རྗེ་ལག་ "RKANG.LAG 'limbs' (Lit. 'leg-arm'); In Lhasa Tibetan and many dialects, རྗེ་ "WA-MO 'fox' and ལག་ "ZHWA.MO 'hat' form རྗེ་ལག་ "WAZHWA 'fox [fur] hat'.

In the various Tibetic languages, lexical nouns often share the same stem but differ in their affixes.

Depending on the languages, some monosyllabic nouns may appear with different suffixes or without suffixes altogether and in some rare cases with a prefix instead of a suffix.

The following examples illustrate the dialectal variation of suffixes:

རྗེ་ "RKANG.PA vs. རྗེ་ "RKANG.MA 'leg' (Balti) vs. འུལ་ "LUS.PO 'body' (Ko); རྗེ་ "ZHWA (Am, Kh) vs. རྗེ་ "ZHWA.MO (Ü, Ts) vs. རྗེ་ "ZHWA.YE (Ho) vs. རྗེ་ "ZHWA.GO (Kh, Sharkhok); རྗེ་ "GLANG.NGU (Am) vs. རྗེ་ "GLANG.GOG 'ox' (Ü, Ts).

The following examples illustrate the fact that some lexical roots may appear with a suffix or with a prefix or even a reduplicated stem depending on the language:

ཨ་སྲུ་ "A.SRU (Hor) vs. ཨ་ "SRU.MO (Ü, Ts) 'maternal aunt';

ཨ་ "A.ZHANG/ashang/ (Ü, Ts) vs. ཨ་ "ZHANG.PO/šangpo/ (Hor) vs. ཨ་ "ZHANG-ZHANG/šoxo/ 'maternal uncle' (Tm).

ཨ་ "A.NE (Ü, Ts) vs. ཨ་ "NE.NE (Ba) 'paternal aunt';

ཨ་ "A.MYES (Am) 'grandfather, forefather' vs. ཨ་ "MES.PO (Ü, Ts) 'ancestor'; ཨ་ "MES.MES.< "MES.MES.

The list of examples below illustrate words which appear with or without suffixes depending on the languages and dialects: རྗེ་ "ZAM.zam/ (Am) vs. རྗེ་ "ZAM.PA/šamba/ (Ü, Ts) 'bridge'; སྤྲིན་ "SPRIN/šwen/ 'fan' (Am) vs. སྤྲིན་ "SPRIN.PA/štön/ (Ü) 'cloud'; རྗེ་ "HAG/hag/ (Am) < CT རྗེ་ "PHAG.PA/šp’akpa/ (Ü, Ts) or རྗེ་ "PHAG.LU (Am) 'pig'; སྤྲིན་ "STON/šten/ (Sherpa) vs. སྤྲིན་ "STON.KHA/štön/a/ 'autumn, fall' (Ü), སྤྲིན་ "GCIN.PA/ščin/ (Ü, Ts), སྤྲིན་ "GCIN/ščin/ (Am, Kh) 'urine'.
Most of the suffixes found in the modern Tibetic languages are derived from the CT nominal suffixes: གོ་ PO, ལོ་ BO, བོ་ MO, བ་ PA, བ་ MA, བ་ KA (and its variants), ཆ་ CHA, ལོ་ SO, བུ་ BU and its variants པ་ GU, གྷ་ NGU, ད་ GU and ན་ U. (see chap 6).

However, there are a few exceptions such as the suffixes ཡེ་ YE, ལེ་ GO, ལེ་ GOG: See for examples བཞྭ་ཡེ་ ZHWA YE (Ho) and བཞྭ་ལེ་ ZHWA GO (Kh, Sharkhok) 'hat'.

The suffixes མོ་ MO and བ་ MA become a nasalized vowel, respectively པ/ and ཐ/. In Dzongkha, Sherpa and Lhoke, the suffixes of many words have lost the final vowel. For example འབོ་ PA, བོ་ MA are realized as འི་ PA and འི་ MA. The suffix of the word བོ་ བུ་ MO is written བུ་ MO and pronounced /b'um/. The suffix of the word པོ་ བུ་ MO is written བུ་ MO and pronounced /b'um/. The suffix of the word བུ་ བུ་ MO is written བུ་ MO and pronounced /b'um/. The suffix of the word བུ་ བུ་ MO is written བུ་ MO and pronounced /b'um/. The suffix of the word བུ་ བུ་ MO is written བུ་ MO and pronounced /b'um/.

Class terms

DeLancey (1986b) and Henderson (2006) define the class terms as follow: class terms occur as part of endocentric nominal compounds in which the class term is taken from a higher position in the taxonomy than the other element in the compound, which specifies the type. In English, for example, snake functions as a class term in compounds like rattlesnake, king snake, and grass snake where snake denotes the basic
category and rattle, king, and grass denote the type of snake.\(^7\)

Tibetic languages use a lot of class terms such as སྒྲ་ BYA 'bird', རྒྱུན། ME.TOG 'flower', རུ། BU 'bug, worm', ཞིང། SHING 'wood', etc.

The class terms are placed in front of the root or after the root depending on the words and the languages. The position of the class term also depends on the dialects.\(^8\)

Here are some examples: རྒྱ་ིན། PER.SHING 'oak', འབྲ་ BYA KRUNG 'crane', ཀྱིམ། BYA KHRA 'hawk', རྒྱ་ིན། BYA RGOD 'vulture', རྡོད། BYA GLAG 'eagle', རྒྱ་ིན། BU BLA.MA.MA.NI 'dragonfly', རྒྱ་ིན། BU DZING.DZING 'cicada', རྡོད། BU CHA, GAPA 'locust, grasshopper', རྡོད། BU GRÖG.MA 'ant', etc. In Zangskar, the class term BU is often after: རྒྱ་ིན། CHA, GA 'locust, grasshopper', རྡོད། BLA.MA.BU 'kind of insect', རྡོད། THAGS, REN.BU 'spider', རྡོད། GRE, MA.BU 'ant'.

In some languages, the class terms are compulsory and thus constitute a part of the word, whereas in other languages, they are optional and may be dropped: In some western and eastern regions, the words ཕྲ སྒྲ་ KHRA 'hawk', རྒྱུན། RLAG 'eagle', རྒྱུན། RGOD 'vulture', may occur alone, but in Central Tibet, the class term must be present: རྒྱུན། BYA KHRA 'hawk', རྒྱུན། BYA GLAG, རྒྱུན། BYA RGOD.

Compounding

Among the compounding strategies found in the Tibetic languages and in Classical Tibetan, one finds the polar compounds.

དྲོད་བསིལ། DROD.BSIL 'temperature' (Lit. 'hot-cool') (Dz), རྒྱུན། TSHA.GRANG 'temperature' (Lit. 'hot-cold') (Û), རྒྱུན། SROM.CHUNG 'size' (Lit. 'big-small') (Dz), རྒྱུན། CHE.CHUNG 'size' (Lit. 'big-small') (Û, Ts, Spi, etc.), རྒྱུན། RING.THUNG 'length' (Lit. 'long-short') (Û, Ts, Spi, La, Am, etc.).

---

\(^7\) For Henderson (2006), nominal classification is a broad cover term include class terms, measure terms, classifiers, and noun class markers (http://www.linguistics.ucsb.edu/sites/secure.itsit.ucsb.edu.sit/files/sitefiles/research/papers/17/Henderson_vol17.pdf)

\(^8\) There are also some variations inside a given dialect. In some words, the class term precedes the root whereas in other words, it follows it.
Echo words

A type of nominal reduplication often referred to as ‘echo-words’ is attested in some Tibetic languages. Echo-words also called *mühmele* are frequent in some language families such as Turkic, Semitic, Indo-Iranian and Mongolic. There are also marginally found in the Germanic (Yiddish, German, English) and Slavic languages. Here are some examples: *helter-skelter, pell-mell, money-shmoney, Techtelmechtel* (German) ‘fling’, *öksel-moksel* (Russian) ‘holy mackerel’, *liebe schmiehe* (Yiddish) ‘love and such things’. In languages such as Turkish, Persian or Hindi, they are highly productive: *ketap-metab* (Persian) ‘all kind of books’, *kelid-melid* ‘key, etc.’, *shâdi-vâdi* (Hindi) ‘wedding, etc.’, *prem-vrem* ‘love, etc.’. Echo words are well attested and productive in two Tibetic-speaking areas: The western region of Ladakh and Baltistan and the northeastern region of Amdo. In the first case, the Tibetic languages are in contact with Indo-Iranian languages and Turkic and in the second case, they are in contact with Turkic and Mongolic languages. As mentioned above, Turkic, Mongolic and Indo-Iranian have developed these morphological constructions and thus it is clear than the Tibetic languages have acquired this phenomenon through contact.

Here are some examples in Amdo: ◊ བྲཀར་ཡོལ་མར་ཡོལ་ *DKAR.YOL.MAR.YOL* ‘all kinds of cups’, ◊ གོན་རྒྱུ་མོན་རྒྱུ་ *GON.RGYU MON.RGYU* ‘all kinds of clothes’, ◊ བེ་རྟོད་རེ་རོད་ *DPE.RJOD ME.RJOD* ‘various examples’ (ex. from Simon 2016). In Ladaks, this construction is very productive: ◊ བྲག་ལྕང་བྲག་ལྕང་ *BA.RGLANG.WA.GLANG* ‘cows, etc., all sorts of cattle’, ◊ ཁྱག་ལྕང་ཁྱག་ལྕང་ *CALA.RGLANG.WA.LGALANG* ‘all sorts of things’, ◊ བེ་དྲེ་ཁྲི། *PE.NE.WE.NE* ‘some money’, ◊ བྲི་ིོ་བྲི་ *TLBI.WLB* ‘all sorts of hats’, ◊ བི་དྲི་ིོ་ *GHARI.WARI* ‘all sorts of vehicles’, ◊ བི་དྲི་ིོ་ *BUMOWO.BOMO* ‘all sorts of girls’.

8.1.2. Personal Pronoun

Three non-honorific pronouns are almost pandialectal: ◊ ག ག *NGA* ‘I’, ◊ དྲ *KHYOD* ‘you (sg)’, ◊ རེ་ *KHO* ‘s/he. In some languages (Am, Sh), the 3rd person pronoun may be expressed by a demonstrative pronoun or by a form derived from a demonstrative pronoun.⁹ For example, in a few dialects of Amdo, Drugchu, Sherpa, etc. the 3rd person

⁹ This strategy is frequently attested in the world languages. For ex. the French personal pronoun “il” is derived from the Latin distal demonstrative *ille*. 
Grammatical outline of the Tibetic languages

Pronouns are identical to the demonstrative pronouns འདི་’DI ‘this’, དེ་’DE that’ or the archaic form གན་’GAN ‘that’ (see Bacot 1948). For the 2nd person singular, the term རང་’RANG ‘self’ is used in some dialects, to mean ‘you’ (Ü, Ts, Lho, La).

Most Tibetic languages form the three ‘plural pronouns’: ‘we’, ‘you (pl)’, ‘they’ by adding a collective marker (see the section on number below), such as ཡི་TSHO, ཐོ་TSHANG, རུ་KUN, སུ་CAG, ཡི་RNAMS, འོ་CHO, འོ་THAMS, འེ་RIGS, etc. to the above mentioned pronouns (see 8.1.10 about number).

The category of dual is grammaticalized in a number of languages: རང་གཉིས་’NGA-GNYIS ‘the two of us, we’ (Ü), རང་སྟོན་’NGA-CHAM (Kh: Derge) ‘the two of us, we (exclusive dual)’ (see Häsl er 1999) or: རང་སྟོན་’NGA-CHAM ‘the three of them, they’. In some languages, a collective marker may be inserted between the pronoun and the number as in Kham (Derge): རང་སྟོན་གཉིས་’NGA-GNYIS-GNYIS ‘dual (inclusive), see Häsl er (1999).

In a few languages (Dz, Sp, Ga, YK, La, etc.), one finds plural pronouns that correspond to the honorific singular pronouns found in CT: བོད་KHONG ‘he/she (H)’, བོད་KHYED ‘you’ (H)’ or the archaic singular pronoun བོད་NGED ‘I’ (formal). Thus, in these southern and western languages, the forms corresponding to བོད་KHONG, བོད་KHYED or བོད་NGED mean respectively ‘they’, ‘you (pl)’ and ‘we’. It is likely that the honorific meanings found in CT are in fact derived from a plural meaning that has been preserved in some of the modern languages (just as the pluralis majestatis ‘majestic plural’ attested for example in the European languages). See also Hill (2010b).

The basic construction of the noun phrase with a pronoun is:

PRO-(NUM/COL)-case.

Along with the distinctions between the three personal pronouns (1st, 2nd, 3rd) singular and plural, most languages distinguish the following categories: gender (for 3rd person only), inclusive/exclusive (for 1st person non-singular only), honorific versus ordinary.
Gender distinction:

- In many languages, one encounters a gender distinction on the 3rd person. When this is the case (e.g. in Ü, Ts, Sh, Lho, Dz, BL (Kh), Ba, etc.), the root མོ MO is used, sometimes followed by a suffix. In some dialects (cf. Amdo), the 3rd person 'she' does not have a plural form. There is a distinction of gender in CT for the 1st person: ཞུས་པོ GUS.PO 'I (masc)', ཞུས་མོ GUS.MO 'I (fem)'. This distinction does not seem to be attested in the modern languages.

Inclusive/exclusive distinction:

- For 1st person non-singular, one frequently finds a distinction between inclusive and exclusive. The inclusive pronoun includes all the addresses while the exclusive pronoun excludes them. Most of the inclusive pronouns in the various languages are derived from the root ཝ ཁོ། O or མི་ OGAG / མི་ ཁོ། O.CAG, མི་ ཁོ། ASKL 'we' found in OT and CT, but one also finds the root རང་ RANG 'self': རང་ སྐོལ་ RANG.SKOL 'we'. They are found in Kham, Amdo, Thewo, Shkarhok, Zhikatse,10 Sherpa, Kyirong, Yolmo, etc. Ebihara (2014: 127) notes that:

  "In Amdo Tibetan, many dialects show the inclusive-exclusive distinction in the first-person plural and dual [...]. All the exclusive pronouns are derived from 1SG pronoun ཁོ་ (WT: ngo). Inclusive pronouns are those starting with vowels: ཁོ་, ཁོ་, ཁོ་, ་(ong, o, abko (the only exception is the Hongyuan)" [derived from ཁོ་ 'v'].

In Ladaks and Purik, the opposition is also attested between the exclusive ག ཕ སྐོལ NGAZHA (La) / ག ཕ སྐོལ NGACA (Pur) and the inclusive ག ཕ སྐོལ NGADANG /ngatang/. (See Koshal 1979 and Zemp 2018.)

Respectful forms:

- Many languages and dialects use distinct pronouns for the respectful forms of the three persons. For the 1st person, a humilific form is used in a few languages to show respect to the addressee: རྡི་ BDAG (Zhikatse) and ག ཕ་ B'A11 pronounced /mˈba/ or /ma/ (Spiti and Khunu.)

10. In Zhikatse, the difference between inclusive and exclusive is marked by a specific suffix (see Haller 2000).
11. The origin of this form is not clear. It could be derived from 'bA.ZHIG 'alone'.
Many languages have a special form for the 2nd person honorific, which is identical to CT Khoyed (often followed by a suffix). In a few languages, the 1st singular pronoun is used for 2nd person honorific. In Ladakhs, Balti and Spiti: ◇ Nyir-Trang ‘you’ (H) is clearly derived from the archaic form: ◇ Nyid-(Trang) ‘I’ which is maybe related to the root Nyid ‘self’ found in CT. In Sherpa, the form ◇ O-Trang (1st incl) is also used for ‘you’ (H). The Dzongkha honorific ◇ Na/na/ ‘you/he/she (H)’ and its plural form ◇ NaBu/naBu/ ‘they, you’ have unclear origins, but could well be derived from ◇ Rnam.Pa “honorific term used for second and third person” (see Valby 2003). The form ◇ Lhan.Rgyas < CT ‘together, common’ is frequently used in Lhoke; and in Common Tibetan especially for the second person plural high honorific: ◇ Lhan.Rgyas Rnam.Pa Tsho.

For the 3rd person honorific, many languages (Ü, Ts, La, etc.) use the root Khong sometimes followed by a suffix. Finally, it should be noted that some dialects (southern Kham as well as Drugchu, Baima, etc.) lack honorific pronouns entirely.

When people talk about a member of the family, in some regions (Ü, Tsang, southern Himalayas, etc.) they may use kinship terms to replace the personal pronouns, such as རྫུ་པུ ‘(maternal) uncle’, སྦ་མོ ‘(maternal) aunt’, སྦ་ཆེ ‘elder sister’, སྦ་ཇོ ‘elder brother’, etc. As noted by Huber (2002) about the Kyirong dialect, the kinship term is selected according to a set of rules depending on age, status and relationship of the person with the family of the speaker.

8.1.3. Interrogative proforms

The interrogative proforms include who, what, when, where, why, how, how much / many. The roots are normally derived from CT.

12. This type of usage is found in other languages, for example, the Kansai dialect of Japanese uses the word zibun (Lit. ‘self’) to mention ‘I’ as well as ‘you’.

The proform Who

One encounters only two forms in the various languages. The most widespread form is derived from CT སུ་ SU. It is present in the eight sections but pronounced in various ways: /su/, /sa/, /sʔu/, etc. However, in a few languages, the proform ‘who’ is derived from CT གང་ GANG ‘what, which’. This is the case in Dzongkha གཱ་ /ˈg’aː/ and Thewo-mi གི། /ˈkoː/ and some Southern Kham languages such as Derong /ˈkwoː/. An exceptional form /ˈsh’oː/ is used in Yunnan.

The proform What

There are two main roots found in the Tibetic languages both derived from CT ལོ་ CI (or its variant ལོ་ CHI) ‘what’, and གང་ GANG (or its variant ར་ GA) ‘which’, ‘what’, ‘who’. The first form ལོ་ CI is pronounced /či/ (Ba, La), /ˈči/ (Tö), /ˈči/ (Ho), /ˈčo/ (Ko) and /ˈčaː/, /ˈčaː tə/ (in Southern Kham). In Amdo this root is normally followed by the indefinite suffix /zək/; བོད་སྤྲུལ་ CHLZIG /ˈdzzək/ or བོད་སྤྲུལ་ CLZIG /ˈdzək/ (see Simon 2016). The other form གང་ GANG is realized as /kang/ (Ts, Sh), རེ་ GARE /ˈk’àːre/ (Û) and other forms in Southern Kham /ˈkə tə/, རེ་ GARE /ˈkə də/, རེ་ རེ་ GARE /ˈkə na/, རེ་ GARA /ˈkə la/, རེ་ རེ་ GABZO /ˈkə zo/. In some languages such as Dzongkha, the two roots are combined: རེ་ལོ་ GACI /ˈg’aːtə/). In CT, one additionally finds the form ལོ་ JI (which is obviously related to ལོ་ CI) in combination with other formant (see below).

The proform When

In most Tibetan areas, one finds a form, which is derived from CT ལོ་ NAM ‘when’ (Tö, Ho, Kh, Sp, La, Ba, Dz, Sh, etc.). This form is sometimes followed by another root such as ཁེ་ TSHAD or ཁེ་ TSHOD ‘measure’, e.g. in northern Kham /namtsä/, or Amdo /namts’ot/. In some areas, one finds compound words such as རེ་ལོ་ རེ་ལོ་ GADUS (Lit. ‘what time’), རེ་ རེ་ རེ་ རེ་ DUS, TSHOD GATSHAD also meaning ‘what time’: e.g. /ˈk’aːtə/ (Û), /ˈtöːsə’/ kaze/ (Kh), etc.
The proform *Where*

All the languages use the same root for the proform ‘where’. It is derived from the word བྱེད་*GANG* or its variant བྱེད་ CI ‘which’, ‘what’, ‘who’ followed by various suffixes: e.g. བྱེད་ CI *GANG.NA* /kangna/ (Am), བྱེད་ CI *GAPAR* /k’apar/ (Ü), བྱེད་ CI *GANA* /kana/ (Nkh, Lhoke), བྱེད་ CI *GAR* /kar/ (Sn, Dj), /gar/ (Ba), བྱེད་ CI *GANAS* /kane/ (La), བྱེད་ CI *GARU* /karu/ (Ts, Zhikatse), /karu/ (Drug, Čone, Th-m), བྱེད་ CI *GA* /ka/ (Thewo-mā), བྱེད་ CI *GARA* /k’ara/ (Ho), བྱེད་ CI *GALA* /k’ala/ (Sp), བྱེད་ CI *GATE* /’g’ate/ (Dz), བྱེད་ CI *GANI* /k’ani/ (Sh), བོད་སི་ *GANG,GSHIS* /kosh/ (Khopokhok), etc.

Some languages (such as Amdo, Lhoke, Zhikatse, Thewo-mā, etc.) distinguish the proform ‘where’ indicating a direction from the proform indicating a location. The difference is marked by the suffix: བྱེད་ CI /kangnga/ (Am), བྱེད་ CI /k’ala/ (Zhikatse), བྱེད་ CI /koná/ (Th-m), བྱེད་ CI /karu/ (La).

The proform ‘where from’ is also formed with the same basic root བྱེད་ CI *GA* ‘which’ followed by various suffixes depending on the language. In many regions, the suffix is derived from the elative case བོད་སི /NAS/ or more rarely from the ablative case བོད་སི /LAS/. Here are some examples of the various forms: བྱེད་ CI *GANAS* /k’aná/ (Ü), /kane/ (La), /’kana/ (Sp), བྱེད་ CI *GARNA* /karna/ (Ba), བྱེད་ CI *GALAS* /k’alá/ (Thewo-mā), བྱེད་ CI *GARS* /k’árá/ (Ho), བོད་སི /GARAGIS* /karagi/ (Čone), བོད་སི /GANG,GTI* /kanggo/ (Am), བོད་སི /GATE,LAS’ /’g’atélá/ (Dz), བོད་སི /GANIMA’ /k’aníma/ (Sh).

The proform *How*

The proform ‘how’ is a compound made of ‘what’ — བོད་སི CI (or བོད་སི CHI, བོད་སི JI) བྱེད་ CI *GANG* (or བོད་སི *GA*) — followed by various formants such as བོད་སི DRA’ similar’, བོད་སི TSUG or the equivalent of the verb ‘to do’ (in the given language). Here are some examples: བོད་སི *GADRAZE* /k’a’džás/ (Ü, Zhikatse), བོད་སི *GADRA* /’ka’dʒá/ (Bathang, Lithang), བོད་སི *GATSUG* /katsuk, kazuk/ (La), /’katsuk/ (Sp), བོད་སི *TSUG,GYA’A’* /’tsukkya/ (Sh), བོད་སི /GADE,BAD’ /’g’adebe/ (Dz), བོད་སི /GADE’ /’k’atá/ (Lho), བོད་སི’ *GII* /’g’atebe/ (Dz), བོད་སི’ *GII* /’g’atebe/ (Dz), བོད་སི’ *GII* /’g’atebe/ (Dz), བོད་སི’ *CHBYASE* /’čibyase/ (Ba), བོད་སི’ /’čibyase/ (Ba), བོད་སི’ /’čibyase/ all contain the verb ‘to do’ in the given language (kya, be, bya) respectively derived from CT བོད་སི /NID’ to do’ and བོད་སི /BID’ to do’.
The proform How much, how many

The proform 'how much' or 'how many' is sometimes made of single word, which is derived from CT words: ༠་ "how many", ག་ "a little, about, around" and བ ག་ "some". Here are some examples: ག་ "how many", བ ག་ "quantity" or བ ག་ "measure", བ ག་ "some". Many languages have a compound word that may be translated as 'what measure': བ ག་ /'k'atsä:/ (Ü), བ ག་ /'k'atsö:/ (Zhikatse), བ ག་ /'či/ (Ho), བ ག་ /'čitsä:/ (Nk). Finally, one also encounters forms that are identical or similar to 'how': བ ག་ /'imoz/ (Am).

The proform Why

The meaning corresponding to the proform 'why' is rendered in the Tibetic languages by various expressions (often corresponding to phrases) that literally mean: 'what has been done?', 'what has happened?', 'what caused (this)? 'for what reason/root?' or simply 'for what?'. For example, in several languages, the verb 'to do' is included in the expression, e.g. in Ü, Tsang and Dzongkha: བ ག་ /'k'are/ (Ü), བ ག་ /'k'ang/ (Zhi), བ ག་ /'či/ (Ho), བ ག་ /'čitsä:/ (Nk). Finally, one also encounters forms that are identical or similar to 'how': བ ག་ /'imoz/ (Am).

In Amdo, two forms are encountered for 'how much', 'how many': བ ག་ /'tə/ and བ ག་ /'čimoz/. It seems, the former is more frequently used by pastoralists, whereas the latter is used by cultivators (Camille Simon, pers. comm.).
Most Tibetic languages make a distinction between proximal and distal demonstratives. Some languages have a threefold distinction: proximal, medial and distal. In some languages, one finds also an obviative pronoun.\textsuperscript{16}

Historically, Tibetic demonstratives are derived from five demonstrative markers found in CT: རྡོ་ ‘proimal’ (and variants that sometimes appear in the modern written forms such as དེ་ ‘that’ or དེ་ ‘that (medial)’ (used in Am and Pa), དེ་ ‘that (medial)’ or ‘that (obviative)’, དེ་ ‘archaic determiner’ (see Huber 2005 and Bielmeier 2000), དེ་ ‘that over there (distal)’ (the variant དེ་ ‘HA occurs in Amdo). Some of these morphemes may combine together.

The three-fold distinction between proximal, medial and distal is well preserved in some languages. This is for example the case in Common Tibetan: རྡོ་ ‘this’ དེ་ ‘that’, དེ་ ‘PHAGI that (over there)’. Additionally this language also has a distinction between དེ་ ‘YAGI that (up there), དེ་ ‘MAGI that (down there). Dzongkha has even a four-fold distinction: རྡོ་ ‘this (very proximal)’, རྡོ་ ‘that (proimal)’, དེ་ ‘DE that (neutral), དེ་ ‘that over yonder’ (distal)’ (van Driem 1992). Some languages have not maintained the distinction between medial and distal (e.g. Kham or Spiti).

As in many world languages, the demonstratives have two main functions: they may serve as pronouns or as noun determiners.

For the demonstrative determiners, we find the following syntactic patterns across the Tibetic languages. Note that the representation scheme of the noun phrase is abbreviated below (for the complete version see section 8.1):

(a) \textbf{NOUN-DEM-(COL) CASE}

(b) \textbf{DEM-NOUN-(DEF)-(COL) CASE}

\textsuperscript{16} The obviative designates a third person that is less salient compared to another third person that is more central to the story or is close by (proximal).
Note that in (b), the demonstrative may co-occur with a definite marker. The combination of a demonstrative and an article is found in some other languages, that is for example the case in Hebrew but the word order is the reverse of the Tibetic order: Definite article-N-DEM.

In the great majority of Tibetic languages within Tibet (Ü, Tsang, Tö Ngari, northern Kham and Amdo), the demonstrative determiners occur after the head noun as in (a).

Conversely, in most Tibetic languages spoken outside Tibet, in the southern and western Himalayas (such as Balti, Ladakhi, Spiti, Sherpa, Yolmo, Dzongkha, Chochangca, etc.) as well as a few languages of Tibet such as Kyirong and some southern Kham dialects (Sn(Kh), Mi(Kh), YK), etc. the demonstrative determiners are placed before the head noun as in (b). In this case, they often appear together with a definite article that follows the noun.

The following example in Yolmo (Gawne 2013) illustrates the anteposition of the demonstrative:

(3')  ཀྲུང་ཁ ལེགས་ ཀཱ་པོ་ འོད། (འོད་)
/‘o.DI MI YAG.PU YED
that person+ABS good CPV
‘That person is nice.’

See also in Dzongkha:

(4)  ཀྲུང་ཁ ལེགས་ ཀཱ་པོ་ འོད། (འོད་)
/‘o.PHI MI DGE.SLONg HIN
/‘o’i mi gelong ing/
that person ordained monk CPV
‘That man is a monk.’

In Common Tibetan, the demonstrative is always placed after the head noun:

(5)  ཀྲུང་ཁ ལེགས་ ཀཱ་པོ་ འོད།
/MI ‘DH YAG.PO YED
person+ABS this good CPV
‘This person is nice.’ (Ü, ComTib)
In a couple of languages spoken in southern Kham, the demonstrative may occur either before or after the noun depending on syntactic criteria. If the noun is followed by a modifier, the demonstrative is postponed but otherwise it occurs before the head noun (see Suzuki 2011).

8.1.5. Definiteness markers

Tibetic languages do not have real articles since the grammatical category of definiteness is not obligatory in these languages. The fact that “Tibetan has no articles per se” has been noted by some authors such as Goldstein (1991: 48). However, in some previous works, definiteness markers have received various labels such as “article” (Kesang Gyurmé 1992; van Driem 1998; Tournadre & Sangda Dorje 1998; Haller 2000; Huber 2002; Gawne 2016;17 Graves 2007; Robin & Simon forthcoming), “specifier” (Koshal 1982; Beyer 1992), “determiner” and “selector” (Beyer 1992), “numerator” (Denwood 1999). In general, most authors have noticed that the definiteness markers are not obligatory and their frequency is much lower than the articles in European languages.

In order to clearly distinguish these grammatical categories from the category of the “article”, the terms definite marker and indefinite marker (or the variants definiteness marker, indefiniteness marker) are more suitable. We will use them in the book following several authors such as Häslar (1999) and Zeisler (2007).

Indefinite marker

Modern languages have an indefinite marker, which is either derived from the CT indefinite marker ་ཞིག་ZHIG or its allomorph ་ཅིག་CIG (note that the literary article has three variants depending on the last letter of the word: ་ཞིག་ZHIG, ་ཅིག་CIG and ་ཤིག་SHIG). In CT, this marker is cognate with the numeral ་གཅིག‘one’ and is clearly derived from it. Some modern languages such as Amdo maintain a distinction between the numeral ་གཅིག‘GCIG’ and the indefinite marker ་ཟིག་ZIG (reflex of CT ་ཞིག་ZHIG).

17. Gawne (2016: 98) notes in her Sketch grammar of Lamjung Yolmo that “there are no words in Lamjung Yolmo that are specifically used as articles, however there are words that do function as articles. They are not always used and in/definiteness cannot be inferred from their absence.”
However, in most modern languages, the indefinite marker is formally identical to the numeral ‘one’.

The indefinite marker is pronounced in various ways in the modern languages: རྗེ་ CIG /čik/ in Balti and Purik (sometimes reduced to -čik/), /či/ in Ü, Tsang; /či/ in Dzongkha and Kham; ZIG /zak/ in Amdo; or reduced to /yι/ as in Sherpa.

(6) གྱི་ རྗེ་/ རྗེ་/ CIG

\(^{`} \text{MI} \quad \text{PHYUG.PU YI}

\(^{`} \text{mi} \quad \text{č} \text{ukpu-ι/}

\text{man rich-IND}

‘A rich person.’ (Sh; adapted from Graves 2007: 191)

(7) གཟིག་ རྗེ་ ZIG

\(^{`} \text{NGAS} \quad \text{KHYIM-CIG} \quad \text{BZOS-ZIN}

\(^{`} \text{ŋä} \quad \text{č} \text{ung-či} \quad \text{zu-a-zin/}

\text{1SG+ERG house-IND make-AUX}

‘I built a house.’ (Gyalthang Kham; adapted from Hongladarom 2007b: 128)

**Definite marker**

The majority of modern languages, just as Classical Tibetan, do not have a definite marker. This category is also absent in CT. In order to convey definiteness, a bare noun or a noun with a demonstrative is used.

However, some languages with demonstratives preceding the noun have developed definite markers, which are always postponed to the noun. The definite marker may be used alone or co-occur with the demonstrative or possessive nouns. This corresponds to the structure: (DEM/ POSS) NOUN-DEF.

Such structures, which combine a demonstrative or possessive and a definite article (or marker), are also attested in other languages such as Hebrew, Italian, Romanian, Scandinavian languages and Greek.\(^\text{18}\)

---

18. Cf. e.g.: for example the phrase: ‘my friend’ translates as: *Il mio amico* (Italian), *prietenul meu* (Romanian), *ילש רבחה* ha xaver sheli (Hebrew).
Tibetic languages which have developed definite markers are mostly found in southern and western regions. They include Kyirong, Yolmo, Sherpa, Balti, Purik, Ladaks, Spiti, etc.

The definite marker is usually derived from CT demonstratives: usually the medial demonstrative དེ་ *DE* (e.g. Sherpa and Kyirong) or the proximal ཤ་ *DI* sometimes.

In Ladaks, Purik and Balti, the definite marker /po/ *པོ* (and its variant /bo/ *བོ*) may be derived from ཤ་ *O*, an archaic determiner. In Amdo, a morpheme used in the relative clause structure /o/ may also be related to this archaic determiner. In Tsang, the definite marker is derived from the 3rd person pronoun མ་ *KHO* (Haller 2000).

Here are some examples of definite markers in Ladaks:

(8) སྣ་པོ།  *work-DEF*

*LAS-PO*  
/las-po/  
The work. *(La)*

(9) མ་ཚོགས་སི་གནས་སྟངས་པོ།  *situation-DEF*

*SPYI-TSHOGS-SI*  
/netangs-po/  
The situation of the society. *(Ladags Melong, 2002, vol 1, issue 5)*

In the following examples in Sherpa (see Graves 2007; Tournadre et al. 2009) and in Kyirong (Huber 2002) the demonstrative and the definite marker coexist:

(10) ཤ་  *DEM*

*DI*  
/p’edza-ti/  
That child.*
8.1.6. Quantifiers and Numerals

In the Tibetic languages, quantifiers and numerals are always postponed to the noun. The quantifiers are never used with the numerals.

A few quantifiers are frequently found across the Tibetic area. They include: བལ་ལ་ /lala/ 'some' (Ts, Am) or བལ་ལུ་ /lu/ 'some' (Dz), ཀྲ་ཧི་ས /khas/ 'some' (Ü, Ts, Am), ཆུ་མོ /chum-mo/ 'each' (pandialectal), etc. For detail, see the HCTL.

There are two basic numeral systems found across the area: a decimal and a vigesimal system. The decimal system is pervasive. Numerals are fundamentally composed of the words from 1 to 10 and 10\(x\). The cardinal numbers from 1 to 10 are derived from the same roots in all the Tibetic languages:

- གཅིག /gcig/ 'one'
- སྤྱིར /gnyis/ 'two'
- གསུམ /gsum/ 'three'
- ཚུ་ /bzhi/ 'four'
- ལྣ /lnga/ 'five'
- ཉི /drug/ 'six'
- བདུན /bdun/ 'seven'
- རྒྱ བ /brgyad/ 'eight'
- དགུ /dgu/ 'nine'
- རྒྱ /bcu/ 'ten'

In the decimal system, the rounded numbers can be followed by the word ཐམ་པ་ /thampa/ 'multiple of ten.' The unrounded numbers normally need a connecting element to express the units, for example, དྲུ་ /rtsa/ for 20 + x (x = 1 to 9), ཉི་ཤུ /nyi-shu/ '21.' In Balti, the connecting element is ཉི /nyi/ in Kham. For example, ཉིས་བརྒྱ་ ཉི /nyis-brgya-nyi/ '220' (Ts).

Other connecting elements used after རྒྱ /brgya/ 'hundred', ཡི་ /bya/ 'thousand', etc. include བ མ་ /ra/ in Amdo, བ ཡ ལ ལ དང /dang/ in Ü, ཁ ཆ ཐ ཕ /ni/ in Kham. For example, in Amo:

- ཉིས་བརྒྱ་ ཉི /nyis-brgya-nyi/ '220' (Am), ཉིས་བཞི /nyis-brgya-dang/ '220' (Ü, Ts).

The element དྲུ /rtsa/ can be used for all the unrounded numbers from 21 to 99, but in many languages, the connecting element exhibits different forms depending on the number of tens digit. For example in Ü and Dzongkha we find:
In these languages, the connecting element is clearly derived from the tens digit:


In a few dialects such as Khöpokhok and Zhollam (Sn(Kh)), this connecting element is not needed.

There are specific words for multiple of tens for hundred and higher figures: BRGYA 'hundred', STONG 'thousand', KHRI 'ten thousand', BUM 'hundred thousand', SAYA 'million', BYE 'ten million', DUNG.PHYUR 'hundred millions', THER.BUM 'billion'. These words are usually pandialectal however, after BUM, the higher figures are rarely known or used.

The vigesimal system is found in many languages particularly at the periphery of the linguistic area in the southern and western languages (Dzongkha, Lhoke, Chočanga, Sherpa, Yolmo, Balti, etc.) and in the east (Thewo-mä). It is probable that the vigesimal system corresponds to a more ancient tradition. Although it is not attested in CT, it is also found in some East Bodish languages such as Bumthang, Kheng, Dzala, 'Ole. It is also found in Lepcha (van Driem 2001: 818).

19. In Lhasa, NYER is replaced by RTSA.

20. For Yolmo (see Hari and Lama 2004), for Sherpa see Tournadre et al. (2009), for Chočanga, see Tournadre & Karma Rigzin 2015), for Balti (Tournadre's unpublished data and field recording in Skardo, Shigar and Khapulu).
The vigesimal system usually coexists with the decimal system. Sometimes, the former is reserved for the expression of the age and for measuring grain, etc.

The vigesimal system uses the word མཁལ་ KHAL which originally means 'load' (on a yak) and has become a weight measure of 'one score' equivalent to twenty units.

ཁལ་གཅིག་ KHAL GCIG 'twenty', མཁལ་གཉིས་ KHAL GNYIS 'forty', མཁལ་གསུམ་ KHAL GSUM 'sixty', མཁལ་བཞི་ KHAL BZHI 'eighty', མཁལ་ལྔ་ KHAL LNGA 'one hundred'. It is interesting to note that the word མཁལ་ KHAL is used in nearly all the regions where the vigesimal system is still found. However, in Baltistan, the word མཁལ་ KHAL is replaced by ཉི་ཤུ་ NYI SHU 'twenty'. Thus, in Balti, we have instead ཉི་ཤུ་ KHAL GNYIS 'forty' (Lit. 'two twenty'), ཉི་ཤུ་ KHAL GSUM 'sixty' (Lit. 'three twenty'). Every region using the vigesimal system has a different way of counting the figures 'thirty', 'fifty', 'seventy', etc. Dzongkha for example uses the word ཨོ་ཐེད་ PHYED 'half' in combination with མཁལ་ KHAL (see van Driem 1998). For example, མཁལ་ཐེད་དང་གཉིས་ KHAL PHYED-DANG GNYIS 'thirty' (Lit. 'twoscore (minus) halfscore' or 'forty (minus) half twenty'), མཁལ་ཐེད་དང་གསུམ་ KHAL PHYED-DANG GSUM 'fifty' (Lit. 'threescore (minus) halfscore', or sixty (minus) half twenty'), etc. Thewo-mā uses the word རོ་ཞིག་ཐམ་པ་ THAM PA 'ten' (Lit. multiple of ten). For example, བཀྲ་ཚོགས་ཐམ་པ་ KHAL GNYIS THAM PA 'fifty' (Lit. 'two score ten'), etc.

Finally, to count the number of years, some dialects such as Purik use the word བཀྲུས་ SKOR which refers to the 'twelve year animal cycle' of the Tibetan calendar. Thus བཀྲུས་ནི་ཤྱེར་ GNYIS.SKOR means '24 years', བཀྲུས་ནི་ཤྱེར་ GSUM.SKOR '36 years', etc.

Classifiers, except for the measurement terms, are rarely used in the Tibetic languages and their grammatical status is at most marginal. They always follow the head noun in the following way: N-CL-NUM. Here are some examples of classifiers used for measurement: རྒྱང་ RKANG 'hair', བྲེ་ BRE 'a measure unit' (for cereal or liquid roughly equivalent to one liter), སྦོར་ MGO 'head' (Am; see Simon 2016).

Ex. Ü, Ts (and CT): བཀྲ་ཚོགས་ SKRA RKANG GCIG 'one strand of hair', རྒྱང་དི་། བཀྲ་ཚོགས་ ME.TO G S K RA RKANG GSUM 'three flowers', བཀྲ་ཚོགས་ NAS BRE GSUM 'three dre
of barley', ◊ མི་མགོ་གཅིག MI MO GCIG 'one person' (Kh), ◊ མྱི་མགོ་གནོད་ MYI MO GNIS 'two persons' (Am). See Simon (2016).

In the dialect of Rongdrak (Kh), ◊ ད་ DA, a special classifier not found in CT is used:
◊ མི་ད་གཅིག MI DA GCIG /da^h/ 'one person' (Kh).

In the dialect of Lhagang (Kh), ◊ རྡོག་ RDOG, a special classifier is used only for 'one':
◊ མི་རྡོག་གཅིག MI RDOG GCIG /^hdo^h/: 'one person' (Kh).

A couple of specific numerals, such as ◊ གང་ GANG 'one' or ◊ དོ་ DO 'two' are specific to count massive nouns, usually liquids or grain: for example ◊ ཆུ་ཕོར་པ་གང CHU PHOR PA GANG 'one bowl of water' (Lit. '(A) full bowl of water'), ◊ ཆང་དཀར་ཡོལ་དོ CHANG DKA R YOL DO 'two cups of chang'.

In some dialects of Thewo-Tö, one finds the numeral ◊ འཁན་ KHAN /^nk'ã/ whose origin is unclear: ex. ◊ མྱི་འཁན་ MYI KHAN 'one person'.

The ordinal numbers are very similar in most languages and dialects. They are made by the adjunction of the suffix PA also found in CT:
◊ གཉིས་པ་ GNYIS PA 'second', ◊ གསུམ་པ་ GSUM PA 'third', ◊ བཞི་པ་ BZHI PA 'fourth', ◊ དྲུག་པ་ DRUG PA 'fifth', ◊ བདུན་པ་ BDUN PA 'sixth', ◊ བཅུ་པ་ BCU PA 'seventh', ◊ བདུན་པ་ BDUN PA 'eighth', ◊ འབྱུང་་་ DGU PA 'ninth', ◊ བཅུ་པ་ BCU PA 'tenth', etc.

The word 'first' has a special form usually derived from CT ◊ དང་པོ DANG PO 'first'.

The ordinals in some languages are usually preceded by the word འདོ་ PAN 'number'.

In southern Kham, some dialects have yet another way to express ordinals.

8.1.7. Adjective

The adjectives in the Tibetic languages, as many world languages, have two main grammatical functions: they function as a modifier of a head noun and as an adjectival predicate. Here is a description of their morphology.

The Tibetic adjectival morphology consists of four main types of formation:
(a) The stem is followed by a grammatical suffix.
(b) The stem is reduplicated.
(c) The stem (or compound adjective) is followed by a reduplicated syllable.
(d) The compound adjective is made of a noun and an adjective.

Additionally, nominalized verb phrases are used in CT to translate Sanskrit adjectives.

One should note that in most cases corresponding to (a), (b), and (c), the adjectival stem is of verbal origin.

**Stem followed by a grammatical suffix**

The adjectival morphology corresponding to the type (a) is characterized by a set of suffixes. The number of suffixes varies according to the languages. As in many world languages the Tibetan languages distinguish three degrees for the adjective: positive, comparative and superlative. Some Tibetic languages additionally have a series of specific grammatical suffixes (see below).

For the positive forms, the main suffixes are derived from the CT adjective suffixes: བོ PO, བོ BO, མོ MO, བ་ PA, བ་ BA, བ་ MA. One also encounters other suffixes such as ཀན CAN or ངན DRAGS also attested in CT but with slightly different meanings.

Some southern dialects such as Dzongkha, suffixes may be subject to phonological reductions such as: PO/PA → /p/, BO/BA → /w/, MA/MA → /m/. The reduction of the suffix has also an impact on the stem which is also sometimes reduced: e.g. རིང་པོ RING.MO /'rimp/ 'long' (Dz) < རིང་པོ RING.MO, གཡུག PB /'pbyug/ 'rich' (Dz) < གཡུག PB PHYUG.PO. In some Tö and Tsang dialects, the suffix PA may be reduced to /a/: སྐྱིད་དা SKYID.DA /'skyita/ (see other ex. below), གར་ར་ GAR,RA /'k'ara/'strong (alcohol)/.

Additionally, two other suffixes which are not found in CT are attested in various regions and they probably have a Proto-Tibetic origin: ཇེ་ DE /'de/ or /'te/ (or even /'tc/), southern and western regions (Ts, Tö, YK, Sh, Sp, Ga, La), as well as in a few north-
eastern dialects (Am, Tm, etc.). Its variant गे GE e.g. in Tsang, Tö and Sherpa. The morpheme गे GE is used after velar while ्◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌◌สาธารט
'numerous', a lot: མང་པོ MANG.PO / (Ü), མང་ལོ MANG.LO (Ts), མང་ལོ MANG.BA/mangwa/ (Am).

'long': རིང་པོ RING.PO (Ü), རིང་ལོ RING.LO (Hor), རིངེ RINGMO/'rim/ (Dz).

'strong' (tea, alcohol, etc.): བར་པོ GAR.PO /'k'a:po/ (Ü), བར་ལོ GAR.LO /'k'ara/ (Ts, Lhatse).

'cold': རྒྱུན་མོ GRANG.MO /grangmo/ (Pur), /graχmo/ (Ba), /ʈ'angmo/ (Ü), རྒྱུན་ལོ GRANG.LO /ʈ'angge/ (Ts, Sh).

'excellent': བཟང་པོ BZANG.PO, བཟང་ལོ BZANG.LO (Ü), བཟང་གེ BZANG.GE (Ts).

'strong' (work): དྱོང་པོ GYONG.PO /'ky'o:ngo/ (Ü), དྱོང་ལོ GYONG.LO /'ky'o:ngge/ (Ts).

Reduplication of the stem

The reduplication of the stem is a common way of forming some adjectives, for example: ཆུང་ཆུང་ CHUNG~CHUNG 'small' (Ü, Ts), ཐུང་ཐུང་ THUNG~THUNG 'short' (Ü, Ts), རིལ་རིལ་ RIL~RIL 'spheric' (Ü, Ts, etc.), སྒོར་སྒོར་ SGOR~SGOR 'round' (Ü, Ts, Kh), ཀྱིར་ཀྱིར་ KYIR~KYIR 'round' (La), ལོད་ལོད་ LHOD~LHOD 'relaxed' (Ü, Ts, Kh, etc.), དམའ་དམའ་ DMA~DMA 'low' (Kh), མྱུལ་ཕྱུལ་ PHRA~PHRA 'thin' (Kh), སྐྱོད་སྐྱོད་ SKYA~SKYA 'grey' (see also Bartee 2007).

In many dialects of Kham, reduplication is used to form color adjectives: སྦྱར་སྦྱར་ DKAR~DKAR 'white', ཤུག་ཏུག་ NAG~NAG 'black', དམར་དམར་ DMAR~DMAR 'red', སེར་སེར་ SER~SER 'yellow', གཉུན་ SNGO~SNGO 'blue', ཕྱུར་ SKYA~SKYA 'grey' (see also Bartee 2007).

Stems followed by a reduplicated syllable

Stems followed by a reduplicated syllable, are a common way of forming adjective in some Tibetic languages such as Ü, Tsang, Kham, Dzongkha or Lhoke. In many

22. 'This adjective is essentially used with an adverbial function. See the section on "Adjectival and adverbial expressions" below.
dialects, these constructions usually convey an expressive meaning and function as ideophones (see below, 8.1.8).

The following examples are from Central Tibetan:

སྔོ་ཐིང་ཐིང་ SNGO THING~THING 'bright blue/green', དམར་ལྷབ་ DMAR LHAB 'blazing, burning red', དམར་ཐིང་ DMAR THING 'bright red (blood)', ལྷིང་ཐིང་ LHING THING 'bright blue/green', ཨིང་ཐིང་ཐིང་ 'bright red (blood)', རྒྱལ་ཐིང་ཐིང་ GRANG SIL~SIL 'biting cold', ངོ་ཐིང་ཐིང་ TSHASOB~SÖB 'warm as toast (pleasant heat)', ཤིང་ཐིང་ JAM THING~THING 'quite calm', ལྷིང་ཧྲིལ་ HRIL~HRIL 'clean as a whistle'.

In Kham, SKAL.BZANG 'GYUR.MED & SKAL.BZANG DBYANG.CAN (2002) have shown that these constructions are extremely productive and convey a lot of semantic nuances. For example, they listed the following expressions just for the various types of ཡོན་ NAG 'black' in Bathang (Kh) but did not provide the translations that could render the various uses and subtle nuances of these ideophones. Here is the list (with some translations that we added):

ནག་ཏིང་ཏིང་ NAG TING~TING 'deep black',ནག་ཁབ་ཁབ་ NAG KHAB~KHAB 'blue, turquoise, etc. (or other color) with shades of black',ནག་ཐོ་ཐོ་ NAG THO~THO 'at night, unclear forms',ནག་དྲིང་དྲིང་ NAG LDING~LDING 'deep water with dark or black color',ནག་རྲི་རྲི་ NAG RBAB~RBAB 'dark, not very clear, at dusk',ནག་མུར་མུར་ NAG MUR~MUR 'a group of people or animal hardly visible in the dark',ནག་པུ་པུ་ NAG SUB~SUB 'dark, not very clear, at dusk',ནག་ལྕྱི་ NAG CHIS~CHIS 'a lot of animals (such as yaks) seen from a distance making a dark shape',ནག་དྲེག་དྲེག་ NAG GO~GO,ནག་དུན་དུན་ NAG 'LA~LA',ནག་པུ་ NAG RDOG~RDOG. One can add ཡོན་ནུང་ནུང་ NAG SLONG~SLONG 'black clouds as a sign of hail'.

In Amdo, we find similar constructions: གནོ་དྭངས་དྭངས་ SNGO DWANGS~DWANGS 'entirely blue (for the sky)', གོ་ལྡེམ་ལྡེམ་ SNGO LDEM~LDEM 'entirely green (for fields)'.

For a detailed description of the same constructions in Lhoke, see Yliniemi (2019).

In Dzongkha and Choça-ngaça, these constructions no longer convey an expressive meaning and are used to form many basic adjectives:
Compound adjective made of noun and an adjective

Adjective compounds are frequently found in Tibetic languages. They are usually made of a noun followed by an adjective. Frequent compositions include in Central Tibetan:

- གསལ་ GSAL TOG~TO 'clear', དྲང་ DRANG TOG~TO 'straight, honest', སྐྱིད་ SKYID TOG~TO 'happy, pleasant', འཇམ་ JAM TOG~TO 'easy', བྱ།་ CHI TOG~TO 'nice', etc.

The stems are respectively derived from CT forms:

- གསལ་ GSAL 'clear', དྲང་ DRANG 'straight, honest', སྐྱིད་ SKYID 'happy, pleasant', etc.

In other modern languages, these constructions are also attested e.g. in Dzongkha and Kham:

- སྤྱིན་པོ་ GYON PO 'important, significant', གླེ་མཚར་ གླེ་མཚར་ RGYA RGYA 'difficult', གོང་ཆེན་ གོང་ཆེན་ GONG CHEN 'expensive', རྒྱུ་མ་ རྒྱུ་མ་ RGYU RGYU 'lymphatic, slow', etc.
Nominalized verb phrases used in CT to translate Sanskrit adjectives

These nominalized verb phrases are used as lexicalized adjectives:

བསམ་ཡུལ་ལས་འདས་པ་ **BSAM.YUL.LAS.DAS.PA** ‘inconceivable unbelievable’;

འཇིགས་སུ་རུང་བ་ **JIGS.SU.RUNG.BA** ‘frightening, scary’;

ཡིད་དུ་འོང་པ་ **YID.DU.ONG.BA** ‘delightful, adorable’;

བསམ་གྱིས་མི་འཁྱབ་པ་ **BSAM.GYIS.MI.KHYAB.PA** ‘inconceivable, unfathomable’;

གྲངས་ཀྱིས་མི་ཆོད་པ་ **GRANGS KYIS.MI.CHOD.PA** ‘uncountable, innumerable’;

མེད་དུ་མི་རུང་བ་ **MED.DU.MI.RUNG.BA** ‘indispensable, essential’;

བསླུ་བ་མེད་པ་ **BSLU.BA.MED.PA** ‘inevitable’;

སྨད་དུ་བྱུང་བ་ **SMAD.DU.BYUNG.BA** ‘wonderful’.

Construction of the adjectival predicate

In the Tibetic languages, adjectives function in a similar way as “intransitive (or monovalent) stative verbs”. Some authors even prefer to use the category of “stative verb” (Hoshi 2003, 2016) or “descriptive verb” as opposed to a class of “real adjectives” (see Bartee 2007). In many languages they are followed by a relator or suffix derived from a nominalizer, plus an auxiliary verb. The attested relators are – **PA/BA** (Ü, Ts, TN, also attested in CT), ◊ **LE** (HN, northern Kh), ◊ ◊ **NI** (Am).

The construction can be represented by the formula below: **ADJ+REL+_AUX**.

For example, ‘is bigger’ translates as མཆེ་བ་འདུག་ **CHE.BA.DUG**, མཆེ་ལེ་རེད་ **CHE.LE.RED**, མཆེ་ནི་རེད་ **CHE.NI.RED** depending on the language. In some languages such as Ü, the auxiliary may be dropped: **ADJ+REL+(AUX)**. In other languages, such as Southern Kham, Dzongkha and Choča-ngacha, etc., the adjective generally occurs without any auxiliary or relator.

Comparative constructions

There is a certain diversity of comparative constructions throughout the Tibetic linguistic area. The comparison implies two elements: the “comparee NP” (abbreviated as **CompNP**) and the “standard NP” (abbreviated as **StandNP**) as in the
following examples respectively from Common Tibetan, Southern Kham, Ladaks and In Choča ngača (in Bhutan):

(12) འབྲོང་གཡག་ལས་ཆེ་བ་ (འདུག)
    BRONG G’YAG-LAS CHE.BA (’DUG)
    (=standNP) (=CompNP)
    drong yak-ABL big-REL AUX

‘Wild yaks (drong) are bigger than (domestic) yaks.’ (Ü, ComTib)

(13) འབྲོང་གཡག་བས་ཆེ།
    BRONG G’YAG-BAS CHE
    drong yak-COMP big

‘Wild yaks (drong) are bigger than (domestic) yaks.’ (SKh)

(14) འབྲོང་གཡག་གི་སང་ཆེ་ཡ་ཡོད་ཀྱག
    BRONG G’YAG-GA’Y.SANG CHE-YA YOD.KYAG
    drong yak-GEN-COMP big-NMLZ EXV+FACT

‘Wild yaks (drong) are bigger than (domestic) yaks.’ (La)

(15) ཉང་ཝ་ཏ་ཁྱོད་རྒས།
    NGA-WA.TA KHYOD RGAS
    1sg-COMP 2sg Old

‘You are older than me.’ (Cho)

The standard NP often comes first and then the comparee NP occurs before the adjective. This order is probably the preferred or unmarked order. However, the reverse order is also attested and depends on the discursive parameters (topic, focus).

Here are the main comparative constructions attested in the Tibetic languages:

1. The comparee is marked by a case (comparative or ablative). The adjective may be marked by a comparative suffix (as in English ‘-er’) or has a positive form (unmarked). Thus in ex.12, the meaning could be rendered as “from the yaks, the wild yaks are bigger” and in ex. 13, as “from/than the yaks, the wild yaks are big.”

2. The comparee occurs in a subordinate clause. In English, it would correspond to: “If we consider the yaks, the wild yaks are bigger.” The verb BLTAS ‘to look’, ‘to consider’ is normally used in the subordinate clause.
3. The comparee occurs in a paratactic clause. The sentence structure could be rendered as “(I) won’t talk about the yak, the wild yaks are bigger.” The verb \textit{ZER} ‘to talk, to say’ is used in the paratactic clause.

The first construction is probably pervasive in the Tibetic area. See the examples above.

There is some diversity in the case used for the comparative: depending on the language, it may be derived from the CT comparative བས་ \textit{BAS} (Dz, Cho, Kh), ablative ལས་ \textit{LAS} (Ü, Ts, YK, Lho), elative ནས \textit{NAS} (Kh), ergative འས་ \textit{YIS} (Kh: Mi) (see below 8.1.9.). Additionally, a few markers are found such as ག ཚ \textit{SANG} (La, Sp, TN: Geyä), ི བ ཟ \textit{SUM} (TN: Gar, Tsanda; Kk), ཡ ལ ཉ ོ \textit{DES} (Kh: Ro, Sn). In a few southern Kham dialects and Sharkhok, one also encounters dissyllabic markers for the comparative: \textit{ko tó} /, \textit{bo} /.

Examples:

(16) \textit{NGA-SANG KHYOD MTHO-BA} 1SG-COMP 2SG tall-CS

‘You are taller than me.’ (Tö Ngari: Geyä; Qu and Tan 1983: 191)

In Ladaks, the compare NP is also followed by the morpheme \textit{SANG} but it is usually preceded by the genitive or, more frequently, the dative case:

(17) \textit{DPAL-L DAN NYE-RANG-NGI-SANG RING-MO DUG}
paldan nyerang-ngo-sang ringmo duk

Paldan 2SG-GEN-COMP tall EXV

‘Paldan is taller than you.’ (La)

The same meaning may be expressed by the dative:

(18) \textit{DPAL-L DAN NYE-RANG-NGA-SANG RING-MO DUG}
Paldan 2SG-DAT-COMP tall EXV

The second construction is frequent in Amdo as shown in the example below:
Let’s illustrate the third construction. It is found in Minyak Rabgang:

(21) གཡག་ཞེར་རྒྱུ་མེད འབྲོང་ཆེ་དོ།

G ′YAG-ZER-RGYU.MED ′BRONG CHE-DO
yak say-NMLZ-NEG+EXV drong big-SENS

‘Drong (wild yaks) are bigger than (domestic) yaks.’ (Kh)

Zeisler (2018c) mentions similar constructions of juxtaposition to convey comparative meanings in various dialects of Ladakh but also in Literary Tibetan. Her article also shows the diversity of available constructions within the Western Tibetic languages.

**Superlative constructions**

**Morphology**

The superlative may function as a predicate adjective or as an attributive adjective. In many Tibetic languages (Ü, Ts, YK, TN, Dz, etc.), the superlative form is composed of the stem followed by a superlative suffix such as ཤོས SHOS usually followed by a copulative verb when it has a predicative function.

For example, in Common Tibetan:

(22) དབེ་འདི་ཡག་ཤོས་རེད།

DEB ′DI YAG-SHOS RED
book DEM good-SUP be

‘This book is the best.’ (Ü, ComTib)
In some languages, the superlative suffix སོས་ SHOS is replaced by another suffix such as བ་ BA as in Derong-Jol (Kh) or བོ་ BO in Ladakh.

(23) གྲོ་ ཆེས་ ལྷེ་སྟེ། ལདྱ་བུ་ ལྷ་ཤེས་ སྡེ་བོ།

DEM book all-GEN in-ABL best-NMLZ be+FACT

‘Among all the books, this book is the best.’ (La)

In a number of languages of eastern Tibet, the superlative is formed by adding the intensive marker ོེས་ CHES ‘the most’ (Lit. ’great(ly)’):

(24) གྲོ་ ོེས་ རྒྱུན་ནི་རེད།

CHES PHYUG-NL.RED
most rich-STAT

‘(It) is the richest.’ (Am)

(25) གྲོ་ ོེས་ རྒྱུན་ རྒྱ་མོ་ ལྷེ།

YL.GE CHES YAG.MO RED
book most good be

‘This book is the best.’ (Kh)

To convey a superlative meaning, other intensive markers are also encountered in some languages, particularly Amdo and Kham.

ལེན་མོ་ཚེས་ A.THUR.GIS (Am) probably of Mongolian origin (Simon, pers. comm. 2020), རྒྱུན་སེའི་ SHED.SE (northern Kh), ནུ་ས། KUN (Kh: Sn) (Lit. ’all!’).

Additionally, in some dialects, the reduplication of the comparative form is also attested for the superlative, as in Bathang (Kham):

(26) གྲོ་ རྒྱ་མོ་ རྒྱ་མོ།

YAG-BA YAG-BA
good-CS good-CS

‘The best.’

Syntax of the superlative construction

The comparee is usually introduced by a postposition བོ་ NANG’in’ followed by a case which varies according to the regions. The case is normally derived from the
following CT cases: the elative ལས་ LAS, the ablative བས ས LAS and the locative བས ད N.A. In some dialects, a specific inessive case is found /nə/.

For example let’s illustrate the attribute function:

(27) 
\[ SLOB.MA-TSHO-TI, NANG-NAS CHUNG-SHOS SU RED \]
student-COL-GEN in-ABL young-most who be
‘Who is the youngest of the students?’ (Lit. ‘from the students’) (Ü, ComTib)

(28) 
\[ SLOB.MA-CHO-TI, NANG-NAS ATHURGIS CHUNG-NI-BO SU RED \]
student-COL-GEN in-ABL most young-NMLZ-DEF who COP(CT)
‘Who is the youngest of the students?’ (Am)

And then the predicative function:

(29) 
\[ SLOB.MA-TSHO-TI, NANG-NAS CHUNG-SHOS RED \]
student-COL-GEN in-ABL young-most be
‘(S/he) is the youngest of the students.’ (Ü)

(30) 
\[ SLOB.MA-CHO-TI, NANG-NAS ATHURGIS CHUNG-NI-BO RED \]
student-COL-GEN in-ABL most young-NMLZ-DEF be
‘(S/he) is the youngest of the students.’ (Am)

Other adjectival constructions and meanings

Several Tibetic languages (e.g. Ü, Ts) have additional suffixes which indicate ‘excessive’ (’too much’), ‘attenuative’ (’a little bit’), ‘interrogative’ (’how+adj+?’), ‘intensive’ (’very+adj’), ‘mirative’ (’how+adj+!’). They directly follow the adjective stem and are not used after reduplicated stems. (See e.g. Tournadre & Sangda Dorje 2003; Yliniemi 2019; Haller 2000, 2007, etc.)

23. The nominalizer NI followed by the definite marker is usually pronounced as /no/ (see Robin & Simon, forthcoming).
For example, ཆེ་དྲགས་ ‘too big’ (Ü), སྦོམ་དྲགས་ ‘too big’ (Dz), ཐག་རིང་ལོས་ ‘how far?’ (Ü), ཇུང་ཙམ་ ‘a bit small’ (Ü), སྦྱིད་པ་ལ་ ‘how nice!’ (Ü).

These suffixes are normally used in predicative constructions. (See also the section on “verb phrase.”)

Adverbial functions of the adjectives

Tibetic languages just as many linguistic families of the world do not have a “true category” of manner adverb. In order to render the meaning of manner adverbs, these languages use adjectives sometimes followed by a case (usually the dative).

\[(31)\] མགྱོགས་པོ (R) ། མཇལ་ཡོང་

‘See you soon (Lit. ‘quick’).’ (Ü, ComTib)

\[(32)\] ནོར་མེར་ མགྱོགས་པོ

‘(A) quick pace.’ (Ü, ComTib)

When the adjective is used as an “adverb” it may not only be marked by a case as mentioned but also has a distinct syntactic position and often occurs before the verb.

The syntax of adjective

In most Tibetic languages, adjectives are postposed to the noun: Noun + Adj. This is also the case in CT.

However, in Balti and Purik, the adjectives usually occur before the noun: Adj + Noun. The position of adjectives is one of the important syntactic features that distinguish Ladaks from Purik and Balti.

The anteposition of the adjective is also possible in CT and even Common Tibetan to indicate a restrictive function, but the adjective must be followed by a genitive: Adj + Genitive + Noun.

In Balti and Purik, since the adjective normally precedes the noun, the genitive is not required unlike in CT.
Thus, for example the sentence མཁྱི་ནག་པོ་ཅིག་ (KHYI NAG.P.O.PO CIG) ‘a black dog’ (Lit. ‘Dog black a’) in Common Tibetan and many languages becomes in Balti and Purik རྣ་པོ་ཁྱི་ཅིག་ (NAG.P.O.KHYI CIG) (Lit. ‘Black dog a’). Here are other examples:


In most languages, as mentioned above, adjectives typically occur as post-head modifiers. However, when a speaker wishes to place special emphasis or focus on an adjective, the adjective may also be placed before the head. The following examples are from Common Tibetan:

(33) དཀར་པོ་འི་མོ་ཊ་གསར་པ་ག་པར་འདུག
DKAR.P.O.-I MO.TA GSAR.PA GA.PAR 'DUG
white-GEN car new where EXV+SENS
‘Where is the new white car?’ (Ü, ComTib)

(34) འགྲེལ་པའི་མོ་ཊ་དཀར་པོ་ག་པར་འདུག
GSAR.PA.-I MO.TA DKAR.P.O GA.PAR 'DUG
new-GEN car white where EXV+SENS
‘Where is the new white car?’ (Ü, ComTib)

8.1.8. Ideophones

A morphological category which is pervasive in the Tibetic languages and often found in many ST languages is the category of “ideophones” which are used as adjective predicates (see above) but also have an adverbial function.

Ideophones convey a vivid representation of an idea in the form of a sound. From a semantic point of view, they have an expressive and emotional function and convey subjective and often intense perceptions of sound (in which case, they are normally derived from onomatopoeias), color, smell, form, or events. For Beck (2008) quoting Doke (1935), the term “ideophone” refers to onomatopoeic or synesthetic expression which “are distinguished as a group by syntactic, morphological, and /or phonological
properties, tend to have an emotive function and are associated with spoken and dramatic registers of speech.”

Ideophones are distinguished in Tibetic languages by their morphology: they are normally quadrisyllabic (with an echo reduplication) but some may be disyllabic (Yliniemi 2019). These ideophones are found in the entire Tibetic area but they may vary a lot in their forms and semantism.
evil', རྡེ་ལི་རྡི་ལེ་ RDE.RI~RDE.RE 'dismayed, downcast' (Am), རྡེ་ལི་ལེ་ RDE.LE 'swaying, wobbling' (Am), རྡེ་ལི་ རྡེ་ལེ་ KHR.LLE~KHROL.LE (CT) 'many ornaments hanging down and jingling together' 'BAG.CIG~BUG.CIG' completely silent, still without flies', etc.

For a detailed description of ideophones in Purik and Lhoke, see respectively Zemp (2018) and Yliniemi (2019).

**Dramatizers**

In some Tibetic languages, one also finds a category of words called dramatizers also called intensifiers that "may be viewed as a subcategory of ideophones" (Zemp 2013a), but have also specific features. They usually precede a verb. Dramatizers are also attested in other Tibetic languages such as Ladaks (Zeisler pers. comm.), Yolmo (Hari and Lama 2004), and Jirel (Strahm and Maibaum 2005).

Here are some examples of dramatizers found in Purik (Zemp 2013a): རྫོར་ /rwar/ 'noise of sudden rain, etc.' together with the verb CHAP.YONG 'to rain', གཙུག /tsug/ (sound of breaking an object) with the verb CHAG 'to break', དམ་ རྱམ /sprug/ 'to shake (apricot trees)', etc.

**8.1.9. Case marking**

The modern languages have preserved to a certain extent fundamental features of the CT grammatical cases (see 6.6.1). They are often multifunctional and some cases may serve as connectives when placed after a verb or a nominalized verb. In some languages, certain cases are optional (see e.g. Tournadre 1991; DeLancey 1991, 2011a). The grammatical cases of the Tibetic languages are morphologically clitics and normally atonal (Huber 2002; Graves 2007; Tournadre 2010; Yliniemi 2019). Their host is the last element of the noun phrase which can be a noun, an adjective, a plural (or collective) marker, a numeral or a determiner. These characteristics of case marking are valid for all the Tibetic languages.

---

24. Sources: Mélač et al. (2014), sKalezang’GYURMED & sKalezang dBYANG.CAN (2002), CTDT, Lama Sangye, Khyungpo (p.c.), Hua & KUL.BUM RGyal (1993) and the sK’GDMAS GLEGS R’AM.
The distinction between case affixes, clitics and adpositions is not always easy to make (Spencer & Luis 2012). Tibetic case markers are normally not considered as suffixes: it is easy to demonstrate that the markers occur once at the end of the noun phrase and they do not have other properties of suffixes. Some scholars such as DeLancey (2003a) or Strahm & Maibaum (2005: 809) have considered that Tibetan cases are actually more like postpositions. There are however some arguments suggesting that they behave like enclitics. First, the Tibetan case markers may never occur alone (without the noun) unlike adpositions which often have this property (cf. English, it’s under/after/on, etc.) and they form a prosodic word (together with their host). Tibetan case clitics may not be coordinated unlike some adpositions (in and out, on and off). They often undergo morphophonological alterations depending on the host phonological context, etc. However, the case clitics do sometimes exhibit properties which are also shared by adpositions, such as the possibility of combining together two markers (cf. English in-to, on-to, with-in, up-on, etc.). For both CT and modern Tibetic languages, this phenomenon is attested. Here are some examples: NAS-SU/nesu/ (ablative+purposive) in Spiti, Khunu and upper Ladakh (see also Zeisler 2011, 2018a), /le-gi/ (locative+genitive) in Yolmo, /du-ki/ (purposive+ablative) in Jirel, etc., but such combinations remain quite exceptional.

As we have seen in chapter 6, the CT system can be described as consisting of 10 cases. They include grammatical cases: absolutive (Ø), ergative ལྷི་ (GYIS), dative ལ་ (LA), and local cases, purposive དུ་ (DU), comitative དང་ (DANG), elative ཀྲན་ (NAS), ablative ང་ (LAS), genitive གྱི་ (GYI), locative ཉ་ (NA), and comparative ར་ (BAS).

These markers have been inherited in the modern languages, although no language has preserved the totality of the cases.

**Ergative**

Ergative marking derived from the CT form ལྷི་ (GYIS) or its conditional allomorphs ལྷི་ (GYIS), ལྷི་ (YIS), etc. (see chapt. 6) is attested in all the modern languages. Various reflexes are found: /-gi:/ (Ü, Ts, etc.), /-gi/ (Dz), /-ka/ (Kh) /-γa/ (Sk), /-γa/ (Sn: Gyt), /-i/s/ (Ba, Pu), /-ze/ (Sham), etc. In many languages, the forms of the ergative forms have merged with those of the genitive or are very similar with the latter. In most languages, the ergative is morphologically identical to the instrumental but the ergative
and instrumental roles differ in their syntax. In some languages such as Ladaks, the comitative case *DANG* is used instead to convey the instrumental function (ex. ‘to write *with* a pencil’).

Ergative constructions are present in all the Tibetic languages with very few exceptions (such as Baima), but the grammatical functions of the ergative may vary substantially. First the number of the "action verbs" that trigger the ergative constructions may differ for the various languages, but "action verbs" always constitute the major verb class. For example, from a semantic point of view, some verbs which denote perceptions such as ‘to see, to hear’ or even emotions ‘to love, to fear’ are treated as "action verbs" (with a subject in the ergative) in some languages, while they are considered respectively as “reception verb” (with a subject in the dative/aesthetive case) and as “emotion verbs” (with a “subject” in the absolutive).

In a canonical way, the ergative marks the Agent of a transitive verb (A), i.e in our terminology an "action verb", see below the verb classes). It is used with both controllable and non-controllable verbs (see below the section on lexical verbs). In a few languages such as Amdo, the ergative is essentially used in this canonical way. In some languages such as Purik, Balti and Ladaks, the ergative construction is restricted to controllable verbs (see Zeisler 2007, 2011).

However, in many languages, there are restrictions to this pattern. The ergative may be only compulsory with the completed\(^{25}\) aspect and optional with the uncompleted aspect and the future. This “optionality” of the ergative in some environments has been described by several authors (see Tournadre 1991, 2010; DeLancey 2011a). Whenever the ergative is not compulsory, it may indicate a pragmatic sense of emphasis or contrast. This “pragmatic ergativity” attested in many Tibetic languages is also found in other TB languages (LaPolla 1995). In some languages such as Baima, the ergative

---

\(^{25}\) About verbal aspect in Tibetic see section 8.4.1 We use the terms “completed” and “uncompleted” avoiding the terms “perfective” and “imperfective” which correspond to notions found in some specific aspectual systems such as Slavonic languages or some Indic languages (see Guentchéva 2016).
or agentive is even more restricted and used mainly to disambiguate two potential agents (Chirkova 2005).

Finally, the ergative may also be used with intransitive controllable verbs, to create an emphasis on the agent. This specific function is attested in some languages as Ü (see Tournadre 1991, 1996a), Tsang or Kyirong (Huber 2002), Sherpa (Graves 2007). It is also found (albeit not frequently) in the Classical language (see Kesang Gyurmé 1992; Tournadre 1996a, 2010; Hoshi 2016).

Absolutive

The absolutive marks both the intransitive “subject”, S and the grammatical patient of a transitive verb, P. It is always unmarked in all Tibetic languages. As we have mentioned above, in some languages, the absolutive marking of an intransitive subject (S) may alternate with the ergative. When there is an emphasis on S or P, the absolutive may be replaced in some languages by a dative (see Tournadre 1996a; Simon 2011; Zeisler 2007, 2012a).

Dative

The dative ṽ LA has been inherited in many languages, but it has a lot of realizations in modern languages such as: /-la/ (Ü, Ts, Lo, etc.), /-lo/ (Lh), /lv (Dz), /lv (Kh, Cho), /-lo/ (Kh), /-a/ (La, Am), /-e/ (Cho), etc. The dative is used to mark the grammatical beneficiary (also called recipient) and in some cases the grammatical patient. Additionally, in virtually all the Tibetic languages, the “subject” of the possessive constructions is generally marked by the dative and the possessed object is indicated by the absolutive case. For this special dative function, Zeisler (2007) has coined the term “aesthetive” case26 (see also Tournadre 2010 about the various functions of the dative and the traditional grammatical treatment of this case).

Purposive

Reflexes of the purposive case (also called “terminative case”, Hill 2011) ṽ DU and its allomorphs ṽ RU, ṽ R, ṽ SU, etc. (see chapt. 6) are rarely found in modern languages.

26. Tournadre (2010, 2012) has proposed to call this grammatical role “recipient”, “possessor” or “ceptor”.
with the notable exceptions of Purik and Balti (see Zemp 2018). Forms derived from the CT purposive case also appear at least marginally in Zhikatse (see Haller 2000) and they also occur in other dialects as fossilized expressions such as ལྷག་པར་དུ་ LHAG-PAR-DU ‘particularly’, གཉེས་པར་དུ་ NGES-PAR-DU ‘certainly, definitely’ (Ü) or རིན་ རིག་ YID-DU-YONG ‘to remember’ (Lit. ‘come in mind’) (La).

Comitative

The comitative case conveys the meaning of ‘with’ in English. Forms derived from the comitative དང་ DANG are attested in many languages across the Tibetic area, however in most modern languages, it is mainly used as a connective. In some languages, it is still used as a grammatical comitative case (Ü, La). In Ladaks and Purik, the comitative is also used to convey the instrumental meaning (“by’). The form མ་ LA or its variant ར་ R which corresponds to the dative in CT, are used in Amdo for this function (see Simon 2016).

Genitive

Reflexes of the genitive ཡི་ GI or its conditional allomorphs ཡི་ GYI or ཡི་ GI, are found in all modern languages under various forms: found /-gi/ (Ü, Ts, Dz), /-kə/ (Kh), /-ɣə/ (Sk), /-yə/ (Sn: Gyälthang), /-i/ or /e/ (La, Ba, Pu), etc. The genitive (together with a nominalizer) is also used to form relative clauses in most modern languages (as in CT).

Ablative

Tournadre (2010) provided the following description of the ablative LAS in CT: The meaning of the ablative is rather specific and much more restricted semantically than the elative NAS. Semantically, it indicates the spatial origin when the figure is on the surface of the referent (and not inside). Thus, for example སྲིད་པའི་ཐམས་ཅད RTA-LAS BABS ‘(s/he) dismounted the horse’, X was on the back of the horse (and not inside the horse!). It is the same with: སྲིད་པའི་ཐམས་ཅད BRAG-LAS-LHUNG ‘(s/he) has fallen from the rock’ For these meanings, LAS and NAS are practically equivalent: སྲིད་པའི་ཐམས་ཅད BRAG-NAS LUNG ‘(s/he) has fallen from the rock’ and སྲིད་པའི་ཐམས་ཅད RTA-NAS BABS ‘(s/he) dismounted the horse.’ However, in most examples, LAS does not simply indicate the spatial origin but rather the origin of a transformation: the object (or the being) from which, something is extracted, generated or produced. In these cases, the replacement
of Las by Nas is impossible as show the examples below. མ་ལས་མང་ལས་ཐེན་ལས་འོ་མ་བྱུང་ 0MA-ABL:ORI 0BYUNG 'The milk comes from the cow', འོ་མ་ལས་མར་བླངས་ 0MA-ABL:ORI MAR+Ø BLANGS 'One takes (extract) butter from milk'. (See also Hill 2011, 2012a.)

Many modern languages no longer make a distinction between ablative and elative. However, reflexes of the ablative ཟས་ Las are attested in a few modern languages such as Lhoke and Dzongkha. They are generally not found in Kham. In Ú and Tsang, Las is attested but only has a comparative function (see below).

Elative

The elative indicates the provenance or the source. The elative བདག Nas is quite pervasive in modern languages, but some southern languages, such as Dzongkha, Choča-ngā, etc. do not have a case derived from Nas. In some languages, the elative may also be used (rarely) to convey an ergative meaning. This is the case in formal speech in Common Tibetan (see Tournadre and Sangda Dorje 1998, 2003) and Amdo (see Simon 2016).

Locative

Forms derived from the locative བདག Na are attested nearly all the modern languages, however the locative case meaning has been preserved only in some languages such as Amdo, Lhoke, Dzongkha (in a marginal way).

In most other languages (Ú, Ts, Kh, Hor, etc.), བདག Na only functions as the conditional 'if' (after a verb) and no longer conveys a locative case function. In some languages, the conditional is marked by a construction which includes a form historically derived from the case marker Na such as: བདག་ བལོས་པོ་NE (Lho) (see Yliniemi 2019), བདག་ བྲི་(DZ) < ? CT བར་མཇི་Na =NMLZ+EXV 'if' lit. 'If there is'.

It is worth noting that many Amdo varieties have developed an opposition for the locative depending on the verbal aspect: བདག Na is used in the uncompleted aspect whereas བདག Nas /ni/ is used in the completed aspect. (See Sung & LHA.BYAM.RGYAL 2005: 254.)
Comparative

Forms reflecting the comparative བས་ BAS are only found in a few languages such as Dzongkha under the form ར་-wā (van Driem 1998), Choća-ngāča with -wata/ (Tournadre & Karma Rigzin 2015) and some dialects of Kham under the form -peʔ/. For the comparative function, some languages (Ū, Ts, Lho) use the ablative བས་ LAS. A number of Western languages use the comparative case བས་ SANG which is not attested in CT (La, Sp, TN: Purang, Gegye), བས་ SANG (Sham).

Summary of the characteristic of modern case systems

Most modern languages have a simplified case system and have undergone syncretism. For example, as mentioned above, several languages mark the ergative GIS in the same way as the genitive GI in some morphological contexts, thus exhibiting a partial syncretism (cf. Am, YK, Zkt, Dz), but personal pronouns usually maintain the distinction. The elative NAS and the ablative LAS functions are no longer distinct. The locative NA and dative LA have usually merged into a single function. The function of the comparative ablative case LAS and the comparative case BAS are also no longer found together in a single language.27

The minimum number of case markers is probably 4: absolutive, ergative/genitive, dative/locative and ablative. Many case systems consist of 5 cases: absolutive, ergative,28 dative, genitive, and ablative. In some languages (Gyālthang, Kongpo, Spiti, etc.), the ergative has a very marginal status and mainly a pragmatic function.

Others case systems may have up to 7 or 8 cases (La, Kh: Ro, Kh: Zhollam, Dz). The frequently attested case markers include absolutive, ergative, dative, genitive, locative, ablative, comparative, comitative and instrumental.

Some languages are currently developing new casual morphemes for the inessive (‘in, on’), the adessive (‘close to’), the comparative (‘more than’) or even the ablative

27. Even Classical, or Old Tibetan, no longer had an opposition between these two cases, and it seems the use of BAS versus LAS is more a matter of style (archaic versus innovative).

28. In CT, the “agentive” subsumes both ergative and instrumental functions. However in modern Tibetic languages, one encounters systems whereby the ergative is morphologically different from the instrumental. Thus we use here the term “ergative”.
However, the grammaticalization is often not complete. These morphemes are derived from the following terms: 'GO'top' or 'MGO'head' > 'on', 'NANG'inside' > 'in' (both in Kh, Dz), 'RTSA'root' > 'close' (Lho) and 'SANG'comparative'(La, Sp, TN: Purang, Gegye) which may be derived from CT 'RTSA-NA; 'KHA'mouth, surface' > 'K.A., a locative case in Purik and Ladaks (see e.g. Zemp 2018). A special form for the ablative is used in 'DAS' in Rongdrak Kham and Zhollam, but its origin is unclear.

8.1.10. Number

Number is usually not compulsory in the Tibetic languages and it is not comparable to the opposition singular/plural in European languages.

For these reasons, we will use here the term “collective marker” (COL) rather than “plural marker” (see Jan Rijkhoff 2001a-b).

In fact, there are reasons to consider that number is not entirely grammaticalized in Tibetic languages. There is no specific form for the singular. There are markers that do indicate “plurality”; however, this type of marking is not obligatory as in the European languages and it is always absent with a numeral.

(35) ✐BU.MO.GSUM
 girl three
 ‘(The) three girls.’

(36) ✐BU.MO.GSUM-TSHO
girl three-COL

Intended meaning: ‘(the) three girls.’

(37) ✐BU.MO-TSHO.GSUM
girl-COL three

Intended meaning: ‘(the) three girls.’

Additionally, collective markers indicate that the referent is a non-singleton set and not simply a multiple of singular objects (books) or collectives (families). In many
cases, the use of collective markers (or “plural” markers) is restricted to definite and animate beings when it is directly affixed to the head nouns, but they may be used with inanimate entities when preceded by a demonstrative.

The collective marking is more systematic in the Western regions (Ladakh, Baltistan). In all the Tibetic languages, collective markers appear as clitics that are either postponed to demonstrative or the definite article and appear before the case morpheme:

- **NOUN-DEM-COL-CASE**
- **dem-NOUN-def-col-case**

or directly to the noun as:

- **NOUN-COL-CASE**

Thus, in Common Tibetan we have:

(38) བུ་མོ་ཚོ།  girl-COL
girl-COL

‘The girls.’

(39) དགེ་རྒན་ཚོ།  teacher-COL

‘The teachers.’

(40) བླ་མ་ཚོ།  lama-COL

‘The lamas.’

(41) བུ་མོ་  girl-DEM

‘Those girls.’

(42) དགེ་རྒན་  teacher-DEM

‘Those teachers.’

(43) བླ་མ་  lama-DEM

‘These lamas.’

The collective markers may be derived from one of the morphemes used in literary Tibetan for this function. They include the following markers:

- **ཀུན་ KUN** particularly in the western regions of Ladakh, Baltistan and Spiti. In Baltistan and many dialects of Ladakh KUN has some variants /kun/, /gun/,
and after vowels: /un/, /ung/. KUN is also attested in southern Kham such as Gyālthang and sDerong-nJol (/kɛ/), and the variant /küntsʔ/ ◊ KUN-TSHAD is used in Ngari (Gar).

- TSHO particularly in Central Tibet and Dzongkha (where it is written ◊ TSHU) but also attested in Amdo: /zo/ (Am: Xunhua),
- TSHAMS-(CAD) in Kham < CT ‘all’.
- TSHANG, MA or simply ◊ TSHANG are widespread in the Tibetic area,
- RNAMS used e.g. in Kham,
- DAG in Ü and Tsang.
- CAG. The various forms ◊ kyaʔ/ (in Tö and Nagchu), /kyaʔN/ ◊ kyaʔr (Hor: Amdo), /-kya/ ◊ kyo/ (Kyirong), /-yo/ ◊ ya/ ◊ ya/ (Chamdo), /-ya/ ◊ ya/ (Yolmo, Mustang), and possibly /-kyaʔ/ ◊ kyaʔ/ (Lhasa) as well as ◊ sak/ (Ladaks, Purik)29 are probably all derived from CT ◊ CAG.
- GANG.PO < CT ‘all’ is used for ‘collective for animate beings’ (Cho).
- RIGS /rak/ (Am, Kham: Derge; see Häsl 1999) < CT ‘type’.
- TSHANG /sang/ or /zang/ < CT ‘nest’. Amdo: Xunhua and many other dialects (see Simon 2016).
- GRAL ‘row, line’ (Tsang).
- CHABO (Am) or ◊ CHO (Amdo) < CT ◊ CHA’a pair, a couple’.
- SLO.SKOR /logo/ (Thewo). The origin of this collective marker is unclear.

Most collective markers are clearly derived from nouns. That is the case of ◊ TSHO and ◊ TSHANG, respectively derived from ◊ TSHOGS’ assembly, group’,30 and ◊ TSHANG’nest’. The Amdo marker /čao/, /čo/ ◊ CHABO may be derived from ◊ CHA-TSHANG’entirely, all’ or from ◊ CHA’pair’. The collective marker ◊ RIGS

29. According to Koshal (1979, 2005), /sak/ is used as a free variant of /kun ~ gun/. According to Zeisler (pers. comm.) this is restricted to the Sham dialects.

30. This type of derivation is attested in many languages throughout the world. See e.g. Mauritian Creole (cf. plural marker /ban/ < French bande ‘a set, a gang’).
is derived from the noun ‘type’. The Tsang collective marker ཡོང་  རྒལ་  ‘GRAL’ comes from the noun ‘row, line’.

Here are some examples respectively with the markers RNAMS, CHABO, GRAL, NG (KUN), RIGS:

(44) རྟ་རྣམས།  ‘the horses’ (Kh: Derge)
(45) གོས་རྣམས།  ‘the clothes’ (Kh: Derge)
(46) བེ་ཆ་ཆ་བོ།  ‘the books’ (Am: Labrang)
(47) བེ་ཆ་གྲལ།  ‘the books’ (Ts: Zhikatse)
(48) སྤེན་པོ།  ‘the girls’ (Ba)
(49) མྱི་རིགས།  ‘the people’ (Am).

In some languages, several types of collective markers may coexist (Simon 2016). For example, in some Amdo dialects (Yadzi, Hualong), ཕྱ རིགས་  ‘RIGS’, སྒྲ་ ’TSANG’, སྒྲ་  ‘GRAL’, བོ་ སྒྲ་  are used with slightly a different meaning.

### 8.2. Postposition phrase

Just as CT, all the modern Tibetic languages have postpositions. The use of prepositions is not attested in the Tibetic languages. Postpositions are diachronically derived from relator nouns. The term “relator noun” has been used by some authors such as DeLancey (2005) to describe a functional category which is neither reducible to the category of lexical noun nor to that of adposition. We chose to use the term of “postpositions” rather than “relator nouns” because they show a higher degree of grammaticalization, similar to postpositions. For example, depending on the languages, the postposition is not always followed by a case and often occurs in a construction which corresponds to:

NP-POST

However, the postpositions clearly retain some nominal properties: in many cases, the postpositions govern a genitive case on the noun and are often marked by a spatial

31. The collective probably indicates definite reference in most cases but this needs further research.
The basic structure of postposition phrases are as follows:

NP- (case: GEN) POST-(case: DAT/LOC/ELA/ABL)-(TOP/ADM)

For example:

(50) བཀྲ་ཁང་(ཀི) རང་ལ། ZA.KHANG(-GI) NANG.LA
    restaurant (-GEN) interior-DAT
    ‘In the restaurant.’ (Ü, ComTib)

(51) བཀྲ་ཁང་(ཀི) རང་ལས RDZONG.KHAG(-GI) NANG-LAS
    district (-GEN) in-ABL
    ‘Among the districts.’ (Dz)

(52) བཀྲ་ཁང་(ཀི) རང་ལ། KHANG.PA-’I RGYAB-LA
    house-GEN back-DAT
    ‘Behind the house.’ (Ü, ComTib)

(53) བཀྲ་ཁང་(ཀི) རང་ལ། KHANG.BA-(GI) LTAG.GA-NA
    house (-GEN) back-LOC
    ‘Behind the house.’ (Am)

(54) བཀྲ་ཁང་(ཀི) BR.’GO-STENG
    mountain top-on
    ‘On the top of the mountain.’ (Kh)

In many languages, the genitive case following the noun is optional and in other languages the use of the genitive is ungrammatical. The optionality depends on various morphophonological paremeters (final vowel or consonant and the number of syllable) as well as the style (formal or casual). This optionality is also found in CT.
The spatial case following the postposition is also omitted in many languages, as shown in the Kham example above ('on the mountain'). Thus, the structure **NP-POST** is quite frequent.

It is interesting to note that this strategy to develop postpositions from nouns has been followed in the entire family but using different lexical items (see the list below).

The main postpositions are derived from the following CT relator nouns: **ནང་** NANG ‘inside’ > ‘in’, **ཕྱི** PHYI ‘outside’ > ‘out of’, **ཀླད་པ་** KLAD ‘top part’ (related to **ཀླད་** ‘brain’ on the ‘top of the body’) > ‘on, at the top of’ (a flat surface), **ཕྱིག** ‘lower part’ > ‘below, under’, **གོང་** GONG upper part < ‘before’, **འོག་** OG ‘face’ > ‘in front of, before’, probably related to **མདུན་** MDUN ‘in front of’, **དོན་** DON ‘meaning’ > ‘for, on behalf of, in order to’, **ཐོག་** THOG ‘roof’ > ‘on’, **ཕྱི’** PHYI ‘mouth’, ‘surface’ > ‘on’, **མགོ་** MGO ‘head’ > ‘on’, **སྒང་** SGANG ‘hill’ > ‘on’, **སྟེང་** STENG ‘upper part’ > ‘above’, **ཀླད་** KLAD ‘top part’ (related to **ཀླད་པ་** ‘brain’ on the ‘top of the body’) > ‘on, at the top of’ (a flat surface), **པར་** BAR ‘middle part’ > ‘between, until’, **དཀྱིལ་** DKYIL ‘center’ > ‘in the middle of, among’, **ཐོག་** THOG ‘head’ > ‘on’, **མ་གཏོགས་** MA.GTOGS ‘except’ < CT NEG+verb ‘belong’, **ནང་བཞིན་** NANG.BZHIN ‘alike, like, same’ < CT, **བཞིན་** BZHIN ‘alike, like, same’ < OT འད་ (? OT འད་) STGOS(LA) ‘alike, like, same’ < OT འད་ འད་ (LA) STGOS(-PA)‘and the like, etc.’.  

As one can notice from the list above, a number of relator nouns indicate body parts such as ‘head’, ‘mouth’, ‘back’, ‘cheek’, ‘heel’, ‘rib’, which reflect the general anthropomorphic trend of human languages.

In a more marginal way, postpositions may derive from other categories such as verbs and include compound postpositions: **སྐོར་** SKOR ‘about’ < CT verb ‘to circle, go around’, **མ་གཏོགས་** MA.GTOGS ‘except’ < CT NEG+verb ‘belong’, **ནང་བཞིན་** NANG.BZHIN ‘alike, like, same’ < CT, **ཅིམ་** TSOS(LA) ‘alike, like, same’ < OT འད་ འད་ (LA) STGOS(-PA)‘and the like, etc.’.
Most of the spatial and temporal postpositions such as སྔོན་ SNGON ‘before’, གོང་ GONG ‘before’, རྗེས་ RJES ‘after’, etc. also function as connectives and occur after nominalized verbs.

### 8.3. Verbal predicate

In the Tibetic languages, the verb may be preceded or followed by various markers indicating direction, negation as well as tag, direct questions or jussive. The verb is often followed by an auxiliary. Auxiliation plays a major role in the Tibetic languages. Additionally the verb phrase may also include a modal verb.

Two structures should be distinguished for the verbal predicate:

- the verb occurs without an auxiliary
- the verb is followed by an auxiliary which indicates tense, aspect, modality and evidentiality.

#### 8.3.1. Predicate without auxiliary verb

In the former case, which is less frequent, one essentially finds two basic structures may be preceded or followed by various clitic markers indicating the direction (DIR), the negation (NEG), interrogation (Q), marked by prefixed interrogative markers (PQ) or final interrogative markers (FQ), tag questions (TAG) and jussive mood (JUS) as shown below. In some languages of the family, the verb itself may be inflected (FLEX), and the inflections may be inherited from Classical Tibetan (see chap. 6) or innovative (see 8.3.2., below).

Depending on the languages and on the various tenses (or TAM), one finds the following sequences of markers for the verbal predicate:

a) (DIR)-(NEG)-VERB[FLEX]-(FQ/TAG/JUS)
b) (DIR)-(PQ)-VERB[FLEX]-TAG

The main difference between the constructions (a) and (b) is the position of the interrogative marker.

In (a), the interrogation is indicated by a sentence-final question marker (FQ) and it may co-occur with the negation which is placed in the position before the verb stem. In (b)
a prefixed question marker (PQ) precedes the verb and occupy the same slot as the negation in (a), and in this case the negation and the interrogative markers are incompatible.

Both the suffixed and the prefixed interrogative markers are found in CT, respectively as དམ་ (or its allomorphs) and འེ. Most modern languages have either prefixed interrogative marker derived from འེ or a postverbal interrogative marker. Some languages use དམ་ instead of འེ. Postverbal interrogative marking is found in most languages whereas prefixed interrogative marking is dominant in Kham and Amdo. In the languages with a preverbal marking, one also finds postverbal markers, because as mentioned above, the prefixed interrogative marker is incompatible with the negation.

Thus, for example, for the copulative verbsཡིན་ and རེད་, one finds respectively the interrogative: འེ-ཡིན་ and འེ-རེད་ as well as the negative forms མིན་ (the contraction of མ་ཡིན་) and མ་རེད་, but the combinations འེ-མིན་ and འེ-མ་ཡིན་ are not attested. In order to utter a negative interrogative sentence, one has to use final question markers: མིན་ཟིམ་ (CT), མིན་ (Am), མིན་པས་/min-pä/ (Ü, Ts), མིན་ (La), མ་ཐྲེད་དམ་ (CT), མ་ཐྲེད་པ་ (Kh).

This incompatibility has also been noted for Amdo by Sung and lHa byams rgyal (2005: 65), but it can be generalized to all the eastern Tibetic languages, which have preserved prefixed question markers (PQ). Thus, negative interrogative sentences are normally expressed by the structure (a), i.e. those using sentence final question markers.

However, the prefixed interrogative marker is compatible with a tag question:

(55) འེ-ཟིམ་ཆ་ (CHI)
PQ-be tasty-TAG
'I wonder if it is tasty.' (Am)

(56) མིན་ (Am)
PQ-CPV-TAG
'I wonder if it is the case.' (Am)
Another peculiarity of the negative marker is that it can be inserted inside a disyllabic verb as shown in the following example which is inserted between the syllables HA and GO of the verb: ཡ་གོ ‘to know, understand’.

(57) ༩༽ ཡ་མི་གོ
HA MI-GO
know-NEG-STEM
'I don’t know.' (Kh)

The predicate may be reduced to the verb alone. It is particularly frequent in the imperative constructions.

(58) ༩༽ འགྲོ།
'GRO
go (PRS)
'Go! / Let’s go!’ (Û, ComTib)

The various tenses are often marked by auxiliaries, but, in some languages, the verb alone may indicate tenses or aspects such as the aoristic past, the intentional future or the present.

(59) ༩༽ མ་སོང་།
MA-SONG
NEG-go (PST)
'I did not go.' (Am)

(60) ༩༽ ག་བྱེད།
NGA BYED
1SG do (PRS)
'I will do it.' (Kh)

(61) ༩༽ མི་བྱིན།
MI-BYIN
NEG-give (PST)
'(I) won’t give (it).’ (Dz)
8.3.2. Predicate with auxiliary verb

In the Tibetic languages, the verb is usually followed by an auxiliary (the various auxiliaries are listed below in 8.3.3 and 8.3.4). The auxiliary may occur alone after the verb but is often accompanied by a relator (REL) which corresponds to a nominalizer (see 8.3.12) or to a connective such as གི, སྟེ or ནས (see e.g. DeLancey 1991, 2011b; Tournadre & Konchok Jiatso 2001) linking it to the lexical verb. The most frequent predicate structure attested in the Tibetic languages may thus be formulated in the following way:

VERB[LEX]-REL-AUX

Auxiliaries are attested for all tenses, aspects and modalities including imperative and prohibitive as in the example below: predicate involving a directional, a negation and an imperative auxiliary:

(65) ཡར་མ་འགྲོ་རོགས་
     YAR-MA-'GRO-ROGS
     DIR-NEG-go-JUS
     ‘Please, don’t come up here (towards me)!’ (Kh)

The relator (REL, in the scheme above) is sometimes morphologically fused with the auxiliary, but it may be separated from the auxiliary by the negation (NEG) or by the preverbal question marker (PQ). In various dialects (Ü, Ts, SKh, Am), the auxiliary
may be dropped (particularly in affirmative sentences) and the verb is then only followed by a connective.

Auxiliaries may also be used alone, without the verb, when answering questions as a way of saying 'yes' or 'no'. On the other hand, a zero relator is also found in a paradigm of auxiliaries, e.g., in southern Kham, in which the zero relator is related to a completed aspect.32

In modern Tibetic languages, the relator and the auxiliary have often merged and one may analyze these forms as suffixes or verb endings as shown by Zeisler (2004). However, from a diachronic and comparative point of view,33 it is better to use the concept of auxiliary and we will continue to use this term, even if from a strictly synchronic point of view, the terms verb endings or verb suffixes are often more appropriate.

Another frequently attested structure includes a "secondary verb"34 which is inserted between the lexical verb and the final auxiliary. The secondary verb (see SKAL.BZANG 'GYUR.MED 1992; Tournadre & Konchok Jiatso 2001; Bartee 2007; Yliniemi 2019) is a modal verb or an aspectual or directional auxiliary35 (see also below section 8.3.10):

VERB[flex]-SEC-(REL)-AUX

In some cases, the secondary verb occurs without auxiliaries (see examples below). In a few languages such as Amdo, one finds sometimes two secondary verbs in a sequence:

VERB[flex]-SEC1-SEC2-(REL)-AUX

---

32. This is also true in the variety of Common Tibetan spoken in the Diaspora. Depending on the analysis and the cases, some may prefer to consider that the relator is simply absent.

33. The auxiliary may also occur alone in some specific contexts, e.g. when answering questions. So it has retained some syntactic autonomy.

34. The term "secondary verb" (see SKAL.BZANG 'GYUR.MED 1992) was proposed by N. Tournadre to translate SKAL.BZANG 'GYUR.MED's term BYATSHIG.PHAL.PA.

35. From a phonological and prosodic point of view, the various languages seem to behave differently. In Lhasa, the secondary (even preceded by the negation) forms one prosodic word while in Kham or Amdo, the verb and the secondary may constitute two prosodic words.
Additionally, after the auxiliary, one finds final interrogative suffixes (FQ), tag questions (TAG) and jussive markers (JUS)

\[ \text{VERB[flex]} - (\text{sec}) - (\text{rel}) - \text{AUX} - (FQ/TAG/JUS) \]

As for the elements preceding a verb stem or an auxiliary, interrogative prefixes (PQ) and negative prefixes (NEG) are also frequently attested. The latter are found in all the Tibetic languages (see section 8.4.9), while the interrogative prefixes are mainly found in the eastern Tibetan languages of Amdo and Kham.

Interrogative and negative prefixes may occur in three syntactic positions: either before the lexical verb:

\[ \text{NEG/PQ-VERB[flex]} - (\text{sec}) - (\text{rel}) - \text{AUX} \]

or before the auxiliary verb:

\[ \text{VERB[flex]} - (\text{sec}) - (\text{rel}) - \text{NEG/PQ-AUX} \]

And finally, when a secondary verb is present, NEG and PQ may precede it:

\[ \text{VERB[flex]} - \text{NEG/PQ-SEC-(rel)-AUX} \]

The position of negative and interrogative prefixed markers depends both on the TAME meaning of a given language and varies from a language to the other.

Note again that since the negative and interrogative prefixes occupy the same syntactic slot, they may not co-occur. For negative interrogative questions, one has essentially two options:

a) One uses final interrogative suffixes with the negation before the auxiliary or the verb.

b) One uses interrogative prefixes before the auxiliary and raises the negative prefix before the verb (or secondary verb).

Additionally in some languages, the verb may be preceded by a directional marker (DIR).
Whenever a causative verb is present, it follows the main verb. This may be expressed by the formula:

**VERB**[**FLEX**]-**(REL)**-CAUS-(SEC)-(REL)-AUX

Here is an example of this complex structure:

(67) ◊ བེ་རིང་ འཇིག་ བོད་ལོ་ བོད་ འཇུག་ མ་ཐུབ་ཐལ།

`DE.RING` `ʔA.YIS` `LO.LON-NA` `Z.A.MA` `Z.A-GI` `JUGMA-THUB-THAL.

today lady old-DAT meal eat-REL CAUS NEG-SEC-AUX

'Today I could not make the old lady eat.' (Am)

**ZA** is the main verb, **JUG** corresponds to the causative verb, **THUB** is a modal verb functioning as a secondary verb and **THAL** is the auxiliary.

In summary, the complete schematized expression of the predicate is:

( **DIR**)-VERB[ **FLEX**]-**(DIR)**-( **REL**)-( **CAUS**)-(SEC)-(REL)-(NEG/PQ)-AUX-( **FQ**/TAG/JUS)

(Note that in some rare cases, there may be two secondary verbs, as above. Also, the negation may occur also before the lexical verb, the secondary verb or the causative verb).

In terms of morphosyntactic and morphosemantic categories, the main categories present in the verbal predicate are: (a) the lexical verb and the auxiliary (AUX) and (b) optionally causative and secondary verbs, (c) directional, relator, negative and interrogative markers.

This formula corresponds to the maximal expansion of the predicate.

Note that for Tormarong (Dongwang) Tibetan, Bartee provides a similar formula for the following expanded schematized expression (we added the bold outline):

**(DIR)** (NEG-) **VERB** ( **CAUS**) (DIR) (MOD) (ASP) (QST) (FNL AUX) (EV/VAL)

However, this formula includes semantic categories which are not always manifested by autonomous morphological units, such as "modality", "aspect" or "evidential". These operators may be expressed by the final auxiliary or by the secondary verb.
Here are some examples:

(68) གླུ་ལེན་རྒྱུ་མ་རེད།
GLU LEN-RGYU.MA.RED
song take[PRES]-REL+NEG+_AUX
‘(He) won’t sing.’ (Am)

(69)◊སྒོར་མོ་མ་བྱིན་ན་ངས་ལས་ཀ་ལས་ནི་མ་རེད།
SGOR.MO MA-BYIN-NA NGA-S LAS.KA LAS-NL.MA.RED
1SG NEG-give[PAST]-if I-ERG work work-REL+NEG+_AUX
‘If he does give money, I won’t work.’ (Am)

A relator without any auxiliaries can appear depending on language. This is true for example, in Amdo and some eastern languages:

(70)ང་འགྲོ་རྒྱུ་ཡིན་(ཡིན་)
NGA 'GRO-RGYU(-YIN)
1SG go-REL-NEG-AUX
‘I will go.’ (Kh)

The following example illustrates the double auxiliation:

(71)བཞུགས་ཡོད་(པ་)མ་རེད།
BZHUGS-YOD-(PA)-MA.RED
sit-AUX-REL-NEG-AUX
‘(S/he) did not stay.’ (Ü, ComTib)

Other examples with prefixed interrogative and negative markers.

(72)◊འགྱོ་རྒྱུ་ཨེ་རེད།
GYO-RGYU-ʔE-RED
go[flex]-REL-PQ-AUX
‘Are you going?’ (Am)

(73)མ་བཞུགས་པ་རེད།
MA-BZHUGS-PA.RED
NEG-sit-REL+_AUX
‘(S/he) did not stay.’ [S/he refused to stay] (Ü, ComTib)
In some marginal cases, the verb may be separated from the secondary verb by a relator as in the following example:
8.3.3. Copulative and existential verbs

Concerning verb types in the Tibetic languages, it is necessary to distinguish between copulative and existential verbs (CEV), auxiliary verbs, secondary verbs and lexical verbs. As we will see the CEV verbs as well as auxiliary verbs play a fundamental role in the expression of evidentiality and epistemic modalities.

Copulative and existential verbs are essentially derived from equative verbs (‘to be’) as well as existence, posture and perception verbs.

As we will see, lexical verbs in section 8.3.6 may be divided into various verb classes based on their semantics and their syntactic structures. These verb classes have an impact on the behavior of evidential and epistemic auxiliaries. Some verb classes do not combine with certain types of auxiliaries. Thus, one should say that both CEV and lexical verb classes play a fundamental role in the functioning of the Tibetic evidentiality and epistemicity.

8.3.3.1. Copulative verbs

Copulative verbs function as a syntactic copula and generally correspond in English to the verb ‘to be’ but they additionally convey grammatical semantic values, namely evidential and/or epistemic meanings as well as the speaker’s stance. Thus, depending on the various Tibetic languages, they may indicate egophoricity, inference, factuality, etc. as shown below in 8.4.

This is one of the main characteristic features in the Tibetic languages. Each language and dialect may differ in the grammatical semantics conveyed by copulative verbs as well as the form of the verbs.

The copulative verb ཁྱིན་ .Requires additional context or symbols to represent the notation here. to be’ attested in OT is found in all the Tibetic languages (with virtually no exception). In Dzongkha and Lhoke, it is spelled as ཁྱིན་. It appears in various pronunciations such as /’工商银行/, /’工行/, /’工商银行/ and /’工商银行/. ཁྱིན་ has a
specific negative form མིན་MIN although the form མ་ཡིན་MA-YIN is marginally attested (in Dzayül).

Many compound copulative and auxiliary verbs are derived from this verb (see below 8.3.3.3):

Other frequent copulative verbs include:

- རེད་RED 'to be' [Ü, Kh, Hor, Am, Tö(P)]. The form /rak/ in Lo Mönthang (Mustang, see Kretschmar 1995) is probably cognate with /reʔ/ attested in Tö pastoralists' dialects (Qu & Tan 1983).

- བེད་BED for Ts, Lho) /pa:/, /po:/, /pa:/ [Ts: Nyemo, Lhokha] < CT བྱ་BYA 'to do' (fut.) are also attested. Concerning the origin of this auxiliary other hypotheses have been proposed such as འབད་bad (Haller 2000) 'to make an effort, to strive', འབབ་bab 'to descend to coincide form དབའ་DBA' (BOD.RGYA TSHIG.MDZOD CHEN.MO) a dialectal form for emphatic assertion, བྱེད་BYED 'to do' or the future form བྱ་BYA which are already used as an auxiliary in CT (see SKAL.BZANG’GYUR.MED 1992).

- དག་DAG 'to be' /'dak, "da?/ (Tö[C])< CT 'to be correct' and the variants: དེད་/‘dāi/, ‘de:/ (Tö[C], LJ), ལག་NAG /nak/ (Lo), ལ་NA /na/ 'to be' (Bro, in Bhutan) and probably Sherpa ཉ་DZA /dza/ (see also /da/ in Naaba, eastern Nepal), in /་NOG/nok/ (Sp, La, Eastern Sham). In some cases, the

36. This orthography RED would be the more appropriate to show the derivation from BYED. In many areas of the Western Tibetosphere, the yatak /y/ is dropped in the spoken dialects. The preinitial S in SBED is not etymological and used only for the pronunciation sake.

37. In the case of NOG, another possible etymology would be the verb YUG. Two arguments are in favor of this hypothesis: the change of the vowel /a/ to /o/ and the fact that in Ladak, YIN.NO has two entirely different meanings: one is the equivalent of YIN.DAG 'factual' and the other is the sensory visual inferential (see 8.4.3) which could thus be reconstructed as YIN:DUG.
verb is preceded by *YIN*: ཡིན་འདག *YIN_ Dag* (LJ), ཡིན་ད་ *YIN_Da* (Bro) (this form is ‘inferential’ (or ‘acquired’) as opposed to ཡིན་‘factual’ (or ‘assimilated’), see Funk 2020), ཡིན་ཛ་ *YIN_Dza* (Sh), ཡིན་ཉོག་ *YIN_Nog* (La, Sp).

Alternative candidates could be proposed such as ཡིན་ད་ ‘to pass’ (past: ཡིན་དེ་ ད་/ past ‘to pass’) or ཡིན་ད་ ‘to desire, wish’ but in both there is problem to explain the existence of a final /k/ in western Tö and Ladakh. So the best candidate is probably the verb ཡིན་‘Dag’ ‘to be correct’.

In a marginal way, one also finds the following copulative verbs:

- ཀིས་ *GIS* ‘to be’ (Thewo),
- རྗིས་ *Zin* ‘to be’ (Kh:Sn) and བྱང་ *Snang* ‘to be’ (Kh:Sn) < CT ‘to appear’,
- དྲག་ *Grag* ‘to be’ (Kh:Sn, Derge).

As we have seen above there are numerous forms for the copulative verb ‘to be’ in the Tibetic languages. However, in the great majority of languages, the verb ‘to be’ is derived from one (or several) of the 4 following CT roots: ཡིན་, རེད་, བྱེད་ and ཡིན་‘Dag’.

8.3.3.2. Existential verbs

In the languages of the world existence, location, possession and attribution are often expressed by different verbs (copulative, existential, possessive, stative, etc.) and various syntactic constructions. However, in many Tibetic languages, existence, location, possession and attribution are expressed by existential verbs. It seems a pervasive characteristic of the language family to express both the existence and the location through the same existential verbs.

Even when the four functions are conveyed by a single existential verb, they require different syntactic constructions. For example, the existential verb *Yod* or *Dug* indicate existence, location, possession and attributive, but each of these meanings is often related to a different argument position and case marking. See below the examples 82-85 in Common Tibetan:

38. This verb is also used in the traditional Buddhist debate to mean "it is right."
Existence:

(82) གཡག་དཀར་པོ་གཅིག་འདུག

‘There is one white yak.’

Location:

(83) གཡག་དཀར་པོ་དེ་སྤང་རི་ལ་འདུག

‘The white yak is on the pasture.’

Possession:

(84) ང་བོ་ལ་གཡག་དཀར་པོ་གཅིག་འདུག

‘Tshering has one white yak.’

Attribution:

(85) དཀོན་པོ་དཀར་པོ་གཅིག་འདུག

‘White yaks are rare.’

In some languages, the dative and the locative are marked in the same way and thus the possession and location constructions may sometimes look similar. However, the possessor usually occurs in the first position while the location complement occurs before the verb (in a neutral statement). This similarity of construction is not found in some languages (Amdo, Lhoke, Dzongkha, etc.) since the possession and the location markers have different forms.

Based on the analysis of Tokpe Gola and Common Tibetan, Caplow (2000) has coined the abbreviation ELPA to refer to the various function of the existential verb (Existence, Location, Possession and Attribution). This term has also been taken up by
various scholars such as Garrett (2001) and Gawne (2016) but, even if this term is useful, it cannot apply to all the Tibetic languages for various reasons.

(a) As mentioned above, the existential verb does not exhibit all the four functions in the Tibetic languages. For example, some Kham languages (such as Gyalthang and Lhagang) or the Sham dialect of Ladakh (Zeisler, pers. comm. 2020) do not use existential verbs to express the attributive meaning and use instead a copulative verb or something else.

(b) In some languages of southern Kham, the possession is expressed by an existential verb which is distinct from the existential verb expressing the other functions.

(c) In some languages such as Ladaks, the copulative verb *YIN* ‘to be’ is sometimes used to express location.

We will thus continue to use the traditional and more transparent term “existential verb” in this book.

Additionally, just as with copulative verbs, existential verbs usually also convey grammatical semantic values, namely evidential and/or epistemic meanings. We observe different morphosyntactic constructions in the modern languages. The main differences between existential verbs in the various languages are related to the use of distinct verbs (*མེད* `YOD`, *དུག* `DUG`, *དག་* `GDA`, *སྣང* `SNANG`, *གྲག* `GRAG`, see below), case marking and word order. It is worth noting that a few languages merely have one lexical verb, which is often the old existential CT verb *མེད* `YOD`, but most languages have a paradigm of suppletive existential verbs to convey the various evidential and epistemic modalities.

The main existential verbs which convey the meaning ‘to exist, to have’ in the various languages are: *མེད* `YOD` ‘to exist, to have’ (Ba, La, Ü, Ts, Dz, Am, E, Kh, Lho, etc.) or

39. Moreover, in many languages around the world existential verbs are used for a variety of functions, and thus this is not specific to Tibetic languages. For ex. in Russian the verb *есть* is used for the existence, the possession and marginally the location. In Hebrew, *יְשָׁה* is also used for the existence and possession. What is specific to most Tibetic languages is to cumulate various syntactic and semantic functions and play an important role in the expression of evidentiality and epistemic modalities.
its archaic form འོད་ ’OD (Tö, Kh, Hor, etc.) and its specific negative forms: མེད་ MED ‘NEG+exist’ (Ba, La, Ü, Ts, Dz, Am, Kh, Lho, etc.) or its archaic form མྱེད་ MYED ‘NEG+exist’ (Kh, E, etc.). The verb འོད་ YOD is already attested in Old Tibetan.

The verb འདུག་ DUG or its variant འདུག་ NUG < CT ‘to stay, to sit, to exist’ or other derived forms is also nearly pervasive. Reflexes of འདུག་ DUG are found in most Tibetic languages (Pur, La, Sp, Tö, Yol, Ts, Ü, Kh, Sh, Dz, Lho). One should note that when the verb འདུག་ DUG is used as an auxiliary (see 8.3.4), or as a component of a compound CEV (see 8.3.3.3), it may take a lot of forms such as འདུག་ NUG (Dz), འོད་ TUG (La), འདུག་ SUG or the variants འདུག་ TSUG and འདུག་ TSHUG (La, Pur, Ba, Kh: Derge; see Koshal 1979; Zeisler 2017; Häsler 1999), འཇི་ SI (Kh; SK.BZANG ’GYUR.MED & SK.BZANG.DYANGS.CAN 2002), འཇི་ ZUG, often transcribed as འཇི་ ZIG (Am; ibid.).

Other verbs include གདའ་ GDA’ (Kh, Hor) < CT ‘to exist’ and maybe originally ‘to sit’ (see the derived form གདན་ GDAN ‘carpet’), འཇི་ GIG or འཇི་ GGE (Kh: Northern route), མཆིས་ MCHIS < CT ‘to exist’. The verb རྒྱ་ GRAG < CT ‘to sound, to be heard of’ (East Purik, Ladaks, Ladakh Jangthang, Garzha, Spiti, Tö, Kham) is also widely attested. ཚགས་ SNANG < CT ‘to appear, to shine, to become visible’ is found in Phânpo (central Tibet), in the Pari (Hwari) dialect of Amdo in many varieties of Kham (Bathang, Gyälthang, Chathreng, Zhollam, Tormarong, etc.), in some languages of the northeast region such as Thewo, Čone, Drugchu, Sharkhok, and Khöpokhok in Gansu and Sichuan (see Suzuki 2012d; RIG’DZIN DBANG.MO 2013; Ebihara 2017), in the Nubra dialects of Ladakh and in Balti, notably Turtuk, Khaplu and Tyakshi dialects (see Ebihara 2014). This existential copula and auxiliary is also attested in the Pangi dialect spoken in Himachal Pradesh.

A few languages use different verbs to indicate “possession” vs. “existence/location” (e.g. respectively འོད་ YOD versus ཚགས་ SNANG in Gyälthang). Additionally, in the southern area of Kham, ‘animacy’ is also reflected in the system of existential verbs (see Bartee 2007). Both copulative and existential verbs also function as auxiliary to indicate evidential and epistemic values. (see 8.3.10 and 8.4.3).
8.3.3.3. Compound CEV

Among the specificities of the Tibetic verb systems, we ought mention the existence of many compound CEV made of copulative and existential verbs. The old copulative and existential verbs འདབ 'to be' and འོད 'YOD combine with other auxiliaries according to the following patterns:

**Copulative:** འདབ 'YIN + (NMLZ/CONN.) + AUX

**Existential:** འོད 'YOD + (NMLZ/CONN.) + AUX

These patterns are attested in most (if not all) the Tibetic languages:

Ex. of compound copulative verbs: [factual] འདབ་པ་ རེད་ (Ü), འདབ་ནི་ རེད་ (Am), འདབ་ལེ་ རེད་ (Kh), འདབ་གྲག་ (La), འདབ་ཁྱེགས་ (Dag, Lh), འདབ་ཐིག་ (La), འདབ་ཐིག་བཤད་ (La).

Ex. of compound existential verbs: [factual] འོད་པ་ རེད་ (Ü), འོད་ནི་ རེད་ (Am), འོད་ལེ་ རེད་ (Kh), འོད་གྲག་ (La), འོད་ཁྱེགས་ (Dag, Lh), འོད་ཁྱེགས་བཤད་ (La), འོད་ཁྱེགས་ (Dag, Lh), འོད་ཁྱེགས་ (La).

8.3.3.4. The main morphological and functional differences in the CEV

In the comparative charts below we find the main CEV attested in the Tibetic languages. The first chart illustrates the CEV in Western languages and dialects such
as Tö Ngari, Spiti, Ladaks, Eastern Sham, Western Sham, Eastern Purik, Balti (Turtuk, Hardas and Skardo).\(^{41}\)

### Chart VIII.1. – Differences in the CEV in the Western regions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>དགུད /duk/</td>
<td>དགུད /duk/</td>
<td>དགུད /duk/</td>
<td>དགུད /duk/</td>
<td>དགུད /duk/</td>
</tr>
<tr>
<td>Non vis.</td>
<td>གྲག /rak/, /dak/ /ṭak/</td>
<td>གྲག /rak/, /dak/ /ṭak/</td>
<td>དགུད /duk/</td>
<td>དགུད /duk/</td>
<td>དགུད /duk/</td>
</tr>
<tr>
<td>Endo.</td>
<td>གྲག /rak/, /dak/ /ṭak/</td>
<td>གྲག /rak/, /dak/ /ṭak/</td>
<td>དགུད /duk/</td>
<td>དགུད /duk/</td>
<td>དགུད /duk/</td>
</tr>
<tr>
<td>Auth. or Ego. Cop.</td>
<td>ཡིན /yin/</td>
<td>ཡིན /yin/</td>
<td>ཡིན /yin/</td>
<td>ཡིན /yin/</td>
<td>ཡིན /yin/</td>
</tr>
<tr>
<td>Factual Cop.</td>
<td>ཡིན /yin/</td>
<td>ཡིན /yin/</td>
<td>ཡིན /yin/</td>
<td>ཡིན /yin/</td>
<td>ཡིན /yin/</td>
</tr>
</tbody>
</table>

\(^{41}\) Some of the squares are left empty. This is the case of the factual category. This is due to the lack of precise data. The copulative verb ཡིན སྟེ /yin-st/ and existential verb ཡོད སྟེ /yod-st/ do exist in Balti and Purik as well as Ladaks. Despite the fact that the forms are identical or very similar, their grammatical meaning is quite different. They clearly convey an inferential meaning (often associated with as past tense) and not a factual one (see Zemp 2018). The difficulty is also due to the polysemy of some forms. For example, ཡིན སྟེ /yin-st/ in Ladaks has two entirely different meanings: it functions as a factual marker and as a visual inferential. (Sec 8.4.3.) For simplicity’s sake, we have grouped together the authoritative and egophoric categories (see also 8.4.3.5).

\(^{42}\) This form may originally come from ཡོད སྟེ /yod-st/.
In the chart below are the main CEV attested in the Eastern and Central languages and dialects such as Thewo (eastern section), Amdo, Northern Kham, Hor, Tsang and Ü.

**Chart VIII.2. – Differences in the CEV in the Central and Eastern regions**

<table>
<thead>
<tr>
<th>Eastern S.</th>
<th>Amdo</th>
<th>NKham (Derge)</th>
<th>Hor, NKham</th>
<th>Ü, Phänpo</th>
<th>Ü, Tsang</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sens (Vis., Non-vis. Endo.)</td>
<td>སྣང་ / n’õ/</td>
<td>YODGI / yo/</td>
<td>GI /‘go/</td>
<td>GI / da/ etc.</td>
<td>སྣང་ / nang/</td>
</tr>
<tr>
<td>Egophoric cop.</td>
<td>YIN</td>
<td>YIN</td>
<td>YIN</td>
<td>YIN</td>
<td>YIN</td>
</tr>
<tr>
<td>Factual cop.</td>
<td>GI /‘gi/</td>
<td>RED /‘re/</td>
<td>RED /‘re/</td>
<td>RED /‘re/</td>
<td>RED /‘re/, /ba/</td>
</tr>
<tr>
<td>Ego. exist.</td>
<td>YOD /‘ye/</td>
<td>YOD / yot/, yol/</td>
<td>YOD /‘yot/</td>
<td>YOD /‘yot/</td>
<td>YOD /‘yot</td>
</tr>
<tr>
<td>Factual Exist.</td>
<td>YOD,LE,GI /‘yelegi/</td>
<td>YOD,LE,GI /yonire/</td>
<td>YOD,LE,RED /‘yolec/</td>
<td>YOD,LE,RED /‘olerc/</td>
<td>YOD,LE,RED /‘yore/</td>
</tr>
</tbody>
</table>

Finally in the southern languages such as Sherpa, Dzongkha and Lhoke, we find the following CEV:
As we can see from the above comparative charts, the main CEV present both morphological and functional variations. However, the main CEV found in the majority of the Tibetic languages are essentially derived from the CT verbs འིན་ YIN, འགྲེང་ YOD, གོང་ DUG, གྲག་ SNANG, གཅིག་ GRAG, གུ་མ་ GDA’, རྡེ་ RED, སྨོད་ BYED and དབང་ DAG. Concerning the evidential functions of the CEV see 8.4.3.

### 8.3.4. Auxiliary verbs

Auxiliary verbs play a central role in the marking of tenses, aspects and modal categories. They are used with all the tenses and convey evidential and epistemic values as well as the speaker’s stance. Additionally they also may convey intentionality, directionality and deontic meanings.
Auxiliaries are mainly derived from the copulative and existential verbs (CEV, see above) as well as motion or transfer verbs. As mentioned in 8.3.2., the auxiliary may occur alone after the verb but is often preceded by a nominalizer or a connective.

Aside from the CEV, the list of the main auxiliaries found in the various Tibetic languages includes:

- **evidential auxiliaries**: སོང་ SONG < CT ‘to go’ (Ü, Ts, Sh, Lj), ཞིང་ THAL < CT ‘to go’ (Hor, Kh, Am, Sharkhok, etc.), ཡོང་ YONG < CT ‘to come’ (Ü, Ts), ཟི་ ZHE (or the variant སྲོ་ CE (Lho, Cho) as well as སྲི་ YI or the variant མྦ CI (Dz) < ? CT ཤི་ MCHIS ‘to exist, to come’, ཞིང་ BYUNG (Ü, Kh, Lj) < CT ‘to become’, རིན་ BZHAG < CT ‘to put’ (Am, Ü), ཡོང་ MYONG < CT ‘to taste’, རོ་ རོ་ DGOS < CT ‘to need’, རེ་ THON ‘to come/ go out’ (Kh: Gyālthang).

- **epistemic auxiliaries**: འོང་ ONG / འོངས་ ONGS < CT ‘to come’ and their variants ཡོང་ YONG, འོངས་ YONGS, རོ་ GRO < CT ‘to go’

- **jussive auxiliary**: ཤོག་ SHOG < CT རྟོགས་ GSHEGS ‘to go’ (H).

Most auxiliaries verbs are already found in CT. Only a few verbs found in the modern languages have no obvious correspondences with Classical forms (they appear with the sign ◊ in the list above and in the CEV list in 8.3.3.). One should note that a number of these verbs were already used as auxiliaries in CT: འབུ། YIN, འབྲ་ BYED, སྣང་ SNANG, རད་ DUG, རག་ GDA, ཤི་ MCHIS, པྱི་ GRAG (see 8.3.3.1. and 8.3.3.2.).

The auxiliary རེ་ RED ‘to be’, attested in Central Tibetan, Kham and Amdo, appeared in some texts, already in the 14th century (Shao 2016) but was not widespread in CT. The copulative verb འདག DAG < CT ‘to be correct’ used in Tö Ngari, Spiti and Ladakh is not reported as an auxiliary verb in CT.

Some auxiliaries such as སྲི་ ZHE or the variant སྲོ་ CE (Lho, Cho) as well as སྲི་ YI or the variant མྦ CI (Dz). These forms may derive from the CT verb ཤི་ MCHIS (van Driem 1998).
8.3.5. Secondary verbs

Secondary verbs play an important role in the marking of modal categories. They mainly indicate deontic, intentional as well as directional values.

The term secondary verb (see section 8.3.2) was coined by Tournadre in his translation of Kesang Gyurmé’s grammar (1992) to render the expression བྱ་ཚིག་ཕལ་པ་ BYATSHIG PHAL.PA used by SKAL.BZANG’GYUR.MED (1981). From a syntactic point of view, secondary verbs correspond to the second verb of serial verb constructions and are derived from converbs (see DeLancey 1991, 2011b; Yliniemi 2019; Zeisler 2004, 2019). As we have seen earlier (8.3.2), the secondary verbs occur after the lexical verb and before the auxiliary.

Although secondary verbs resemble auxiliaries and, in some cases, may even function as auxiliaries, it is important to distinguish them for several reasons:

(a) They generally convey modal, aspectual as well as directional values, but unlike auxiliaries, do not convey epistemic and evidential values.

(b) The secondary verbs are in most cases not autonomous and must be followed by auxiliaries.

(c) Unlike many “true auxiliaries”, secondary verbs are normally not preceded by relators.

(d) They are slightly less grammaticalized than auxiliaries and are more related to the lexical meaning of the verb.

Some verbs may function as auxiliaries in some languages whereas they function as secondary verbs in other languages. For example, the verb SONG ‘to go’ is an auxiliary in Central Tibetan and Sherpa while it is a secondary verb in Amdo and Dzongkha. Both functions may coexist within a single language. Also, དགོས DGOS ‘to need’ is both a secondary verb and an auxiliary in Common Tibetan as appears in the examples 86 and 87 below:

(86) ཉང་བྱས་དགོས། NGAS BYAS-DGOS 1SG-ERG do+FUT+BEN+EGO
‘I will do it (for you).’ [benefactive meaning]
In the first example above, Dgos has a benefactive function and refers to an action which the speaker proposes to achieve for the hearer’s benefit. In this function, Dgos is a future benefactive auxiliary conveying an egophoric intentional meaning (only compatible with the 1st person singular, and sometimes plural, of a controllable verb) whereas in the second example, it is a secondary verb with the modal meaning ‘need’.

Secondary verbs are attested in most modern languages such as Amdo (see Ebihara 2005; Tournadre & Shao, forthcoming), Dzongkha (van Driem 1998), Spiti (Hein 2007), Common Tibetan (Tournadre & Sangda Dorje 1998, 2003), (Tournadre & Konchok Jitso 2001), etc.

Frequently attested verbs include བཏང་ BTANG < CT ‘to send’, སོང་ SONG < CT ‘to go’ (past, both attested in Amdo and Dzongkha), ཡོང་ YONG / ལོང་ LONG ‘to come’. Additionally, van Driem (1998) mentions for Dzongkha the verb མཆིས་ MCHIS (< CT ‘to exist, to go’) and provides explanations that show the status of “secondary verb” although he does not use this notion and calls it an auxiliary.

In the following examples of serial verb constructions in Common Tibetan and Amdo (but also many other dialect groups), the secondary verb occurs after the lexical verb: མཁྱེར་འགྲོ་ KHYER GRO ‘to take away’, མཁྱེར་ཡོང་ KHYER YONG (contracted as མཁྱོང་ KHYONG in Ladakh) ‘to bring’, སྟོགས་ལོག་ LOG GRO ‘to return (away)’, སྟོགས་ལོག་ LOG YONG ‘to return (back)’ and in Amdo: རྡོ་ཐོན་ THON BTANG ‘to arrive’, རྡོ་ཐོན་ THON SONG ‘to arrive’.

Here is a list of the secondary verbs in the various languages.

**Motion and transfer verbs:** བཏང་ BTANG < CT ‘to send’ and the variant ཤོང་ YONG, སོང་ SONG < CT ‘to go (past)’, ལོག་ SONG < CT ‘to go (back)’ (Hor, Kham, Amdo, Sharkhok, etc.), རབ་ BZHAG < CT ‘to put’ (Am, Lho), རློང་ BTANG < CT ‘to send’ (Am, Dz, Sp, Lho), 0
Modal and psychological verbs: མོ་'to be allowed', ཨེ་‘to want’, ཉན་‘can’, འཛན་‘may’, རྣམ་‘want’, བྱུང་‘to experience, to taste’, རྒྱུད་‘to finish’, དྲུག་‘to finish’, ཁྱུན་‘to finish’ (Dz), འབང་‘to be time to’, བེཟུ་‘to know’, དུང་‘to dare’, བཟླ་‘to count’, to have the intention to’, མོ་‘to be about to’, འབུ་‘to think’, འཕོ་‘to think’, གདུམ་‘to be about to’, དྲག་‘to have the time to’, དུ་‘to be able’ (Sh) or the related form found in Kham དུ་‘to be able’ (Kh: Nagkerak, Tormarong), དུ་‘can, to be able’ (Am), དུ་‘to be able’ (Dz)

As shown above most secondary verbs found in the modern languages have direct correspondences in CT, with only a few exceptions (appearing with the sign ◊ in the list above) that do not seem to be reflexes of CT forms.

8.3.6. Lexical verbs

For lexical verbs, we will examine various semantic, syntactic and morphological categories such as controllability, valency, and inflection.

As in other languages of the world, the verbs may be intransitive/monovalent or transitive/bivalent. In the case of bivalent verbs, from a semantic and syntactic point of view, it is necessary to distinguish the various classes of lexical verbs: verbs of action, reception, emotion, and association. (See the section 8.3.4.4.)

8.3.6.1. Verb classes

Action verbs designate physical, verbal and mental actions but extend to other semantic types such as sensory perceptions or psychological states. Action verbs constitute
the great majority of lexical verbs in the various languages, probably up to 90% of the verbs. E.g. རྡུང་ 'to beat', བསད་ 'to kill', ཡ་ 'to eat', བི་ 'to write', གྲུད་ 'to wash', དེན་ 'to take', གྱེར་ 'to take', དེར་ 'to say', བཤད་ 'to say, explain', གོ་ 'to repeat', ལབ་ 'to recite', ཏེར་ 'to think', གྲུད་ 'to forget', དཔོ་ 'to remember', ཆོ་ 'to see, to know', སྣུབ་ 'to see', གོ་ 'to hear', དོན་ 'to recite', རིག་ 'to think', རིག་ 'to feel', ཚོར་ 'to smell'. They also include a great number of compound verbs such as བསམ་བློ་བཏང་ 'to reflect (upon)', སྐྱེ ཤོས་ 'to give (to a child)' etc. Existential verbs (see section 8.3.1) such as ཡོད་ 'to exist', etc. which convey the meaning ‘to be born’, etc. also exhibit the same syntactic structure as reception verbs when they indicate possession. As we will see below, the ‘subject’ of these verbs is marked by the “aesthetive” or dative (see 8.1.7). In some western languages (Ba, Pur, Za, La), verbs of perception are also marked with the aesthetive and thus function in a similar way as the reception verbs.

Reception verbs designate processes directed towards a goal or receiver and also serve to indicate the possession. The list of reception verbs is very limited and comprises of a dozen verbs. However, these verbs are very frequent. They include: ཤོབ་ 'to obtain', ཨ་ དགོས་-'to need', ཀན། རྣག་-'to obtain', ཕོ་ རྣའེ་-'to find', རོ་ སྐྱེ 'to give birth (to a child)', ཀན། ཧོའི་-'to receive', ཀྲོ་ ལོག 'to be struck’ as well as some compound verbs such as ཀན། ཞྲིལ་ བཙང་ 'to dream', ཀན། མེ་ཤོར་ 'to catch fire', etc. Existential verbs (see section 8.3.1) such as ཤོམ་ 'YOD, ཤོམ་ རེད་ 'YOD, RED, རུ་ དུག, etc. which convey the meaning ‘to exist’, etc. also exhibit the same syntactic structure as reception verbs when they indicate possession. As we will see below, the ‘subject’ of these verbs is marked by the “aesthetive” or dative (see 8.1.7). In some western languages (Ba, Pur, Za, La), verbs of perception are also marked with the aesthetive and thus function in a similar way as the reception verbs.

Emotion verbs serve to indicate psychological attitudes or emotions such as ‘fear’, ‘love’, ‘surprise’, ‘wrath’, etc. As in the case of reception verbs, the list of emotion verbs is also very limited and contains a dozen verbs, some of which are very frequent. They

44. The category of “subject” is not adapted to many Asian languages but we use it here for simplicity’s sake. See below the grammatical roles.
45. The first meaning of སྐྱེ is ‘to be born’. However when used with a dative complement, it means ‘to give birth’. (Lit. ‘To be born to someone’.)
include འགོགས་ 'to love', བྱམས་ 'to love', བསྟུད་ 'to fear', ཕྱགས་ 'to fear', བསྟུད་ 'to fear', as well as some compound verbs such as བཀྲོང་ཁྲོ། 'être en colère', སེམས་པ་ཤོར་ 'être amoureux'. In CT one also finds verbs such as ཕྱང་ 'hate', བྱམས་ 'être gentil', to love', ཕྱགས་ 'to desire, to love', etc. As we will see below, the "subject" or experiencer of these verbs is marked by the absolutive and the object is marked by the dative.

**Association verbs** are used to indicate a separation or a contact between the "subject" and the "object" of the verb. These verbs have been noted by various scholars such as Beyer (1992), Kesang Gyurmé (1992), Wilson (1992), Hackett (2005), Hill (2004), Tournadre & Sangla Dorje (1998, 2003), Zeisler (2007). The list of association verbs also has a dozen verbs, but these verbs are less frequent than emotion or reception verbs. The list includes the following verbs: ལྡན་ 'to contradict, violate', སྲུང་ 'to be similar, comparable (with)', རྒྱུན་ 'to meet (with)', སྲིད་ 'to meet (with)' (H), རྒྱུན་ 'to be in accord (with)', རྒྱུན་ 'to fight, to quarrel (with)', རྒྱུན་ 'to fight (with)', སྲུང་ 'to associate (with)', རྒྱུན་ 'to tune with, to be in accord (with)', སྲུང་ 'to mix (with)', རྒྱུན་ 'to mix (with)', སྲུང་ 'to mix (with)', རྒྱུན་ 'to mix (with)', སྲུང་ 'to mix (with)', རྒྱུན་ 'to mix (with)', སྲུང་ 'to mix (with)', སྲུང་ 'to mix (with)', རྒྱུན་ 'to mix (with)', སྲུང་ 'to mix (with)', etc. As we will see below, the "object" of these verbs is marked by a special case the "comitative".

As we have briefly mentioned, the lexical semantics has also an impact on the syntactic patterns and the grammatical cases. We will come back to this issue with more details.

It should also be noted that not all the modern Tibetic have preserved these four basic classes of verbs (see also below the section 8.3.6.4 on valency).

46. In Kesang Gyurmé (1992), Tournadre's commentary (in italics) of the Tibetan translation mentions the grammatical role of བཀྲོང་ and provide a list of examples.
8.3.6.2. Controllability

One important feature of Tibetic languages is the semantic notion of "control" as noted by many authors (Tournadre 1996a; Tournadre and Sangda Dorje 1998, 2003; DeLancey 1986a, 1990; Huber 2002; Zeisler 2004, 2007; Haller 2000). Every verb that involves a potential agent or an undergoer can be either controllable or non-controllable and this constitutes a lexical property of the verb. As mentioned by Hoshi (2016), the concept of controllability does not apply to verbs that do not involve a potential agent. The marked category is "controllable" (+Ctr) and this category should be indicated in dictionaries and other lexicographic works.

The majority of action verbs are controllable verbs, but this class also includes non-controllable verbs. Perception verbs as well as physiological verbs (see example below) may be either controllable or non-controllable. Association verbs are either controllable or non-controllable, emotion and reception verbs are always non-controllable.

A controllable verb (abbreviated as +Ctr V) designates "controllable actions" i.e. actions which may be mastered by the agent and are intentionally performed, while non-controllable verbs (abbreviated as -Ctr V) refer to actions or situations that cannot be controlled by the agent or verbs that do not imply any agent. For example, verbs such as བོལ། LTA 'to look at', མ་ NYAN 'to listen', ད་ GRO 'to go', རྣ་ SLOB 'to learn', འི་ ZA 'to eat' are controllable while verbs such as རོ་ MTHONG 'to see', ཀོ GO 'to hear', བཱ་ HAGO 'to understand', བཱ་ SHI 'to die' are non-controllable.

Some authors such as DeLancey (1986a, 1990), Tournadre (1996a), Tournadre & Sangda Dorje (2003), Hoshi (2003, 2016) have used the terms "volitional" versus "non-volitional" to designate the opposition "controllable" versus "non controllable", however these two classes of verb do not only depend on the agent's will or volition, they also depend on the controllability of the given process or situation. For this reason, verbs like བུ། DGA 'to like', སྐྲ་ RNJED 'to find' are non-controllable because the "subject" can not entirely control the process even if the latter implies some degree of volition: one may 'want to find/to love'.

Non-controllable verbs in the Tibetan have morphological and syntactic properties that are different from controllable verbs:
a) In CT, non-controllable verbs never have an imperative stem, while controllable verbs often have a special imperative form: as ལོས་ ‘look!’, ཉོན་ ‘listen!’, གོ་ ‘go!’, རོ་ ‘eat!’.

b) In CT and modern languages, jussive particle such as ཤིག་ may not be used with non-controllable verbs while they are used with controllable verbs: ལོས་ཤིག་ ‘look!’, ཉོན་དང་ ‘listen!’, གོ་ ‘go!’, རོ་ ‘eat!’.

But forms such as ལོ་ ‘see!’,

གོ་ ‘hear!’,

ཧ་ ‘understand!’,

ཤི་ ‘die!’ are unacceptable. However, it seems that in a few dialects, the imperative and the prohibitive are acceptable with some verbs that allow a certain degree of controllability.

c) In CT and in many modern Tibetic languages, the ergative case generally is not used with non-controllable monovalent verbs (or “intransitive” verb) but is acceptable with controllable monovalent verbs: རོ་ ‘go!’ but *མཐོང་ ‘see!’;

གོ་ ‘hear!’;

ཧ་ ‘understand!’;

ཤི་ ‘die!’

his not acceptable. This optional marking on the agent of the monovalent verb generally indicates an emphasis. This type of agentive marking is not acceptable in some languages particularly in Amdo and Southern Kham.

d) In a number of modern languages (U, Ts, TN, Am, Ho, Northern Kh), intentional auxiliaries are not compatible with non-controllable verbs and may only co-occur with controllable verbs: བལྟས་པ་ ‘I (have) looked!’ but not ལོ་ ‘I (have) seen!’;

ལྟ་ ‘I will look (at it)’ but not རོ་ ‘I will see’;

ཉན་ ‘I (have) listened!’ but not དང་ ‘I (have) heard’;

ཤི་ ‘I (have) heard’ but not ཤི་ ‘I will hear’.

This phenomenon has been described by many authors such as Haller (2000), Qu & Tang (1983), Tournadre (1996), Huber (2002) and Zeisler (2004).

e) An additional distinction between controllable and non-controllable verbs is that the former may have up to three core or grammatical arguments while
the latter usually have only maximally two arguments.

In some languages, such as Balti and Ladaks, the morphological marking of the past with /s/ has been generalized for controllable verbs (Zeisler 2004).

One finds pairs of controllable and non-controllable verbs: བོག་ ‘to hear’, ཕུ་ ‘to look’ (+Ctr) vs. མཐོང་ ‘to see’. Such pairs are of course found in many languages of the world. However, in many Tibetic languages, the number of such pairs is quite high. For example, in Common Tibetan, one finds the following verbs: གད་མོ་བགད་ ‘to laugh’ (+Ctr) vs. གད་མོ་ ‘to burst into laughter’, གད་ ‘to cry/weep intentionally’ (+Ctr) vs. གད་ ‘to burst into tears’, གཅིན་ ‘to urinate’ (+Ctr), གཅིན་ ‘to urinate unintentionally’, གཅིན་ ‘to fart intentionally (+Ctr)’, གཅིན་ ‘to fart unintentionally’, ཕེ་ ‘to shout’ (+Ctr), ཁ་ ‘to scream (out of fear, etc.)’, གཏུ་ ‘to go to sleep’ (+Ctr), ཁ་ ‘to fall asleep’.

As mentioned by Huber (2002), “many factors, including cultural differences, are involved in determining what are controllable versus non-controllable verbs in a given language.” For example, the monovalent verb གད་ ‘to cry/weep’ is considered as a controllable verb in most Tibetic languages, a feature which might be surprising for speakers of European languages. The verb གད་ ‘to cry’ is compatible with the ergative, with the intentional auxiliaries and with the jussive suffixes and also possesses a specific imperative form (see a-d). To give the meaning ‘to burst into tears’ (which is uncontrollable), one has to use a different verb construction: གད་ ‘to remember’. Some rare lexical verbs such as གཏུ་ ‘meet’ may function both as controllable and non-controllable depending on their behavior. This is also the case of some honorific verbs that are ambiguous such as གཞིགས་ ‘to remember’ (-Ctr), གཞིགས་ ‘to see, look at’ or གཞིགས་ ‘to come, go, arrive’. In some cases, the two behaviors may trigger different case markings as shown by Zeisler (2007: 415). As shown above, in some cases, the controllability is marked by the light verb: ཡིད་དུ་གོང་ ‘to remember’ (-Ctr) verb, ཡིད་དུ་བཅོ་ ‘to remember intentionally’.
8.3.6.3. Observability and endopathic marking

Another important characteristic of lexical verbs in the Tibetic languages is the notion of “observability” (see Sun 1993; Garrett 2001; Tournadre 1998). Some body-internal sensations or experiences such as hunger, thirst, inner cold, headache, dream, psychological and emotional states or feelings etc. are not directly observable and may be perceived only by the experiencer.

The verbs that involve this type of “inner sensation” are compatible with “endopathic” auxiliaries (see below). This category of verbs [+Obs], related to non-observable phenomena such as an “inner sensation, an emotional state or mental activity” is essentially a subcategory of non-controllable verbs [-Ctr]. However, some rare controllable verbs [+Ctr] such as those indicating a mental activity also fall into the category of “non-observable”.

As mentioned above, these verbs exhibit a specific morphosyntactic behavior. They allow the use of a sensory marker which has an endopathic function (see 8.4.3.1) in affirmative sentences with 1st person, but does not usually occur with 2nd and 3rd person (see ex. below).

Other verbs function in the reverse way: they allow the use of a sensory marker in affirmative sentences with 2nd and 3rd person but not with the 1st person.

Here is an example of [+Obs V]. The following examples (88-95) are in Common Tibetan:

(88) KHONG.TSHO(-S) CHANG ‘THUNG-GL.’DUG)
3PL(-ERG) chang drink-PROG+SENS
‘They are drinking chang.’ [the speaker sees or has seen them drinking]

But the following sentence is incorrect, unless one sees oneself drinking. The sign # below indicates a marked sentence that is correct but requires a very specific situation.

(89) # NGA CHANG ‘THUNG-GL.’DUG)
1SG chang drink-PROG+SENS
‘I am drinking chang.’
Here is an example of [-Obs]:

(90) གློད་ཁོག་ལྟོགས་ཀྱི་འདུག

NGA GROD.KHOG LTOGS-KYL.(DUG)

1SG belly hungry-STAT+ENDO

'I am hungry.'

But the following sentence is normally not acceptable:

(91) ཀྲོད་ཁོག་ལྟོགས་ཀྱི་འདུག

KHONG/KHYED.RANG GROD.KHOG LTOGS-KYL.(DUG)

3SG/2SG belly hungry-STAT+ENDO

Intended meaning: 'He/ you are hungry.'

The same is true for emotional states:

(92) ིེད་ཀྱི་འདུག

NGA ZHED-KYL.(DUG)

1SG afraid-STAT+ENDO

'I am afraid, scared.'

(93) ཀྲོད་ིབས་ཀྱི་འདུག

KHONG ZHED-KHYAG

3SG afraid-PERF+INF+SENS

Intended meaning: 'He is afraid, scared.'

In order to convey this meaning, an inferential is generally preferable:

(94) གླིང་མི་འདུག

KHONG ZHED-BZHAG

3SG afraid-PERF+INF+SENS

'He is afraid, scared.'

[-Obs] verbs do not always coincide in the various Tibetic languages. For example, in Ü-Tsang, the verb 'KHYAG' to feel cold, freeze conveys an inner sensation whereas the adjective 'GRANG,MO' means 'to be cold' (outside). In Common Tibetan, for example, it is absolutely acceptable to say:

(95) གླིང་མི་འདུག

PHYI LOGS-LA GRANG,MO 'DUG YIN.NA'I NGA 'KHYAG-GI-MI.'DUG

outside.LOC cold EXV but 1SG cold-STAT+NEG+ENDO

'It is cold outside, but I am not cold.'
In Amdo, the situation is the opposite. The verb *GRANG* is [-Obs] and used for “inner sensation” whereas the verb *KHYAG* means ‘to be cold’ (outside).

Note that in some languages, the sensory marker and the endopathic sensory marker are formally identical (but functionally distinct). This is the case for example of the marker above marker *DUG*. In some languages, which have non-visual sensory markers, the endopathic and the external sensory are formally distinct. See below the section 8.4.2.

8.3.6.4. **Valency**

In many World languages, following the European tradition, one can distinguish between transitive verbs, which have a direct object, and intransitive verbs which are deprived of a direct object. However, specialists of linguistic typology do not recognize these categories as universal. In many Asian languages these categories are not very helpful because no argument is compulsory. In the Tibetic languages, a verb can always stand alone without any complement. In other words, arguments of the verb may always be omitted if the context is clear enough. This characteristic is known as “lability” (see Creissels 2014). Since verbs in the Tibetic languages are labile, the notion of transitivity is somewhat problematic.

The notion of valency is more appropriate for these languages and we can distinguish, on a semantic basis, monovalent, bivalent and trivalent verbs. This approach has been used by various authors such as Zeisler (2004, 2007, 2012a), Haller (2007), Tournadre (1996a) and Simon (2016).

Virtually all the Tibetic languages exhibit ergative constructions, i.e. the grammatical patient *P* of a bivalent verb is marked in the same way as the grammatical single argument *S* of a monovalent verb by the absolutive case (zero marked), whereas the grammatical agent *A* is marked by a specific case, the ergative.

---

47. More technically speaking: this lability can be defined as "Argument nonmodifying weak lability" (see Creissels 2014).

48. In some cases, even quadrivalent verbs.
The lama ate meat.’ (Ü, ComTib)

The meat is rotten.’ (Ü, ComTib)

The word शर 'meat’ which is the patient (P) in the first sentence is marked in the same way (with the absolutive) as in the second sentence where it functions as a single argument (S). The word त्स་ BLAMA which corresponds to the agent (A) is marked by the ergative case (BLAMA becomes BLA-MA-Š). For the single argument (S), we need to distinguish between S<sub>A</sub> and S<sub>P</sub>. The S<sub>A</sub> corresponds to an agentive single argument of a controllable verb (see 8.3.6.2). It is usually marked by the absolutive case but in some languages (e.g. Ü, Ts, but not Amdo), in the case of a strong emphasis, it may optionally be marked by an ergative case (see 8.1.9). The S<sub>P</sub> corresponds to a patientive single argument of a non-controllable verb.

Ergativity in the Tibetic languages differs from some other types of ergativity which are more driven by syntax such as Basque (see Laka 1996) or Kartuli (Georgian) (Aronson 1982; Hewitt 1995).

In most Tibetic languages, ergativity depends on semantic and pragmatic factors. There is generally a certain degree of optionality (see Zeisler 2012a; Tournadre 1991, 2010; DeLancey 2011a; Simon 2011, 2016; Suzuki 2014a).

This characteristic is not specific to the ergative and applies to other cases, particularly to the dative. The use of the ergative or the dative in some cases usually indicates a contrastive focus on the given argument (pragmatic or discursive value). Additionally the use of the dative (instead of the absolutive) may also indicate that the patient P is partially affected. Here are some examples in Common Tibetan (98-99) and Ladaks (100-101):
(98) མིང་སྡོང་ ’དི་ མེ་ཤོར་བཞག
SHING.SDONG ’DI ME SHOR-BZHAG
tree DEM+ABS catch fire-PERF+INF+SENS
‘The tree has caught on fire.’ [nothing is left from it] (Simon 2011)

(99) མིང་སྡོང་ ’དིར་ མེ་ཤོར་བཞག
SHING.SDONG ’DI-R ME SHOR-BZHAG
tree DEM-DAT catch fire-PERF+INF+SENS
‘The tree has caught on fire.’ [but it did not die] (ibid.)

(100) མིང་སྡོང་སྒོ་ རྡུངས།
KHO-S SGO RDUNGS
3SG-ERG door beat+PST
‘S/he knocked at the door.’ [in order to get in] (adapted from Zeisler 2007)

(101) མིང་སྡོང་སྒོ་ ’ད་ རྡུངས།
◊ KHO-S SGO-A RDUNGS
3SG-ERG door-dat beat+PST
‘S/he beat the door.’ [accidently or in order to make trouble] (ibid.)

Here are two sentences in Common Tibetan illustrating the optionality of the ergative:

(102) ཉ་ བཟོ་གི་ཡོད།
NGA JA BZO-GLYOD
1SG tea prepare+PRS
‘I prepare tea.’ (adapted from Tournadre & Sangda Dorje 2003)

(103) ཉ་ བཟོ་གི་ཡོད།
NGA-S JA BZO-GLYOD
1SG-ERG tea+ABS prepare+PRS
‘I prepare tea.’ [usually] (ibid.)
The degree of optionality of the ergative differs in the various languages. In some languages such as Amdo, the ergative is more prevalent and grammaticalized and optionality is less present whereas in some languages (Southern Kham, Drugchu and Zanhar), ergative is highly pragmatic and does not appear in unmarked sentences. In a third type, ergative marking may be compulsory with some tenses and aspects particularly (with the completed past, see 8.4.2 on tense and aspect) but optionality is found with other tenses or aspects (Common Tibetan, Dzongkha).

The following formulas illustrate the main syntactic patterns of grammatical arguments found in many languages of the Tibetic area.

**Monovalent verbs**

(a) S\text{-}ABS \quad V \\
(b) S\text{-}ABS (/ERG) \quad V

**Bivalent verbs**

*Action verbs*

(c) A-(ERG) \quad P-ABS \quad V \\
(d) A-(ERG) \quad P-DAT \quad V \\
(e) A-(ERG) \quad P-DAT/ABS \quad V \\
(f) A-(ERG) \quad P-DAT/ASS \quad V \\
(g) A-(ERG) \quad P-ASS \quad V

*Reception verbs*

(h) R-(DAT) \quad P-ABS \quad V

*Emotion verbs*

(i) S-ABS \quad P-DAT \quad V

*Association verbs*

(j) S-ABS \quad P-ASS \quad V \\
(k) S-(ERG) \quad P-ASS \quad V
Trivalent verbs

(l) A-(ERG) B-DAT  P-ABS  V
(m) A-(ERG) P-ABS  B-ASS  V

While the 4 verb classes (action, emotion, reception, and association) are found in several Tibetic languages such as Ü, Tsang, Ladaks, and Amdo, some languages such as Dzongkha, have only two major classes: action and reception verbs. For example, emotion verbs behave as action verbs and the "subject" may be marked by the ergative. Compare below Dzongkha (104) and Ü (105):

(104) རོལ་མི་ཁོ་  རུལ་  དགའ།
KHO-GIS MO-LU DGA'
3SG(M)-ERG 3SG(F)-DAT love
‘He loves her.’ (Dz, van Driem 1998: 194)

(105) རོལ་  རུལ་  དགའ་གི་ཡོད་རེད།
KHO MO-LA DGA'-GL.YOD.RED
3SG(M) 3SG(F)-DAT love-UNCMP+FACT
‘He loves her.’ (Ü, ComTib)

In some languages such as Kham (Sn, Cp, etc.), the ergative is optional and has only a pragmatic function.

nJol Tibetan:

(106) འོ་  བཟོ།
NGA JA BZO
1SG tea prepare
‘I prepare tea.’

An ergative marking of the agent is acceptable, but the neutral form is the absolutive.

(107) འོ་  རོལ་  འདོད་མ།
NGA KHO GDAMS
1SG 3SG love
‘I love him.’
An ergative marking of the agent is non-acceptable. Note that the object of the verb 'to love' does not require a dative marking.

8.3.7. Verbal composition

While many verbs are simple monosyllabic verbs, one finds in CT and in the modern Tibetic languages a lot of compound verbs which may be divided into 4 types: (a) the noun-verb compounds, (b) the noun incorporating verbs and (c) the light verb constructions, (d) lexicalized serial verbs.

8.3.7.1. Noun-verb compounds

In CT, a relatively small number of verbs are compound verbs which usually involve a noun followed by a verb. Most of these verbs are still found in the modern languages. They include the following lexical items:

རག་ལས་ \textit{RAG LAS} ‘to depend upon’, ཡིད་ཆེས་ \textit{YID CHES} ‘to trust’, མདོན་ཁྱེན་ \textit{KHONGS GTOGS} ‘to belong’, རག་པོ་ \textit{RAG BI} ‘to be drunk’, རླ་བོ་ \textit{YA MTSHAN} ‘to be astonished’, རུང་ \textit{HA GO} ‘to understand’, ཡིད་ཆེས་ \textit{THAG BCAD} ‘to decide’, མགོ་སྐོར་ \textit{MGO SKOR} ‘to deceive’, ར་བཟི་ \textit{RA BZI} ‘to obey’, etc.

In some cases, the meanings of the noun and verb are transparent: ར་བཟི་ \textit{THAG BCAD} ‘to decide’ literally means to ‘cut (the) rope’. མགོ་སྐོར་ \textit{MGO SKOR} ‘to deceive’ may be easily analyzed as ‘to turn (somebody’s) head’, ར་བཟི་ \textit{KHA-LA NYAN} ‘to obey’ as ‘listen to the mouth’. However, in other cases, the elements of the composition are no longer meaningful in synchrony. ར་བཟི་ \textit{RABZI} (Ü, Kh, Am, etc.) or རིང་ \textit{RAROS} ‘to be drunk’ (La, Kh: Sn) is no longer analyzed in most languages, however in some southern Kham dialects /ra/ is interpreted as རག་ \textit{RAG} ‘alcohol’ by speakers.

8.3.7.2. Noun incorporating verbs

A type of noun incorporation is a fairly common and original strategy attested in CT and some Tibetic languages.\textsuperscript{49} The compound verb is made of a verb which conveys the main lexical information and a generic noun. In many cases, the noun is derived

\textsuperscript{49} From a typological point of view there are several types of noun incorporation and the precise definition of noun incorporation is still debated. See e.g. Mithun 2009; Jacques 2012b. It is interesting to note that, according to Jacques (ibid.), Japhug, a rGyalrongic language spoken in Sichuan (see 10.7.6), is a polysynthetic language with “incorporation-like constructions” (Jacques, ibid.).
from the same verbal root – a case of figura etymologica – by means of a formative D or S or a nominalizing suffix MA, BA or both:

རྩེད་མོ་རྩེ་ RTSED.MO RTSE ’to play (a game)

ཁ་ལག་ཟ་ BA KHA ZA ’to eat (food)

ཟ་མ་ཟ་ MA ZA ’to eat (food)

ལུད་པ་ལུ་ PA LUD LU ’to cough

གཉིད་ཉལ་ GNYID NYAL ’to sleep

In other cases, the lexical noun may be not related to the verb root but also indicates a generic meaning:

གྲོད་ཁོག་ལྟོགས་ GROD.KHOG.LTOGS ’to be hungry’

ཁ་ལག་ཟ་ BA KHA ZA ’to eat (food)

ཟྣ་མོ་ཟྣ་ MA ZA ’to eat (food)

ལུད་པ་ལུ་ PA LUD LU ’to cough

གཉིད་ཉལ་ GNYID NYAL ’to sleep

Some of these verb constructions are preserved in the modern languages particularly in the eastern regions of Kham and Amdo.

Kham: ZAMA ’to eat (food)

In other cases, the lexical noun may be not related to the verb root but also indicates a generic meaning:

གྲོད་ཁོག་ལྟོགས་ GROD.KHOG.LTOGS ’to be hungry’

ཁ་ལག་ཟ་ BA KHA ZA ’to eat (food)

ཟྣ་མོ་ཟྣ་ MA ZA ’to eat (food)

ལུད་པ་ལུ་ PA LUD LU ’to cough

གཉིད་ཉལ་ GNYID NYAL ’to sleep

Some of these verb constructions are preserved in the modern languages particularly in the eastern regions of Kham and Amdo.

Kham: ZAMA ’to eat (food)

In other cases, the lexical noun may be not related to the verb root but also indicates a generic meaning:

གྲོད་ཁོག་ལྟོགས་ GROD.KHOG.LTOGS ’to be hungry’

ཁ་ལག་ཟ་ BA KHA ZA ’to eat (food)

ཟྣ་མོ་ཟྣ་ MA ZA ’to eat (food)

ལུད་པ་ལུ་ PA LUD LU ’to cough

གཉིད་ཉལ་ GNYID NYAL ’to sleep

Some of these verb constructions are preserved in the modern languages particularly in the eastern regions of Kham and Amdo.

Kham: ZAMA ’to eat (food)

In other cases, the lexical noun may be not related to the verb root but also indicates a generic meaning:

གྲོད་ཁོག་ལྟོགས་ GROD.KHOG.LTOGS ’to be hungry’

ཁ་ལག་ཟ་ BA KHA ZA ’to eat (food)

ཟྣ་མོ་ཟྣ་ MA ZA ’to eat (food)

ལུད་པ་ལུ་ PA LUD LU ’to cough

གཉིད་ཉལ་ GNYID NYAL ’to sleep

Some of these verb constructions are preserved in the modern languages particularly in the eastern regions of Kham and Amdo.

Kham: ZAMA ’to eat (food)

In other cases, the lexical noun may be not related to the verb root but also indicates a generic meaning:

གྲོད་ཁོག་ལྟོགས་ GROD.KHOG.LTOGS ’to be hungry’

ཁ་ལག་ཟ་ BA KHA ZA ’to eat (food)

ཟྣ་མོ་ཟྣ་ MA ZA ’to eat (food)

ལུད་པ་ལུ་ PA LUD LU ’to cough

གཉིད་ཉལ་ GNYID NYAL ’to sleep

Some of these verb constructions are preserved in the modern languages particularly in the eastern regions of Kham and Amdo.

Kham: ZAMA ’to eat (food)

In other cases, the lexical noun may be not related to the verb root but also indicates a generic meaning:

གྲོད་ཁོག་ལྟོགས་ GROD.KHOG.LTOGS ’to be hungry’

ཁ་ལག་ཟ་ BA KHA ZA ’to eat (food)

ཟྣ་མོ་ཟྣ་ MA ZA ’to eat (food)

ལུད་པ་ལུ་ PA LUD LU ’to cough

གཉིད་ཉལ་ GNYID NYAL ’to sleep

Some of these verb constructions are preserved in the modern languages particularly in the eastern regions of Kham and Amdo.

Kham: ZAMA ’to eat (food)

In other cases, the lexical noun may be not related to the verb root but also indicates a generic meaning:

གྲོད་ཁོག་ལྟོགས་ GROD.KHOG.LTOGS ’to be hungry’

ཁ་ལག་ཟ་ BA KHA ZA ’to eat (food)

ཟྣ་མོ་ཟྣ་ MA ZA ’to eat (food)

ལུད་པ་ལུ་ PA LUD LU ’to cough

གཉིད་ཉལ་ GNYID NYAL ’to sleep

Some of these verb constructions are preserved in the modern languages particularly in the eastern regions of Kham and Amdo.

Kham: ZAMA ’to eat (food)

In other cases, the lexical noun may be not related to the verb root but also indicates a generic meaning:

གྲོད་ཁོག་ལྟོགས་ GROD.KHOG.LTOGS ’to be hungry’

ཁ་ལག་ཟ་ BA KHA ZA ’to eat (food)

ཟྣ་མོ་ཟྣ་ MA ZA ’to eat (food)

ལུད་པ་ལུ་ PA LUD LU ’to cough

གཉིད་ཉལ་ GNYID NYAL ’to sleep

Some of these verb constructions are preserved in the modern languages particularly in the eastern regions of Kham and Amdo.

Kham: ZAMA ’to eat (food)

In other cases, the lexical noun may be not related to the verb root but also indicates a generic meaning:

གྲོད་ཁོག་ལྟོགས་ GROD.KHOG.LTOGS ’to be hungry’

ཁ་ལག་ཟ་ BA KHA ZA ’to eat (food)

ཟྣ་མོ་ཟྣ་ MA ZA ’to eat (food)

ལུད་པ་ལུ་ PA LUD LU ’to cough

གཉིད་ཉལ་ GNYID NYAL ’to sleep

Some of these verb constructions are preserved in the modern languages particularly in the eastern regions of Kham and Amdo.
This old strategy of noun incorporation was later replaced by another strategy, the light verb constructions, which is now quite widespread.

8.3.7.3. Light verb constructions

Light verb constructions (hence LVC) correspond in a sense to the reverse situation of the noun incorporation. In the case of LVC, the main lexical load is conveyed by the noun and the verb is essentially generic.

The LVC strategy is less developed in Classical Tibetan, but it is widely used in most Tibetic languages. It is interesting to note that the LVC strategy is also widespread in some neighboring Indo-Iranian languages such as Hindi, Nepali or Persian.

In CT, only a couple of verbs function as light verbs: འདེབས་/ བཏབ་ ’DEBS/BTAB ’to plant’, འེན་/ བླུངས་ ’LEN/BLANGS ’to take’ but their productivity is limited: རྣམས་སུ་ ’to practice’, ལུ་ ’to sing’, དཔེ་ ’to copy, imitate’, གོ་ ’to understand’, གསལ་ ’to remind, make clear’, གོ་ ’to pray’. 

In the modern Tibetic languages, a dozen of light verbs derived from CT lexical verbs are found. These verbs are extremely productive and form hundreds of compound verbs. It is interesting to note that different light verbs are used in different languages. Sometimes, within the same language, two or even three light verbs may be used with the same noun to indicate an identical meaning.

The light verbs come from the following CT lexical verbs:
Each language will select two or three verbs as the main light verbs. Among these forms, we find basically three verbs that are widely used across the Tibetic languages: 

- **བརྒྱབ་ (BRGYAB)** (with the variant རྐྱབ་ (RKYAB) in Dzongkha),
- **བཏང་ (BTANG)** (with the variant གཏོང་ (GTONG) in Len/BLangs)
- **བྱེད་ (BYED)** (with the variant ཡེད་ (YED) in Amdo) or **བགྱིད་ (BGYID)**.

To these verbs, one must add some honorific verbs as 

- **གནང་ (GNANG)** 'to give, make, do, H. of BYED',
- **བཏང་གནང་ (BTANG-GNANG)** 'to give, make, do, H. of BTANG',
- **བསྐྱོན་ (BSKYON)** 'hit, H. of BRGYAB',
- **ཞུ་ (ZHU)** 'to say, tell, take, eat, drink (humilific)',
- **བཞེས་ (BZHES)** 'to take, eat, drink, H',
- **སྩལ་ (STSAL)** (La),
- **མཛད་ (MDZAD)** 'to do, H (La).

The light verbs are also used to create pairs of controllable vs. non-controllable verbs (see 8.3.6.2) in the various languages. See for example the opposition in Sherpa (in the list below) between: 

- **འཇིགས་བ་ལང་ (JIGS.BA LANG)** 'to fear'; 
- **དགོད་བྲ་གྱི་ (DGOD.BRA GYI)** 'to laugh' (controllable), 
- **འཇིགས་བ་ལང་ (JIGS.BA LANG)** 'to fear'; 
- **མགོ་སྐོར་ཐེབས་ (MGO SKOR THEBS)** 'to get deceived' (non-controllable).

Here are some examples of LVC in the modern languages:

**In Sherpa:**

- **< RGYAG/BRGYAB:** 0 གྱི་ Kb.'Thab Rgyag' to quarrel, 0 གནམ་ Kb.'Rgyag' to snow, རྗུ་ Kb.'Brug Rgyag' to thunder; < **BGYID:** 0 ཀྱང་ Kb.'La'ka Gyi' to work', གྲ་ Kb.'Grig Gyi' to prepare', གྲགས་ Kb.'Dgod.bra Gyi' to laugh', < **LANG:** 0 གྲགས་ Kb.'Dgod.bra Lang' 'to laugh', གྲགས་ Kb.'Jigs.Ba Lang' 'to fear', < **SHOR:** གྲགས་ Kb.'Dgod.bra Shor' 'to burst out in laughter'; < **LEN:** གྲགས་ Kb.'Gluling' 'to sing'.

In Dzongkha:

< RGYAG/BRGYAB: རིགས་སྐྱེལ་ MDA’RKYAB ‘to shoot arrows’, རིགས་སྐྱེལ་ Shob RKYAB ‘to lie’, རིགས་སྐྱེལ་ GRASGRIG RKYAB ‘to prepare’, རིགས་སྐྱེལ་ MGO RSHA’RKYAB ‘to share, divide’, རིགས་སྐྱེལ་ BLO’O RKYAB ‘to have a conversation’;
< BYED: བ་འབད་ LA’BAD ‘to work’; < BTANG: སྦྱིན་པ་བཏང་ TSHONBTANG ‘to give alms, practice philanthropy’, དགོ་བཤའ་ RKYAB ‘to share, divide’, བློའོ་ RKYAB ‘to have a conversation’;

In Amdo:

< BYED: བརྒྱད་ NGO BYED /shawa ye/ ‘to work, to rest’; < RGYAG/BRGYAB: རིགས་སྐྱེལ་ SGO RGYAG ‘to close the door’, རིགས་སྐྱེལ་ SGOM RGYAG ‘to meditate’, རིགས་སྐྱེལ་ /njo’e/ ‘to do shopping’;
< DEBS: རིགས་སྐྱེལ་ HAR DEBS ‘to grunt (yak’s sound)’, རིགས་སྐྱེལ་ HE DEBS ‘to cry’;
< BTANG/GTONG: རིགས་སྐྱེལ་ TSHONBTANG ‘to give alms, practice philanthropy’, རིགས་སྐྱེལ་ SGO NGATONG ‘to lay eggs’;
< THEN: རིགས་སྐྱེལ་ BSTOD.PA THEN ‘to sing a praise’, རིགས་སྐྱེལ་ CHU THEN ‘to draw (water)’, རིགས་སྐྱེལ་ SNGUR.PA THEN ‘to snore’, etc.

In Kham:

< BYED: རིགས་སྐྱེལ་ SLOB.SBYONG BYED ‘to study’, རིགས་སྐྱེལ་ NGASG.SGO BYED /nge/ ‘to rest’; < RGYAG/BRGYAB: རིགས་སྐྱེལ་ SKAD RGYEB ‘to shout’, རིགས་སྐྱེལ་ TSHONRGYEB ‘to do business, trade’;
< BTAB: རིགས་སྐྱེལ་ PAR BTAB ‘to photograph’, རིགས་སྐྱེལ་ TSHWA BTAB ‘to salt’;
< THEN: རིགས་སྐྱེལ་ GZHAS THEN ‘to sing’, རིགས་སྐྱེལ་ DUBA THEN ‘to smoke’;
< GUL: རིགས་སྐྱེལ་ LASKA GUL /läka/ ‘to work’ (Gyalthang).

In Balti:

< BYED/BYA: རིགས་སྐྱེལ་ LAS BYA ‘to work’, རིགས་སྐྱེལ་ BAGSTON BYA /bakston bya/ ‘to marry’, རིགས་སྐྱེལ་ BAR BYA ‘to separate’, རིགས་སྐྱེལ་ GSAMPYA BYA ‘to think, meditate’, རིགས་སྐྱེལ་ CHESLUGS BYA /česluks bya/ ‘to believe’, རིགས་སྐྱེལ་ DGA’LUGS BYA /galuks bya/ ‘to love’;
< BTANG: རིགས་སྐྱེལ་ MCHLMA BTANG /čima tang/ ‘to weep, to cry’, རིགས་སྐྱེལ་ CHUBTANG ‘to water’, རིགས་སྐྱེལ་ GOMPA BTANG ‘step out’, རིགས་སྐྱེལ་ GRONBTANG /gron

In Ladaks:


In Common Tibetan:

< BYED: བསྟན་པོ། LAS.KA BYED ‘to work’, བེད་པོ། BED.SPYOD ‘to use’; < BTANG: བསྟན་པོ། GZHAS.BTANG ‘to sing’, བསྟན་པོ། YAR.RGYAS.BTANG ‘to improve’, བསྟན་པོ། MGO.SKOR.BTANG ‘to deceive’, བསྟན་པོ། MGO.SKOR.BTANG ‘to dream’, བསྟན་པོ། MGO.SKOR.BTANG ‘to rain’, བསྟན་པོ། MGO.SKOR.BTANG ‘to reflect upon’, བསྟན་པོ། MGO.SKOR.BTANG ‘to have sexual intercourse before marriage’ (Goldstein 2001); < BRGYAB: བསྟན་པོ། CHANG.SABRGYAB ‘to marry’, བསྟན་པོ། CHAR.PABRGYAB ‘to rain’, བསྟན་པོ། ME.MDA.SBRGYAB ‘to shoot (fire arm)’; < THEBS: བསྟན་པོ། MGO.SKOR.THEBS ‘to get deceived’, བསྟན་པོ། MGO.SKOR.THEBS ‘to get shot (weapon)’; < GRO: བསྟན་པོ། YAR.RGYAS.GRO ‘to get improved’, < BSKYON: བསྟན་པོ། ZHAL.LAG.ZHES ‘to eat a meal (H)’; < GNANG: བསྟན་པོ། PHYAG.LAS GNANG ‘to work’ (H), བསྟན་པོ། GZHAS.BTANG-GNANG ‘to sing’ (H).

8.3.8. Inflectional morphology

Most Tibetic languages have preserved traces of inflectional forms derived from CT to convey the TAM. The only major exception are Dzongkha, Lhoke and southern Kham languages. In some languages, particularly Amdo and northern Kham, reflexes of the irregular classical forms are still well attested. In most other areas (Û, Ts, Ba, Pur, La, Sh, etc.) we find only traces of the Classical morphology that reflects regular forms.
Let’s make a preliminary remark about the inflections: even in the languages, where Classical inflections have been relatively preserved, they play only a secondary role, because tenses, aspects, moods, and evidentiality are mainly indicated by verb auxiliaries or suffixes (see above 8.3).

In CT, as shown in 6.6.3, for certain verbs we find four different forms corresponding to the “past”, “present”, “future”, and “imperative” for controllable verbs and only three forms for non-controllable verbs “past”, “present”, and “future”. From a morphological point of view these inflections in the modern languages correspond to variations of the initial consonants, vowel or the final ʌ, as in CT.

We use here the traditional terminology to label the four potential forms. However, as pointed out by some authors such as Zeisler (2004), Zemp (2014, 2016), and Hoshi (2016), the four stems do not convey purely the tense and mood meanings corresponding to the labels “present, past, future and imperative”. They convey aspectual and relative tenses as well as other modal meaning such as “potentialis”. Zeisler (2004: 260) suggested labeling “imperative” the imperative stem when it used as a command and otherwise “potentialis.” She also suggested that the future stem could be called more accurately “posterior/purposive stem.” The present stem should be termed “simultaneous/posterior” stem and the past stem “anterior stem” (Zeisler 2004: 267). We generally agree with this analysis which applies primarily to Old Tibetan and to a lesser extent to Classical Tibetan. The choice to analyze the four stems in terms of relative tenses rather than in term of absolute tenses is certainly justified. However, the notions of “completed” versus “uncompleted” aspect could also be relevant in certain cases. According to Hoshi’s study on the RGYAL, RABGNAI, BATME, LONG (2016: 88-89) the notion “complete” vs. “uncompleted” seems to be more adequate in Classical Tibetan. She employs the terms “uncompleted agentive” and “uncompleted undergoing” instead of the traditional “present” and “future” respectively. For example, the past habitual is often expressed by the “present stem” which cannot in that case be analyzed as a “simultaneous (relative) tense” and must be accounted in terms of “uncompleted aspect.”

Despite the fact that the meaning of the stem is related to relative tenses and aspects in Old and Classical Tibetan, we still maintain the traditional terminology
(present, past, future, and imperative), because in the modern languages, the stems are essentially a vestige of an ancient morphology and the tenses, aspects, modality, and evidentiality are conveyed by various auxiliaries.

The languages which have maintained traces of the archaic verb morphology have not preserved a specific form for the future and use the present form instead. Thus, we essentially find three forms for the controllable verbs “present-future”, “past”, and “imperative” and only two forms for the non-controllable verbs (“present-future” and “past”). However, systems with four forms for controllable verbs are still marginally encountered in some archaic Amdo dialects (near the Gyärlong area) and in Khyungpo (Kham). But even in these languages the fourfold system concerns less than ten verbs.

Here are some examples of the stem variation in various Tibetic languages.

<table>
<thead>
<tr>
<th>Chart VIII.4. – Stem variations in Common Tibetan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Past</strong></td>
</tr>
<tr>
<td>‘to look at’</td>
</tr>
<tr>
<td>‘to eat’</td>
</tr>
<tr>
<td>‘to make, do’</td>
</tr>
<tr>
<td>‘to lie down, sleep’</td>
</tr>
<tr>
<td>‘to tell (hum)’</td>
</tr>
</tbody>
</table>
### CHART VIII.5. – Stem variations in Chabcha Amdo

<table>
<thead>
<tr>
<th>Past</th>
<th>Present</th>
<th>Future</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>'to look at'</td>
<td>བལྟས་ BLTAS /'ti/</td>
<td>བལྟ་ LTA /'ta/</td>
<td>བལྟ་ BLTA /'ta/</td>
</tr>
<tr>
<td>'to eat'</td>
<td>ཞེས་ BZAS /'si/</td>
<td>ཞེས་ BZA /'sa/</td>
<td>ཞེས་ BZA /'sa/</td>
</tr>
<tr>
<td>'to make, do'</td>
<td>ལངས་ LANGS /lang/</td>
<td>ལངས་ LANG /lang/</td>
<td>ལངས་ LANG /lang/</td>
</tr>
<tr>
<td>'to teach'</td>
<td>བསླབས་ BSLABS /'tsaw/ or /'lav/</td>
<td>སློབ་ SLOB /'tsav/ or /'l'aw/</td>
<td>སློབ་ SLOB /'tsav/ or /'l'aw/</td>
</tr>
<tr>
<td>'to write'</td>
<td>བྲིས་ BRIS /'t/</td>
<td>བྲི་ BRI /'d/</td>
<td>བྲི་ BRI /'d/</td>
</tr>
</tbody>
</table>

(The examples above are adapted from Robin & Simon forthcoming.)

### CHART VIII.6. – Stem variations in Čone ('Oggangdruk')

<table>
<thead>
<tr>
<th>Past</th>
<th>Present</th>
<th>Future</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>'to look at'</td>
<td>བལྟས་ BLTAS /'tä:/</td>
<td>བལྟ་ LTA /'ta/</td>
<td>བལྟ་ BLTA /'ta/</td>
</tr>
<tr>
<td>'to blow'</td>
<td>བུད་ BUD /'hpi:/</td>
<td>བུད་ BUD /'hpi:/</td>
<td>བུད་ BUD /'hpi:/</td>
</tr>
<tr>
<td>'to buy'</td>
<td>ཉོས་ NYOS /'ne:/</td>
<td>ཉོྦ་ NYO /'no/</td>
<td>ཉོ ཉོ ཉོ /'no/</td>
</tr>
<tr>
<td>'to make, do'</td>
<td>བྱས་ BYAS /'shi:/</td>
<td>བྱེད་ BYED /'shi/</td>
<td>བྱེད་ BYED /'shi/</td>
</tr>
</tbody>
</table>
### Chart VIII.7. Stem variations in Sharkhok (Astong)

<table>
<thead>
<tr>
<th></th>
<th>Past</th>
<th>Present</th>
<th>Future</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>'to come'</td>
<td>སྐོད་ /'ongs</td>
<td>སྐོད་ /'ong</td>
<td>སྐོད་ /'ong</td>
<td>སྐོད་ /'shog</td>
</tr>
<tr>
<td>'to drink'</td>
<td>སྤྲུལ་ /'thung</td>
<td>སྤྲུལ་ /'thung</td>
<td>སྤྲུལ་ /'thung</td>
<td>སྤྲུལ་ /'thung</td>
</tr>
</tbody>
</table>

### Chart VIII.8. Stem variations in Khyungpo (Thromtshang)

<table>
<thead>
<tr>
<th></th>
<th>Past</th>
<th>Present</th>
<th>Future</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>'to blow'</td>
<td>འབུད་ /'bud</td>
<td>འབུད་ /'bud</td>
<td>འབུད་ /'bud</td>
<td>འབུད་ /'phud</td>
</tr>
<tr>
<td>'to kill'</td>
<td>གསོད་ /'gsod</td>
<td>གསོད་ /'gsod</td>
<td>གསོད་ /'gsod</td>
<td>གསོད་ /'gsod</td>
</tr>
<tr>
<td>'to do, make'</td>
<td>བྱེད་ /'byas</td>
<td>བྱེད་ /'byed</td>
<td>བྱེད་ /'bya</td>
<td>བྱེད་ /'byos</td>
</tr>
<tr>
<td>'to lie down'</td>
<td>ཉལ་ /'nyal</td>
<td>ཉལ་ /'nyal</td>
<td>ཉལ་ /'nyal</td>
<td>ཉལ་ /'nyal</td>
</tr>
</tbody>
</table>

### Chart VIII.9. Stem variations in Northern Kham (Derge)

<table>
<thead>
<tr>
<th></th>
<th>Past</th>
<th>Present</th>
<th>Future</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>'to look at'</td>
<td>ཟློམ་ /'bltas</td>
<td>ཟློམ་ /'lta</td>
<td>ཟློམ་ /'lta</td>
<td>ཟློམ་ /'ltos</td>
</tr>
<tr>
<td>'to eat'</td>
<td>བྲེག་ /'bza</td>
<td>བྲེག་ /'za</td>
<td>བྲེག་ /'za</td>
<td>བྲེག་ /'zo</td>
</tr>
<tr>
<td>'to buy'</td>
<td>བྲོག་ /'nyos</td>
<td>བྲོག་ /'nyo</td>
<td>བྲོག་ /'nyo</td>
<td>བྲོག་ /'nyos</td>
</tr>
<tr>
<td>'to take'</td>
<td>བྲོག་ /'blangs</td>
<td>བྲོག་ /'len</td>
<td>བྲོག་ /'blang</td>
<td>བྲོག་ /'blong</td>
</tr>
</tbody>
</table>
#### Chart VIII.10. – Stem variations in Ladak (Shamskat)

<table>
<thead>
<tr>
<th>Past</th>
<th>Present</th>
<th>Future</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘to look at’</td>
<td>རལྟ་ /ltas/</td>
<td>རལྟ /lta/</td>
<td>རལྟ /lta/</td>
</tr>
<tr>
<td>‘to eat’</td>
<td>རབས /ras/</td>
<td>རི་ /za/</td>
<td>རལྟ /lts/</td>
</tr>
<tr>
<td>‘to sniff, smell’</td>
<td>ར་ /ltas/</td>
<td>ར་ /lta/</td>
<td>ར་ /lta/</td>
</tr>
<tr>
<td>‘to get up’</td>
<td>རབཞེངས /zhang/</td>
<td>རབཞེངས /zhang/</td>
<td>རབཞེངས /zhong/</td>
</tr>
<tr>
<td>‘to kill’</td>
<td>རསད /sats/</td>
<td>རསད /sats/</td>
<td>རསད /sots/</td>
</tr>
<tr>
<td>‘to say’</td>
<td>རེར /zer/</td>
<td>རེར /zer/</td>
<td>རེར /zer/</td>
</tr>
</tbody>
</table>

(The examples above are adapted from Zeisler 2004.)

#### Chart VIII.11. – Stem variations in Choča-ngača (Tsamang)

<table>
<thead>
<tr>
<th>Past</th>
<th>Present</th>
<th>Future</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘to eat’</td>
<td>རབས /ras/</td>
<td>རི /za/</td>
<td>རལྟ /lts/</td>
</tr>
<tr>
<td>‘to do, make’</td>
<td>རེབ /byas/</td>
<td>རེབ /bya/</td>
<td>རེབ /bya/</td>
</tr>
<tr>
<td>‘to cry’</td>
<td>རུས /ngus/</td>
<td>རུ /ngu/</td>
<td>རུ /ngu/</td>
</tr>
</tbody>
</table>
In the verb inflection systems preserving the CT forms attested in modern languages, we can find the following types: the language has preserved either three distinctions (past vs. present-future vs. imperative), two distinctions (past-present-future vs. imperative; past vs. present-future-imperative; past-imperative vs. present-future), or no morphological differences. We can also find a type of four distinctions marginally, as in Khyungpo, however, it generally includes innovative stem alternations which are not attested in CT. Sherpa has also four stems for the verb BGYID 'to do', but other verbs only have three stems.

Languages which do exhibit the most complex inflections are found in Amdo and Northern Kham areas. Some of these inflections are clearly inherited from CT or even Old Tibetan, but other represent innovations. That is for example the case of the generalized aspirated initial consonants in Amdo and Khyungpo for the imperative
stem (see above ‘write’). It is also the case of the final /s/ for the past in Ladaks and Balti. As noted by Zeisler (2004: 620): “the suffix –s, which, in comparison to Old Tibetan and Classical Tibetan has been generalized for controlled action verbs and appears even after DNRL [suffixes] in the central and western dialects. Since neither Ablaut nor change of the radical or prefix occur, it might not be adequate to speak of different stems in a strict sense […].”

Finally, as mentioned above, languages of southern Kham, Lhoke and Dzongkha have lost all their inflections (except for suppletive forms) or have only minimal traces such as the final vocalic length (reflex of the final s) attested for some verbs in Choća-ngäča (see the chart above).

One should mention a final historical remark. Some authors doubt that the system of verbal inflectional forms found in CT was actually common to all the dialects. According to Bielmeier (1988, 2003), the verbal prefix B for the “past” is not found in Proto-Tibetan [Proto-Tibetic in our terminology], and never appeared in Balti and Ladaks. He made the following remark:

“[Inflectional forms of the verbs] feature mainly in conservative Amdo Tibetan, but are completely absent in Western Archaic Tibetan. The question is again whether we have to consider them as Common Tibetan or as specific developments of certain areas. The comparative evidence from the spoken varieties again shows that certain inflectional forms of the paradigms in Amdo Tibetan are newly-formed, and suppletion and analogy play important roles in the process of their formation. At least in some cases we can show that this mechanism is also at the root of the formation of the inflectional forms of the paradigms of Written Tibetan, whose forms, therefore, might have to be listed separately.” (Bielmeier 1988)

First one can not agree with the absence of trace in Ladaks, Purik, and Balti. It is true that some characteristics of the CT system are indeed absent in the Northwestern languages such as the alternation of initial consonants and the preradicals conveying present, future, or past tenses. For example, the preradical B, for the past stem, G/D for the future stem as well as the preradicals ’ and Ø for the past stem are not attested in these languages. However, the ablaut is attested for the imperative and it is probable that the archaic irregular morphology gradually disappeared. As suggested by Zeisler (2004: 347): “It is more likely that they have generalized the suffix –s for those verbs
that must have looked quite irregular for them.” Even if we agree with the absence of irregular morphology, it is not correct to say that there are no traces of the archaic morphology. There are a few traces of *B* past tense prefix in Ladaks or Balti. For example, in Old Tibetan, the form of the verb ‘to write’ are ང་ ‘DRI (pres), བྲིས་ ‘BRIS (past), བྲི་ ‘BRI (fut) and རིས་ ‘RIS (imp) (Hill 2005) and the past form BRIS yields the reflex /rbi-/ in Balti (by metathesis), which implies the existence of a prefix *B* in Proto-Balti: *B*-RI. (see Hill 2005). Of course, the general absence of the prefix *B* in the Northwestern Tibetic languages is a problem but a way to account for this is to propose that the prefix disappeared at an early stage. In Ü-Tsang or Dzongkha, which have lost all the prefixes, it is also hard to say when the *B* “past” prefix disappeared.

Recently, Zemp (2014: 129) has again questioned the absence of archaic morphology and provided a lot of evidences against this affirmation: “[…] there is in fact evidence for all the four stems in at least the variety of Purik. The reason why this was not noticed, however, is that the ClT [Classical Tibetan] verbal system had indeed undergone considerable changes after it split up from Proto-Tibetan.”

Finally, it is a worthwhile reminder that in some languages of the South and Southeastern languages, the verb has entirely lost the inherited ancient morphology and in some cases developed innovative alternations, as we will see in the next section.

### 8.3.9. Morphophonemic alternations of the verb

Another phenomenon found in some modern Tibetic languages is the innovative morphophonemic variation of the verb stem which depends on the phonological context. This variation is well attested in some southern languages, particularly Dzongkha (van Driem 1998), Sherpa (see Tournadre et al. 2009), Tö Ngari (see Qu and Tan 1983) and Spiti.

This type of alternation has been described by some authors such as van Driem (1998: 209) as an “inflected form of the verb.” However, this variation is very different from the inflection forms discussed in 8.3.4, since the forms do not indicate various tense-aspects and modalities. There are merely related to the morphophonemic variation of the combination of the lexical verb and the verb ending (or relator+auxiliary). According to van Driem there are four types of “inflected stems” for the “factual

a) “Verb stems ending /ng, n, m/ in Roman Dzongkha … They form their inflected stem by changing the final consonant to m.” For example, *NYAN M IN / nyem-ʔing/ ‘is listening’, *THON M IN / thom-ʔing/ ‘is going out’; *BSD M IN / dam-ʔing/ ‘is attaching, is closing’, *GNANG M IN / nam-ʔing/ ‘is giving’ (H).

b) “Verb stems ending in /p/ in Roman Dzongkha … They undergo no change in their inflected stem.” For example, *RKYAB IN / čap-ʔing/ ‘is making, LV’, *THAB IN / chap-ʔing/ ‘is fighting, is clashing’.

c) “Verb [stem] […] ending in a vowel in Roman Dzongkha” and forming their inflected stem “by adding -u.” For example, *ZAW IN / sau-ʔing/ PRS of ‘is eating’, *URW IN / uː-ʔing/ ‘is rubbing’, *SKYESW IN / keu-ʔing/ PRS of ‘is growing’, *BSHAL W IN / sheu-ʔing/ PRS of ‘is having diarrhea’.

d) “Verb [stem] […] ending in a vowel in Roman Dzongkha” and forming their inflected stem “by adding -p”: for example, *SOODP IN / döp-ʔing/ ‘is staying’, *BTSUGS P IN / tsup-ʔing/ ‘is putting, is inserting’.

If we globally may agree with van Driem’s analysis, it poses two problems for which there is an alternative solution.

First, one can consider that the verb is in fact invariable and that the verb ending (or the relator) is the one which undergoes a variation. In fact, this alternative analysis is supported by the Dzongkha orthography. The verb stem never varies but is followed by endings that undergo morphophonemic variation.

Depending on the final consonant of the verb we find the four following endings for the factual present: *+M IN; *IN; *+W IN; *+P IN.
If we adopt this solution, we can present the above verbs in the following way:

- /∅ nang/ + IN /∅\ nang\ + m,\ IN/, /∅\ ינ\ + IN, /∅\ ינ\ + \ צ\, TON + m,\ IN/, /∅\ ינ\ + \ צ\, BIDAM + IN, /∅\ ינ\ + \ צ\, GNANG + m,\ IN, /∅\ ינ\ + \ צ\, THAB + IN, /∅\ ינ\ + \ צ\, RKYAB + IN.

This presentation matches the Dzongkha orthography. It also matches the diachronic evolution since the four forms /∅\ ינ\ + m,\ IN, /∅\ ינ\ + \ צ\, W\ IN, /∅\ ינ\ + \ צ\, P\ IN are historically derived from the relator /∅\ PA followed by the auxiliary /∅\ YIN. The labial plosive of the relator PA has simply been assimilated to a labial nasal /m/ after verbs ending in a nasal and to a labial glide /w/ after a vowel.

This explanation also matches the modern pronunciation since the verb stem is normally invariable (with one exception) and followed by a variable suffix, corresponding to the verb ending, as shown in the examples above. Let us display the above examples again:

- /∅\ ינ\ + \ צ\, SKYES + w\ IN /∅\ ינ\ + \ צ\, RKYAB + w\ IN /∅\ ינ\ + \ צ\, chap\ Ing/, /∅\ ינ\ + \ צ\, ZA + w\ IN /∅\ ינ\ + \ צ\, b\ IN /∅\ ינ\ + \ צ\, s\ DO\ D\ IN /∅\ ינ\ + m,\ IN.

The only exception that shows an alternation of the verb stem occurs in van Driem's presentation matches the Dzongkha orthography. It also matches the diachronic evolution since the four forms /∅\ ינ\ + m,\ IN, /∅\ ינ\ + \ צ\, W\ IN, /∅\ ינ\ + \ צ\, P\ IN are historically derived from the relator /∅\ PA followed by the auxiliary /∅\ YIN. The labial plosive of the relator PA has simply been assimilated to a labial nasal /m/ after verbs ending in a nasal and to a labial glide /w/ after a vowel.

This explanation also matches the modern pronunciation since the verb stem is normally invariable (with one exception) and followed by a variable suffix, corresponding to the verb ending, as shown in the examples above. Let us display the above examples again:

- /∅\ ינ\ + \ צ\, SKYES + w\ IN /∅\ ינ\ + \ צ\, RKYAB + w\ IN /∅\ ינ\ + \ צ\, chap\ Ing/, /∅\ ינ\ + \ צ\, ZA + w\ IN /∅\ ינ\ + \ צ\, b\ IN /∅\ ינ\ + \ צ\, s\ DO\ D\ IN /∅\ ינ\ + m,\ IN.

The only exception that shows an alternation of the verb stem occurs in van Driem's presentation matches the Dzongkha orthography. It also matches the diachronic evolution since the four forms /∅\ ינ\ + m,\ IN, /∅\ ינ\ + \ צ\, W\ IN, /∅\ ינ\ + \ צ\, P\ IN are historically derived from the relator /∅\ PA followed by the auxiliary /∅\ YIN. The labial plosive of the relator PA has simply been assimilated to a labial nasal /m/ after verbs ending in a nasal and to a labial glide /w/ after a vowel.

The same morphophonemic variation is attested 1,000 kms away in Purik and Balti with the nominalizer /∅\ PA. Here is Zemp's comment about the allomorphs of /∅\ PA:

- /∅\ PA\ ma (after -ng, n, e.g. in [∅\ ינ\ + m,\ IN] /∅\ ינ\ + \ צ\, len\ ma /∅\ ינ\ + \ צ\, len\ ma 'to take'), and
- /∅\ PA\ a (after -b and all vowels except -e, e.g. in [∅\ ינ\ + m,\ IN] /∅\ ינ\ + \ צ\, leb\ a 'to arrive', [∅\ ינ\ + \ צ\, tr\ a\ 'to ask', [∅\ ינ\ + \ צ\, yu\ a\ 'to cry', [∅\ ינ\ + \ צ\, sko\ a\ 'to dig', and [∅\ ינ\ + \ צ\, ba\ a\ 'to do'], with its phonetic variant /∅\ a\ (after -e, e.g. in [∅\ ינ\ + m,\ IN] /∅\ ינ\ + \ צ\, gse\ a\ [gse\ a\ 'to play'] (Zemp 2018).
Let’s turn back to the Dzongkha verb morphophonemics. In van Driem’s analysis, the verb ending in an open vowel may construct their inflected form by adding /u/ or by adding /p/. The author adds: “Which of these two endings an open-stem verb takes is a given which must be committed to memory.”

It seems however that there is a rule that allows predicting the type of variation:

a) After the reflexes of CT final vowels and final consonants R, L and S, the relator is /u/ written བུ ག/u/ within the same syllable as the verb stem in Dzongkha.

b) After the reflexes of CT final D and G, the relator is /p/ written བུ ག/u/ within the same syllable as the verb stem in Dzongkha.

c) After the reflexes of CT final NG and N, the relator is /m/ written བུ ག/u/ within the same syllable as the verb stem in Dzongkha.

d) After the reflexes of CT final B and M, the relator is Ø (unmarked).

Since the Dzongkha orthography is generally inspired by the CT orthography, the rule (a-d) is easy to apply. There seems to be only one salient exception to this rule: the verb བུ ག/u/ ཁལ/’ to do' (and LV) which is extremely frequent. In Dzongkha orthography, the final consonant ends in a D and thus, it should be ending P /p/ but the factual present yields a W /u/: བུ ག/u/ ཁལ/’ to do’ violating the rule b).

However, the spelling of this Dzongkha word does not reflect the original CT orthography. The Dzongkha verb བུ ག/u/ ཁལ/’ to do’ is in fact derived from the CT verb བུ ག/u/’ BYED ‘to do’. This verb is found in neighboring languages such as Choça-ngača (in the east) under the form བུ ག/u/ ཁལ/’ bya/ and in Lhoke (in the west) as བུ ག/u/ ཁལ/’ bya/ and the classical verb བུ ག/u/’ BYED has the inflection form བུ ག/u/’ BYAS in the past which is probably the origin of the contemporary Dzongkha verb ‘to do’. Thus, the final consonant of the verb ‘to do’ is probably originally an S (and not a D as written in modern Dzongkha) which explains why it is followed by the ending W /u/.

If this is correct, we have a much more simple approach whereby all the verb stems are invariable in Dzongkha and the factual present ending has four allomorphs depending on the final consonant or vowel.
The Sherpa verb also displays morphophonemic variations. In this case, we find both a variation of the verb ending (relator + auxiliary) and in some cases of the verb stem. For example ཝ/‘sa/ > ཝ/‘se/ ‘to eat’, འྲ/‘dgo/ > འྲ/‘dgi/ ‘to go’, ག/‘ngu/ > ག/‘ngo/ ‘to cry’, etc. འྲ/‘dgi-se-wiʔ/ ‘S/he will go’, འྲ/‘dgi-ngu-ub/ ‘S/he will cry’.

The verb ending and particularly the relator also undergo variation as shown in the Sherpa-English dictionary elaborated by N. Tournadre et al. (2009). For example the factual suffix /uza/ has various allomorphs depending on the final consonant of the verb and becomes /pza/ or /tuza/., etc.

For example ར/‘hril-mu-lha-zza/ ‘HRIL. MULHA+BDZA/’šilmu ‘Ta;pza/’(s/he)’ll watch the show’; ར/‘ki-tap-tab-te-dza/ ‘Kit.Tab Sbin-DUDZA/’kitap ‘bin-tuza/ ‘S/he gave (someone) the book’, etc.

Morphophonemic variations of the stem are also frequently found in some Tö Ngari dialects and Dolpo (Nepal). The vowel changes under the influence of the relator (see also Qu and Tan 1983).

For example in the Purang dialect: ར/‘bras-zza-glred/ ‘BRAS ZA-GLRED translates as: ར/‘bras-zza-da-d/ ‘BRAS ZA+’ ‘DAD /‘dæ ‘sända/ ‘they eat rice’. Thus /‘sa/ ‘to eat’ becomes /‘sä/. Again, it should be emphasized that the umlaut on the vowel only conveys a morphophonemic variation and does not convey a tense-aspect meaning unlike the inflection forms that we examined in 8.3.4.

8.3.10. Suppletive forms

Tense-aspect and imperative inflections are sometimes achieved by using suppletive forms, i.e. TAM may be marked by entirely different verbal roots. This suppletion strategy is found throughout the Tibetic area (see Bielmeier 2004). It is attested in Ladaks, Balti, Spiti, Sherpa, Amdo, Southern Kham, Ü, Tsang, Dzongkha, etc.

51. The verb ར/‘hril-mu-lha/ ‘HRIL. MULHA is derived from: ར/‘ltad-mo lta. The phonological variation can be explained by the regular Sherpa reflex LT > /l/. The form ར/‘hril-mu/’šilmu/ is an irregular reflex of CT LTAD.MO
Among the verbs that often have a suppletive form, one finds ‘to go’, ‘to come’, ‘to give’, ‘to say’, i.e. verbs which have a high frequency. Here is a list of suppletive verbs with high frequency to mark the past, present-future and imperative.

‘to come’: ཕོང་ YONG (pst and prs), ཕོག་ SHOG (imp) (Ub, Ts)
‘to come’: ཀོང་'ONG (pst and prs), ཕོག་ SHOG (imp) (Kh, Ho)
‘to come’: ཀོང་'ONG (prs), ཅོས་ ONGS (pst), ཕོག་ SHOG (imp) (Sh)
‘to go’: ཅོག་ GRO (prs), ཆུ་ PHYIN (pst), རྒྱ་ RGYUGS (imp) (Ub, Ts)
‘to go’: ཅོག་ GRO (prs), ཆུ་ PHYIN (pst), ཕོང་ SONG (imp) (Kyirong)
‘to go’: ཁོད་ YO (prs), ཤུ་ THAL (pst), ཕོང་ SONG (imp) (Thewo-Tö)
‘to go’: ཁོད་ YO (prs), ཕོང་ SONG (pst), ཤུ་ YAR (imp) (Dz)
‘to go’: ཁོད་ GRO (prs), ཕོང་ SONG (pst, imp) (Kh, Ho)
‘to go’: ཁོད་ GYO (prs), ཕོང་ SONG (pst, imp) (Am)
‘to go’: ཁོད་ CHA (prs), ཕོང་ SONG (imp, past) (La)
‘to go’: ཁོད་ CHA (prs), དྲུ་ GRUL ང་ BUD (pst), ཕོང་ SONG (imp) (Garzha)
‘to go’: ཁོད་ GRO (prs), དབྱས་ BRGAL (pst) (Sh)
‘to give’: ཁོད་ STER (prs, imp), དབྱས་ BYIN (pst) (Am)
‘to give’: ཁོད་ STER (prs, imp), དབྱས་ SBYIN (past) (Sh)
‘to say’: འཛེར་ ZER (prs), བྲོ་ BZLAS (pst) (Am)
‘to sit, stay’: རྡུ་ DUG (prs), རྡུ་ BSDAD (pst) (Am)
‘to do’: བྱི་ GYID (prs), བྱི་ BYAS (pst), བྱི་ GYS (imp) (Cone)
‘to be sick’: བྲུ་ ZUG.GZER (prs), བྲུ་ NA (pst) (Sp)

8.3.11. Causative derivation

Modern Tibetic languages have inherited to a certain extent pairs of causative and anticausative or resultative verbs found in OT and CT. This morphological derivation has been described by various authors such as Beyer (1992), Kesang Gyurmé (1992), Tournadre & Sangda Dorje (1998, 2003), Kretschmar (1995), Tournadre & Konchok Jiarso (2001). There are about two hundred verbal pairs in CT but modern
languages often have preserved about thirty frequent verbs. In southern Kham, this contrast is restricted to a very limited number of verbs. The preserved verbal pairs are not identical in the various languages, but the verbs are nearly always derived from their CT correspondences.

Morphologically, the causative verb was historically derived from the basic anticausative form by prefixing an “s” also found in ST. In CT, the causative “s” often appears as a superscript ‘S’ (see the examples below). The superscript ‘S’ is still heard in the western languages of Ladakh and Baltistan or have left a trace as in Amdo. However, in most other languages causative verbs are distinguished from their anticausative correspondent by tone and/or aspiration. In some cases, the spoken forms no longer make a distinction between the two verbs of the pair.

**Chart VIII.13. – Anticausative and causative verb pairs**

<table>
<thead>
<tr>
<th>Anticausative/ Resultative verbs</th>
<th>Causative verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>གོལ་ KHOL ‘to boil’</td>
<td>ལོལ་ SKOL ‘to boil/make boil’</td>
</tr>
<tr>
<td>གོར་ KHOR ‘to turn around, spin’</td>
<td>རོར་ SKOR ‘to turn around, revolve’</td>
</tr>
<tr>
<td>གེལ་ PHEL ‘to increase’</td>
<td>ལེལ་ SPEL ‘to cause to increase, augment’</td>
</tr>
<tr>
<td>བུར་ NYAL ‘to lie down, to go to sleep’</td>
<td>བུལ་ SNYAL ‘to lay down, to put to sleep’</td>
</tr>
<tr>
<td>གོན་ THON ‘to come out, depart’</td>
<td>བོན་ BTON ‘to cause to come out, put forth’</td>
</tr>
<tr>
<td>ལོག་ LOG ‘to return, come/go back’</td>
<td>ལོག་ SLOG ‘to return, send back’</td>
</tr>
<tr>
<td>དུག་ GUM ‘to die’</td>
<td>དུག་ BKUM ‘to kill’</td>
</tr>
<tr>
<td>དུར་ KHUR ‘to carry a load (on the back)’</td>
<td>དུར་ BSKUR ‘to load a pack (on the back)’</td>
</tr>
<tr>
<td>བིབ་ BAB ‘to descend, go down’</td>
<td>བིབ་ PHAB ‘to bring down, cause to fall’</td>
</tr>
</tbody>
</table>

From a morphological point of view, causative verbs may have up to four inflectional forms (see 8.3.4), whereas anticausative verbs have usually two inflection forms, rarely three.
From a syntactic point of view, causative verbs always have an additional argument corresponding to an intentional agent performing the action. For these reasons some authors such as Beyer (1992) or Kretschmar (1995: 129-130) have called this opposition “transitive” versus “intransitive” verbs instead of “causative” versus “anticausative”. However, since the syntactic notion of transitivity is highly problematic in the Tibetic languages (see 8.3.2.2), we do not use this terminology: the semantic opposition between “monovalent” and “bivalent” would be more appropriate. However, a few “anticausative” verbs are bivalent (or “transitive”) whereas their causative correspondents are trivalent (or “ditransitive”). This is for example the case of verbs such as ‘KHUR ’to carry’ (anticausative, bivalent) vs. BSKUR ’to make someone carry, send’ (causative, trivalent) or GON ’to put on, dress’ vs. BSKON ’to dress someone’.

Semantically, causative verbs are always controllable verbs while most resultative verbs are non-controllable and lack imperative. For this reason, some authors such as Huber (2002) have qualified the causative derivation as an opposition between controllable and non-controllable verbs. There are however a few anticausative/resultative verbs, which are controllable (see the above ex. ‘KHUR’ ‘to carry’ and GON ‘to put on, dress’).

Finally, another characteristic feature of the verbal pairs is that they may co-occur in the same sentence to indicate whether the performed action has or does not have a result. For this reason the verb pair has also been called “causative vs. resultative”. Let’s give an example of this co-occurrence from Ü and Common Tibetan:

\[(108)\] SNGAGS.PA-S CHAR.PA PHAB.NAS DNGOS.NAS BABS-SONG-NGA
ngagpa-ERG rain fall-CO really fall-CMP+SENS-TAG
‘The ngagpa (tantric priest) performed a ritual to make the rain fall and it did rain!’ (Ü, ComTib)
This semantic opposition is somewhat analogous to the aspectual opposition found in the Slavonic languages between imperfective versus perfective, and in English, to phrasal verbs\(^{53}\) (\textit{cut} versus \textit{cut down}; \textit{shoot} versus \textit{shoot down}; \textit{eat} versus \textit{eat up}; etc.).

Finally the meanings of the causative and anticausative may seem quite distinct in some cases, although they are historically related: compare गीत्र GOM' to meditate' versus गिर्ना GOMS 'to be used to'; शूत्र SLOB 'to learn, teach' versus शूत्र LOB 'to get the habit'.

8.3.12. Causative verbs

A few causative lexical verbs are used across the Tibetic area with the meaning 'let somebody do something' or 'cause somebody do something': भूत्र BCUG 'to insert', यास BYAS 'to insert', ब्रोज BZOS 'to make', रोज ROGS 'to help'.\(^{54}\) (See Tournadre & Sangda Dorje 1998, 2003: 267; Simon 2011.) Here are some examples (109-111) in Common Tibetan:

(109) གས་སྦོའི་ལིགས་སྨན་ཁང་བསྟེན་པར་བྱས་པ་ཡིན། NGA-S SPO,BO,LAGS SMAN,KHA BSTEN,PA,B YAYS,PA,YIN
1SG-ERG grandfather-H hospital consult-NMLZ-DAT make-CMP+EGOint

'I managed to have grandfather consult at the hospital.' (Simon 2011: 87)

(110) གས་ཁོང་ལ་ཡི་གེ་འབྲི་རུ་བཅུག་པ་ཡིན། NGA-S KHONG,LA YI,GE BRI,RU BCUG,PA,YIN
1SG-ERG 3SG-DAT letter write-NMLZ make-CMP+EGOint

'I made him write a letter.' (Tournadre & Sangda Dorje 2003)

(111) འཛིན་ལགས་ཀྱིས་ང་མཚན་ནག་གུང་ལ་ཞེད་ཡག་བཟོས་བཞག AZHANG,LAGS,KYIS NGA MTSHAN,NAG GUNG,LA ZHED,YAG BZOS,BZHAG
uncle-H-ERG 1SG dark inside-LOC be afraid-NMLZ make-PERF+INF+SENS

'My maternal uncle made me afraid in the dark.' (Simon 2011: 87)

---

53. In Russian (soversennyj/nesoversennyj) or Polish (dokonane/niedokonane) as for example in: Długo \textit{ciąłem} tę gałąź, aż w końcu ją \textit{ociąłem}.' I was \textit{cutting} this branch for a long time until I finally \textit{cut it down}.'

54. Although rogs is not listed as a verb in the great Tibetan Chinese dictionary, it functions as a verb in some Kham dialects.
8.3.13. Nominalizer

In CT as well as in the modern Tibetic languages, the nominalizers always follow
the verb and are derived from nouns. They play a major role in the grammar (Tournadre & Sangda Dorje 1998, 2013; DeLancey 2010, 2011b). In the modern languages, we
find essentially the same nominalizers as in CT. Frequent nominalizers include བ་(nearly pervasive) or the variant མ་(Pur) < CT universal nominalizer, མཁན་< CT 'learned person, agent, abbot of a monastery' agentive nominalizer, རྒྱུ་(Kh, Am, Ü) < CT 'thing, cloth, cause' patientive
nominalizer, ས་< CT 'place', manner nominalizer. Some nominalizers that are not used in CT are attested. They include: འེ་(NKh and Kh: Minyak Rabgang, E: Th-m); འེ་ CHAS (La, Sp, Tö, Ts, Ü, Lho) < CT 'thing' and the related nominalizers འེ་ YAS, འེ་ YAG (Ü, Ts) or འེ་ SHAD (Lho) pronounced as /-čas/, /-čes/, /-čel/, /-shä/, /-čhe/, /-ču/, /-ča/, etc. Another frequent nominalizer is མི MI < CT 'person, man' or the archaic form མྱི MI realized in various ways as འེ /-mə/, /-må/, འེ /-na/ (Am) /ni/ (Dz), འེ /-nä/ (Kh), etc. is found in many eastern and southern dialects Amdo, Kham, Dzongkha. In Amdo, the nominalizer འེ /-na/ is often written as འེ NO because it has merged with the definite marker འེ BO. In Amdo, the compound form འེ /sheko/ is used as a nominalizer for instrument. The form is probably derived from འེ BYED.PO which is attested in CT. 55 More marginally we find in Amdo the
nominalizer རོགས་ ROGS < CT 'companion' used for the co-agent (Simon 2016) as well as འེ KHÁ < CT 'mouth, surface' replacing in some cases རྒྱུ RGYU. (See above.)

Even when the nominalizers are formally identical to their Classical homologues, their grammatical functions may differ from the Classical one. For example, in some
languages, the nominalizer ས་< CT can be used as instrumental. E.g. in Kham, ས་ZHAMS ZAMAZA-SA (Lit. 'food eat-NML') can mean either 'a place to eat food' (restaurant) or an 'instrument to eat food' (e.g. 'chopsticks').

55. This allomorph is also present with the ordinal number བདུན་ GO the seven ones, the
set of seven; CT བདུན་ BDUN.PO.
Nominalizers play an important role in the modern grammars. They serve to form nominal clauses, relative clauses as well as verb endings (in combination with auxiliaries).

8.3.14. Connective

The category of connectives plays also an important role in the grammars of the modern Tibetic languages, just as they did in Classical Tibetan (see chap 6), because they have multiple functions. They serve as clause connectives as well but they also occur in the morphology of the verb endings, just as the nominalizers.

The languages usually make a distinction between noun and verb connectives.

The main connectives occurring after a noun or a noun phrase are: དང་-DANG (Û, Kh, Hor) < CT associative case, or the variant: ང་-NANG (La), ང་-NA (Sham, Pur, Ba) < CT locative-inessive case. In Amdo, the connective ལ་-RA < CT dative case marker ཞ་-LA.

The main verb connectives occurring after a verb are: བ་STE (La, Dz, Ù, Cho) and its variants བ་STE, བ་TE and ང་STI and ང་SDI (Lho). Another frequent connective is བ་NAS (Û, Ts, Am, Kh). The connective བ་IAS derived from the CT ablative case marker is also used in a number of areas such as Lhoke (Yliniemi 2019). The connective ང་-DANG also occurs after nominalized verbs.

8.4. The semantics of the predicate

All the modern Tibetic languages have developed rich systems of auxiliary verbs which convey temporal, aspectual, modal, epistemic and evidential informations. They may also encode valency and direction.

8.4.1. The aspectual opposition “completed” versus “uncompleted”

Before we examine the tense-aspect paradigms, we will briefly discuss the terms “completed” and “uncompleted” used here and explain why we avoid the terms of “perfective” and “imperfective” for the description of the Tibetic languages.

Although the terms “perfective” and “imperfective” are pervasive in the linguistic literature in English, their use is problematic. These terms derived from Latin perfectum/imperfectum are related to the Slavic aspects совершённый вид (soverennyi vid).
“perfective” versus “несовершенный вид (nesoversennyi vid) imperfective” (Russian) or dokonane versus niedokonane (Polish) “imperfective”. However, the Slavic system is not prototypical and thus the use of “imperfective” / “perfective” should be restricted to systems that function in a similar way as the Slavic basic aspectual opposition. This opinion is shared by many scholars: “definitions of perfectivity and imperfectivity that are to be taken seriously are necessarily formulated in the light of Slavic languages” (Zeisler 2004: 72; see also Dahl 1985; Cohen 1989). As mentioned by several authors, such as Tournadre (2004) or Guentchéva (2016: 3) “[…] the Slavic perfective / imperfective distinction […] cannot be considered prototypical of the theory of aspect.”

So, in order to render the aspectual oppositions in the Tibetic languages, we will use the terms completed and uncompleted. The “uncompleted” aspect refers to an ongoing or habitual event or action in the present, past or future, whereas the “completed” aspect refers to an action or event already finished or completed at a given reference point in the past, present or future.

In most languages, as we will see below (see section 8.4.2.2), the “uncompleted past” is formally equivalent to the present. As we will see, only a few languages such as Ladaks, Purik and Balti make a formal difference between “uncompleted” present and past. Additionally, the Tibetic languages also possess a perfect distinct from the completed past.

A number of studies describing the Tibetic languages have also used the notion of “aorist.” This category, which is found in the verb morphology of e.g. Greek, Bulgarian

56. The opposition between imperfective and perfective, found in the Slavic languages, simultaneously conveys an aspectual meaning (e.g. “completed” versus “uncompleted”) as well as an Aktionsart, i.e. a type of event for example telic or atelic, i.e. have (or not) an inherent final limit. While in some languages grammatical aspect and Aktionsart may combine within a single verb form, they may be also appear separately and should be clearly distinguished.

57. In the French tradition, the terminology accompli for “completed” and inaccompli for “uncompleted” is well established (see e.g. Cohen 1989; Feuillet 2001; Tournadre 2004) and generally used for the description of many languages (Romance, Semitic, Sinitic, etc.) as opposed to perfective/imperfective which are used for Slavic some other specific systems. In English several terms have been proposed to avoid the terms perfective/imperfective such as completed versus uncompleted/non-completed, complete versus incomplete, accomplished versus unaccomplished, etc.
or Persian, often corresponds to the bare root of the verb. Semantically, this category conveys a lack of anchoring in the moment of speech and a lack of situational anchoring and may occur both in the past, present and future. It is particularly used in a narrative register as well as in gnomic statements. Concerning most Tibetic languages, it is safer to consider that some forms may have aoristic uses, but the aorist is not a core category of the Tibetic verb systems.

8.4.2. Tense-aspect

From a grammatical point of view, Tibetic languages often distinguish the following tenses and aspects:

(a) completed past,
(b) habitual (uncompleted present and past),
(c) progressive (uncompleted present and past),
(d) perfect,
(e) future.

Each of these tenses and aspects may appear with various evidential-epistemic categories.

While the opposition between progressive and habitual is attested in many languages such as Ladaks, Purik, Sherpa, Kham, Hor, Dzongkha, etc., it is not pervasive. Some languages such as Ü-Tsang or Amdo have only a present form for both progressive and habitual.

One should also note that some Kham languages do not often use a specific future form and use the present form instead.
8.4.2.1. "Completed past"

The "completed past" is usually conveyed by a number of forms, depending on the evidential and epistemic modalities. This tense-aspect is translated in other languages by the preterit but also sometimes by the present perfect or the aorist (or non-perfect).

For example:

(112) ◊ངས་མ་རིག་ཐལ།
NGA-S MA-RIG-THAL
1SG-ERG NEG-see-CMP+SENS
'I did not see it/ I have not seen it.' (Am)

(113) ◊ངས་རིག་མ་ཐལ།
NGA-S RIG-MA-THAL
1SG-ERG see-NEG-CMP+SENS
'I did not see it/ I have not seen it.' (Kh)

(114) སྒམ་མཐོང་མ་སོང་།
NGA-S MTHONG-MA-SONG
1SG-ERG see-NEG+CMP+SENS
'I did not see it/ I have not seen it.' (Ü, ComTib)

8.4.2.2. Habitual present and uncompleted past

The uncompleted past usually corresponds to the French "imparfait" (or "imperfect") or Spanish "imperfecto". In most Tibetic languages, the uncompleted past, or habitual past is equivalent to the habitual present. See Tournadre & Sangla Dorje (2003), Robin and Simon (forthcoming). Some North-Western languages such as Ladak and Purik do distinguish uncompleted past and present.

Below is an example of a habitual present followed by a sentence with a habitual past. The two sentences differ only by the context.

58. Note that the same marker THAL in some Kham languages (e.g. Lhagang), would yield the aspectual meaning of a perfect. In this dialect, the completed past (or "aorist") is expressed by V-ZIN.YIN/RED.
In many languages, the habitual present may be indicated by the future form:
8.4.2.3. Progressive, continuous and durative

Some languages usually make a distinction between habitual and progressive. Languages which have a progressive or a durative aspect include Ladaks, Purik, Sherpa, Kham, Hor, Dzongkha and Lhoke. The progressive form may be reconstructed in several languages (La, Pur, Sh, Lho, etc.) as a suffix *yin may be related the CT progressive suffix བཞིན་ BZHIN or from the suffix གིན་ GIN which indicate simultaneity. Some languages (Kh, Hor) use verbs of posture such as བསྡད་ BSDAD or འདུག་ DUG both meaning ‘to stay, sit’. The progressive and habitual are compatible with various evidential and epistemic markers.

Compare the Sherpa habitual and progressive:

(121) དེང་སང་ངས་ཤ་ཝིད། (I WID)
DENG.SANG NGAS SHA ZAYID ( < ZA.GI.YOD) 'Nowadays I eat meat.' (habitual)

(122) དར་ལྟ་ངས་ཤ་ཡིན་ཝིད། (now)
DA.LTA NGAS SHA ZAYIN.WID now 1SG-ERG meat eat-PROG+EGO 'Now I am eating meat.' (progressive)

Here are some examples of the progressive in Ladaks and Purik:

(123) ལས་བཅོ་འིན་ནང་ཡོད། (Right now) I am working.' (the suffix -nang means 'just')
LAS BCO-'IN.NANG-YOD work do-DUR-AUTH

(124) འོ་ལས་ེ་ཞེ་ཝད། (Right now), he is working.' (Pur)
KHO LAS BE-EN-YOD 3SG work do-DUR-AUTH

(125) ལས་རྟུབ་བེན་འདུག (He) is chopping meat.' (Pur; Adapted from Zemp 2018)
SHA RTUB-BEN.'DUG meat chop-DUR+SENS
In some languages such as Ladaks, the progressive may also convey a durative function:

(126) ◊ མཚན་ཚེ་རེ་ཆང་འཐུང་ངིན་འདུག

MTSHAN TSHE.RE CHANG 'THUNG-NGIN,'DUG
night all chang drink-DUR+SENS

'(They) have been drinking chang all night.' (La)

However, in order to convey the durative function, the auxiliaries 'DUG' and 'BSDAD' together with the progressive (past/present) are often preferred in Ladaks and Purik.

(127) ◊ སྐར་བེན་འདུག་གེད།

SHA RTUB-BEN,'DUG-GED
meat chop-DUR-ego

'He keeps on chopping meat.' [durative] (Pur, ibid.)

(128) ◊ ཉང་ཞག་དང་ཀི་ཏབ་སིལ་ལིན་འདུག་གེད།

NGA ZHAG-DANG KLTAB SIL-LIN,'DUG-GED
1sg every day book read-DUR-ego

'I keep reading books every day.' [durative] (La)

In some languages such as Kham and Hor, the same auxiliaries 'DUG' and 'BSDAD' (past)/'SDOD' (present) convey the progressive. In Dzongkha the progressive form ◊ དོ་ may also be derived from 'SDOD.'

(129) ◊ མ་གྲི་འདུག་རེད།

KHO YLGE BRI-DUG.RED
3sg letter write-PROG+FACT

'S/he is writing a letter.' (Kh: Gyalthang)

In some languages, the progressive is marginal and rarely used. That is the case in Common Tibetan (ex. 131) with the progressive construction: V(past)+'SGANG.YIN/'SGANG.RED('lit. to be on doing something').
8.4.2.4. Perfect

The perfect aspect refers to a past situation that is still relevant at the moment of utterance. This aspect is generally found in the Tibetic languages. Several forms convey the perfect together with various evidential or epistemic modalities. In most languages, the perfect is marked by existential verbs, sometimes preceded by a connective (NAS, TE, etc.). The perfect is compatible with various evidential categories (egophoric, factual, sensory) and epistemic categories.

The non-sensory perfect based on logical inference is conveyed by the existential auxiliary YOD as well as compound existential forms which include this auxiliary (see 8.3.3.3).

Auxiliaries that are frequently attested to convey the “sensory perfect” include: DUG (and the related forms Tog (La) and Tsug (Pur), ZUG (Am, Kh), NUG (Dz), GDA (NH, Kh: Yülshül, Am), BZHAG (Ü), SNANG (E, Kh) and GI (Kh: Northern route). See 8.3.10. In some dialects such as Tsang, the sensory is marked with a connective NAS after the verb (without auxiliary) in affirmative sentences. However, in the negative sentence, the auxiliary is not dropped NAS.MI.DUG.

(131) 1SG letter write-PROG+EGO

'I am writing the letter.' (Ü, ComTib)

'Ve are getting married.' (E, Kh: Yülshül, Am) (adapted from Häslar 1999)

(132) snow fall-PERF+SENS

'It has snowed.' (Kh: Derge) (adapted from Häslar 1999)

(133) door close-PERF+SENS

'The door is closed.' (Lit. 'has been closed.') (Ü, ComTib)
PART 2 – CHAP. 8. Grammatical outline of the Tibetic languages 387

(134) ཚོག་བརྒྱབ་མི་འདུག
   SGO BRGYAB-ML DUG
door close-NEG+PERF+SENS

'The door is not closed.' (Lit. 'has not been closed.') (Ü, ComTib)

Note that in Common Tibetan, the auxiliary used in the negation (DUG) is not the same as the auxiliary used in the affirmative sentence BZHAG.

In Ladaks, the same example is:

(135) ཚོག་བཅུགས་ཏེ་འདུག
   SGO BCUG-TE DUG
door close-PERF+SENS

'The door is closed.' (Lit. 'has been closed.') (La)

8.4.2.5. Future

The future is attested in most languages. In many languages, the future is formed by adding the equative copulative verb ཡིན་YIN or other equative copulative verbs (such as རེད་RED, etc.) after a verb usually followed by a nominalizer (see 8.3.13): V-ཞི་+ཡིན་-GIYIN (Ü), V-ཞི་+ཡིན་-NLYIN (Am), V-ཞི་+ཡིན་-NLYIN (Dz), V-ཞི་+ཡིན་RGYU+YIN (Am, Kh), V-ཞི་+ཡིན་-LE+YIN (Kh), V-ཞི་+ཡིན་SHAD+айн (Lho), V +ཡིན་YIN (La).

Ladaks has also developed a second type of future in ◤ འཁྲོན་CEN derived from the nominalizer འཁྲོན་CES +ཡིན་YIN (Zeisler pers. comm. 2020). Some languages (Ü, Ts, Kh, Am, La, etc.) make a distinction between intentional and non-intentional future (see Tournadre 2016, 2017). Here are some examples:

(136) དེ་ཐོག་ལེ་ཡིན།
   NGA 'ONG-LE YIN
1SG come-FUT+EGO

'I will come.' [general] (Kh: Minyak Rabgang)

(137) དེ་ཐོག་ལིའི།
   NGA 'ONG-LI'
1SG come-FUT+EGO

'I will (definitely) come.' [intentional] (Kh: Minyak Rabgang)
Some exceptional languages, such as Rongdrak Kham do not have a specific form for the future and only distinguish past from non-past.

Here is an example of non-past:

(143) ◊ བོད་ཁུལ་ཁོང་

DE ZAN ZA-MI.YIN
3SG food eat-FUT

'S/he will eat food/ is eating food.'
8.4.3. Evidential modalities

Tibetic languages have developed rich and complex evidential-epistemic systems (E-E system) which are semantico-cognitive in nature although they exhibit some syntactic secondary effects. These systems also involve the category of stance and ethos as we will see below (see the notion of authoritative). For an overview of the evidential systems of Tibetic languages, see Gawne & Hill (2017). See also Bickel (2000, 2001).

In all the modern Tibetic languages, the marking of evidentiality and/or epistemity is compulsory in finite clauses.

Evidential and epistemic modalities are marked by verb ending morphemes often derived from copulative and existential verbs or auxiliaries (see section 8.3 and 8.4) that may combine together. Every verb ending conveys either evidential or epistemic or a combination of evidential and epistemic meanings.

Evidential and epistemic marking is correlated with the various tenses and aspects. As expected, in the Tibetic languages, the E-E paradigm of form is more developed in the past than in the present or the future.

The major categories in Tibetic languages of the E-E system are (a) “sensory”, (b) “inferential”, (c) “authoritative”, (d) “hearsay and reported speech”, (e) “epistemic”. These 5 macro-categories are attested in all the modern Tibetic languages. However, these categories may be divided into several subcategories as we will see below. Some subcategories are specific to a given language or even a given dialect.

For the notion of “evidential”, we propose the following definition: the representation of source and access to information according to the speaker’s perspective and strategy. (Tournadre & LaPolla 2014). This definition of evidentiality makes a clear distinction

---

59. Evidential markers have sometimes been called “mediative markers”. The term “médiatif” in French, proposed by Lazard (1956) for a variety of Tajik, referred mainly to “indirect” source and access of information. Later the use of “médiatif” was mainly associated with the verb systems of various languages such as Bulgarian, Persian, Turkish, etc., to indicate hearsay and inference, hence a “mediate information” (Lazard 1999; Guentchéva 1996). These languages do not have a paradigm of “sensory evidentials” indicating specifically direct sensory perception. The term “médiatif” was also applied to Tibetan by Tournadre (1994a, 1996a). In this book, we stick to the English-speaking tradition and will use only the term “evidential” as defined in Tournadre & LaPolla (2014).
between “source” and “access”. In the Tibetic languages, both source and access to information are grammaticalized. The Tibetic Evidential/Epistemic systems are associated with a number of typological characteristics and morphosyntactic behaviours that play a crucial role in the functioning of the systems. Among these characteristics, one should mention:

(a) the complex correlation with the category of person;
(b) the correlation with tenses and aspects;
(c) the anticipation strategy and perspective shift in interrogative sentences;
(d) additionally, in a number of languages, such as Ü-Tsang, Amdo or Kham, etc. evidentiality interacts with the semantic categories of “intentionality” or “animacy”;
(e) the paradigm of evidential and epistemic categories conveyed by auxiliary or copulative verbs is only available in the main clause, but generally not in subordinate clauses (see Chang & Shefts 1964; Tournadre & Sangda Dorje 1998, 2003: 76, 142; Garrett 2001);
(f) The choice of evidential or epistemic markers is flexible and depends on the speakers’ communicative intention. The flexibility in the use of the markers has been noted by some authors (see e.g. Zeisler 2012b; Tournadre 2008, 2017; Gawne 2013 and Hill 2013b).

Finally, we ought to emphasize that the five major E-E categories – sensory, inferential, authoritative, hearsay and reported speech, epistemic – may include many subcategories, some of which are present only in a few languages. For example, some languages such as Ladaks and Spiti, distinguish “visual sensory” from “non-visual sensory”, whereas the majority of the Tibetic languages do not make this distinction. Within “non-sensory”, the “endopathic” category has usually has a special status (see 8.4.3.1 & 8.4.3.2). Some languages such as Dzongkha and Choća-ngāča have developed a specific category “participatory-sensory” (see 8.4.3.3). Inferential markers also play a very significant role in the various languages (Zeisler 2012b). In most cases, the E-E systems distinguish several subtypes which include “sensory inferential” and “logical inferential” (see 8.4.3.4). The category of “authorative” includes “factual” and
“egophoric” which are present in most Tibetic languages (see 8.4.3.5 and 8.4.3.6). The egophoric category can be itself further divided into several subcategories: “intentional egophoric”, “receptive egophoric”, “benefactive egophoric”, etc. as well as “strict” and “loose” egophoric (see 8.4.3.7). The hearsay and reported has a special status in the Tibetic E-E system (8.3.4.9).

Additionally some marginal categories such as self-corrective (counter-expectation) and mnemonic are also found in some Tibetic languages.

It is important to underline the complexity of the E-E systems and the existence of overlaps between various categories. For example, the evidential inference and epistemic may combine within the same verb endings (see 8.4.4).

In the following sections, we will first deal with the evidential markers, then tackle the epistemic markers and then examine some specific features of the E-E systems.

8.4.3.1. Sensory

The category of “sensory” refers here to both “external” and “internal” sensory access to information. This category is attested in all the major Tibetic languages. It may be acquired through the sensory channels of the five senses of sight, sound, touch, smell, and taste. But the sensory marker may also be used for “endopathic”, a term coined by Tournadre (1996c: 226), to indicate an “internal sensory” access to information. Endopathic marking encodes inner sensations such as cold, pain and hunger, as well as psychological states and emotions such as fear and anger. In some languages, it may also refer to intuitive feelings.

Note that the sensory marker often conveys to information recently acquired through perception (see e.g. van Driem 1998, see also the discussion about the mirative in 8.4.8.5), but it does not need to be the case. The observation may have occurred long ago! This is the case in the following example in Ladaks:

(145) ◊ སི་ཁུལ་ མི་རུག

SILKHUL MI-RUG
school NEG-EXV+SENS

‘There was no school (at that time).’ (Lit. ‘there is no school’, I witnessed this situation.)
The speaker, Morup Namgyal, a famous Ladakhi singer, tells about his childhood and remembers that there was not any school in Wanla, his home village when he was a child, more than sixty years ago. In this interview, he further describes his village using mainly the sensory marker 'DUG' or its variant 'RUG'.

The fact that sensory marker is not always related to recent observation is further illustrated by the existence of the past sensory marker 'DUG.PIN' and the allomorph 'RUG.PIN' in Ladaks. To render the same situation the speaker could have said:

(146) དིག་པིན་མིའདུག

'school NEG-EXV+SENS-PST'

'There was no school.'

The use of sensory marker for observation that occurred long ago is not specific to Ladaks and is also attested in other Tibetic languages. In Common Tibetan, the same sentence is rendered as:

(147) སློབ་གྲོ་མི་འདུག

'school NEG-EXV+SENS

'There was no school (at that time).' (Lit. 'there is no school'.)

The category of “sensory” is pervasive in the Tibetic languages but as we will see it may receive various extensions depending on the language. Modern languages usually have various sensory markers depending on the tense-aspect. For the "present" (or "uncompleted/ habitual past") and "perfect" the same auxiliary is often used, whereas a specific marker occurs with the "completed past" sensory.

For the present and perfect, one encounters various sensory forms derived from different lexical verbs depending on the language. In many central, southern and western languages such as Ü, Tsang, Sherpa, Lhoke, Dzongkha, Ladaks and Purik it is derived from the verb 'DUG' to stay, to sit'. Note that the form 'DUG' had already acquired an evidential meaning in Classical Tibetan (Hill 2013; Oisel 2013).

60. The Indian pronunciation of 'DNGOS.GRUB.'
In Hor and several northern Kham dialects, one finds instead the verb གདའ་GDA’ < CT ‘to sit’. Another form, derived from སྣང་SNANG ‘to shine, to appear’, is used for the sensory access marker in Phānpo (central Tibet), in many varieties of Kham,61 in some dialects of Ladakh and Baltistan as well as in Pangi (see 8.3.10). In Amdo, the suffix གི་GI is used. The origin of the sensory markers གི་GI used in the Kham Derge dialect, as well as the Dzongkha བས་MAS are not clear.

In the completed past auxiliaries derived from གོ་THAL ‘to go’ (Kham, Am) and སོང་SONG ‘to go’ (Ü, Ts, Sher, etc.) are frequently attested to convey a sensory meaning.

The “sensory” marking (except for the endopathic, see below) is normally used in declarative and interrogative sentences with the 3rd (singular or plural) person “subject” (A or R). It is sometimes used with the 2nd person “subject” in declarative sentence, but not normally with the 1st person “subject.” The reason is clearly pragmatic: one cannot be a witness of oneself, except in some specific situations. For example, to say ‘I am eating’ or ‘I am writing’, Tibetic languages do not use sensory markers simply because it would entail ‘I see myself eating’, ‘I see myself writing’. If the speaker sees himself in a mirror, in a dream, in a movie, etc. then the use of a sensory with the first person would be perfectly acceptable.

(148) གཞས་ལ་.nyན་གི་(འདུག)
GZHAS-LA NYAN-GLI (’DUG)
song-DAT listen-UNCMP+SENS
‘S/he is listening to a song.’ [I see or I hear s/he is listening to a song] (Ü, ComTib)

(149) གླུ་འ་.nyན་གོ་གི
GLU-A NYAN-GO.GI
song-DAT listen-UNCMP+SENS
‘S/he is listening to a song.’ [I see or I hear s/he is listening to a song] (Am)

61. Note that in some Kham dialects, SNANG is used as a “non-egophoric” or is restricted to “visual sensory”.

"What is s/he doing?" [have you seen, etc. what s/he is doing now?] (Am)

In some languages, the sensory marker may have a different form depending on the controllability of the verb. For example in Amdo, the form གི GI is used with adjective predicates,62 emotion verbs, etc., whereas the compound form གིཡོད་གི GI-YOD.GI, often shortened as གོ་གི GO.GI (see Sun and LHA.BYAMS.RGYAL 2005: 130) occurs with controllable verbs. Compare for example:

(150) ཁི་ཟིག་ཡེད་གི CHI.ZIG YED-UNCMP+SENS
‘What is s/he is doing?’ (Am)

(151) གླུ་འ་ཉན་གོ་གི GLU-A’ NYAN-UNCMP+SENS
‘S/he is listening to a song.’ [the speaker witnesses the scene] (Am)

(152) ཆོག་གི CHOG-GI
‘It is alright.’ (Am)

(153) མང་གི MANG-GI
‘There is a lot.’ [the speaker sees it] (Am)

Two forms for the sensory are also found in Dzongkha: when the sensory occurs as an existential verb, it appears as བོད། DUG but if it is used as a verb or as an adjective suffix, it appears as ས་ MAS:

(154) རེགས་ཤོམ་འདུག LEGS.SHOM DUG
‘(This) is good.’ (Dz)

62. However, as noted by Camille Simon (pers. comm. 2020), the main opposition might be between "stative verbs" and "action verbs". See also Tribur (2019).
The following examples illustrate the auxiliary SNANG conveying a sensory meaning. Note that the S of SNANG is never pronounced in the various languages, but in Kham may trigger an unvoiced nasal [ŋ]: /n’ang/.

In many languages (Ü, Ts, Sh, Dz, etc.), the “visual sensory” markers are formally identical to the “non-visual” markers (i.e. auditory, gustatory, tactile, olfactory and endopathic). The choice of one sense or another depends on the semantics of the predicate and on the context. For example, in Common Tibetan (ex. 159-164), presenting an object (cloth, food, sound, etc.) to somebody, one may ask:

63. S. Ebihara translated this as ‘the tea is hot’. This is maybe due to the fact that Japanese does not have an expression for ‘warm tea’. Tea is either ‘hot’ or ‘cold’ whereas ‘warm’ and ‘cool’ are reserved for the honorific register. According to our Purik informants, ‘hot’ is /ts’ante/ and /dronmo/ means ‘warm’ in a similar way to Ladakhs and Zanhar. This is also confirmed by Zemp (2018). See also the CTDT.
Or ask about an endopathic information:

(164) **SNYING.RJE.PO** 'DUG-GAS
2SG stomach hungry-STAT+ENDO-Q
‘Are you hungry?’ [endopathic]

By using the sensory marker *DUG*, the speaker invites the addressee to look at her
dress, to taste the dish, to smell the object, to touch the clothes, to listen to the sound
or to tell whether s/he feels hungry.

Conversely, the use of an auxiliary may have an impact on the interpretation of the
verbal or adjectival predicate. For example, in Ladaks which distinguishes a visual
sensory from a non-visual sensory (see 8.4.3.2), the use of one marker instead of the
other may have an incidence on the predicate interpretation, as shown in the sentences 165-166 and 167-168.

(165) གསལ་པོ་འདུག GSAL.PO 'DUG clear EXV+VIS 'It is clear.' [visual]

(166) ◊ གསལ་པོ་རག GSAL.PO RAG clear EXV+NVIS 'It is clear, I understand.' [non-visual]

(167) བདེ་མོ་འདུག BDE.MO 'DUG good EXV+VIS 'It is beautiful.' [visual]

(168) ◊ བདེ་མོ་རག BDE.MO RAG good EXV+NVIS 'It is good.' [I feel, non-visual]

Very often, the use of a sensory marker implies several senses. Here is an example in Common Tibetan:

(169) མེ་ཆེན་པོ་ཅིག་འདུག ME CHEN.PO-CIG64 'DUG fire big-INDEF EXV+SENS 'There is a big fire.'

The speaker sees the fire, but at the same time smells it, feels the smoke in his/her eyes, etc. As we will see in the next section, in languages which have a distinction between visual and non-visual sensory markers, the visual sensory prevails over the other types of perception (including auditory) in the case of perceptions involving various senses.

---

64. The indefinite marker is ཉི་ ZHIG in written Tibetan. However, it is pronounced /či(k)/ in the spoken language (Ü, Common Tibetan), in the same way as the cardinal number ཉི་ GCIG.
Endopathic access

It is worth noting that even the languages that do not make a morphological distinction between "visual sensory" and "non-visual sensory" marking, exhibit a specific syntactic behavior for the endopathic marking. The endopathic function is restricted to the 1st person "subject" (experiencer or "sensory captor" $S_p$) in declarative sentence and the 2nd person "subject" in interrogative sentence. (See the anticipation strategy in 8.4.8.3.)

This clearly shows that the endophatic has always a special status within the various types of cognitive sensory access to information.

Here are some examples:

(170) ◊ གྲོད་ཁོག་ལྟོགས་ཀྱི་འདུག
KHONG GROD.KHOG LTOGS-KYL.(’DUG)  
1SG stomach hungry-STAT+ENDO
‘I am hungry.’ (Ü, ComTib)

Note that in the Lhasa dialect, the auxiliary 'DUG is sometimes dropped in affirmative assertions but it always appears in negative sentences and questions.

(171) ?? གྲོད་ཁོག་ལྟོགས་ཀྱི་འདུག
KHONG GROD.KHOG LTOGS-KYL.(’DUG)  
3SG stomach hungry-STAT+ENDO
Intended meaning: ‘He is hungry.’ (Ü, ComTib)

(172) ◊ གི
NGA (’HO) LTOGS-GI  
1SG (belly) hungry-STAT+ENDO
‘I am hungry.’ (Am)

(173) ?? གི
KHO.DGE (’HO) LTOGS-GI  
3SG (belly) hungry-STAT+ENDO
Intended meaning: ‘He is hungry.’ (Am)

65. /ho/ is the Amdo pronunciation of PHO.BA.
As mentioned above, the endopathic can not be used with the 3rd person nor the 2nd person in declarative sentences. Three strategies are then available. The examples below are from Amdo, but any language which has developed endopathic marking would have these various strategies:

(a) sensory inferential

(177) ◊ མུར་གེ་སྐྱག་གོ་ཁོལ་ཁེ་འཛིག་རེད

MUR.GE MGO KHOL-GO.GI ZER.GI
3SG+F head boil-UNCMP+INF+SENS
say-UNCMP +SENS
'She says she has a headache.' (Am)

(b) reported speech

(178) ◊ མུར་གེ་མགོ་ཁོལ་གོ་ཁ་ཟིག་རེད

MUR.GE MGO KHOL-GO.GI ZER.GI
3SG+F head boil-UNCMP+INF+SENS say-UNCMP +SENS
'She says she has a headache.' (Am)

(c) epistemic

(179) ◊ མུར་གེ་མགོ་ཁོལ་གོ་ཁ་ཟིག་རེད

MUR.GE MGO KHOL-GO.KHAZIG.RED
3SG+F head boil-UNCMP+INF
'She must have a headache (Lit. 'head boils').' (Am)
(180) མུར་ ལྭེ་ མགོ་ ཁོལ་ ཉ་ ཐང་ གི།
MUR. GE MGO KHOL-NA. THANG-GI
3SG+F head boil-UNCMP-EPI-SENS
'She must have a headache.' (Am)

It is worth noting that in some rare languages such as Purik (Kargil), the endopathic access is not indicated with a sensory auxiliary but with the authoritative auxiliary, derived from CT གོ སྦ YOD.

(181) རྒྱ་ བོད་ རི་ སེན།< བོད་ སེན།
NGA LTOGS-SED LTOGS-YOD
1SG be hungry+PRES+AUTH
'I am hungry.' (Pur)

(182) རྒྱ་ ཚུག་ རི་ སེན།
NGA LTOGS-DUG
1SG be hungry+PRES+SENS
Intended meaning: 'I am hungry.' (La)

(183) རྒྱ་ ལོག་ སེན།
NGA 'JIGS-SED
/nga zhiks-et/
1SG fear+PRES+AUTH
'I am afraid.' (Pur)

**Dream narratives**

We will now make some remarks about the use of sensory markers in dream narratives. According to Aikhenvald (2004: 344), "the treatment of dreams varies from culture to culture [...]. Some languages treat dreams on a par with ordinary directly observed experience. [...] In other languages dreams are cast in non-first hand evidentials."

In the Tibetic languages, dream narratives are normally told with the use of sensory markers or sensory inferential markers. When the speakers describe dreams which did not involve their participation, *sensory* and *inferential auxiliaries* are used as if the speakers had really seen the given situation.

However, when the speakers describe what happened to them in a dream, they also normally use sensory markers as if they perceive themselves as "actors" of a dream (see
e.g. Oisel 2017a; Tournadre 2008; Tournadre & Sangda Dorje 1998). This shift in the consciousness has an immediate grammatical consequence: in most languages, they do not use the same auxiliaries that would be required to describe an ordinary state of consciousness. Thus, dream narratives have specific grammatical features. In particular in dream narratives describing their own actions and states, speakers do not use the intentional egophoric markers (that convey the intention of the speaker, see below 8.4.3.7), nor do they use endopathic markers (that indicate an inner feeling or sensation, see above). The reason is that they do not have access to the intentions nor the inner feelings of the “self” which performs in the dream. These two types of auxiliaries are replaced by sensory or inferential markers.

8.4.3.2. Visual sensory versus non-visual sensory

Some Tibetic languages make a morphological distinction between “visual” and “non-visual sensory” (i.e. gustative, auditory, tactile, olfactory, and endopathic). That is for example the case in Ladaks, Tö Ngari, Dolpo or Spiti-Khun-Garzha, and some Kham dialects (Bathang, Derge, Gyālthang). The visual evidential is usually marked with the auxiliary སུགག ’DUG, whereas the non-visual evidential is indicated by an auxiliary derived from ང་ཟིལ་ GRAG ’to sound’ (SKAL.BZANG ’GYUR.MED & SKAL.BZANG DBYANGS.CAN 2002; Häsl 1999; Hongladarom 2007a; Bartee 2007; Tournadre & LaPolla 2014).

Compare for example the “visual” and “non-visual” evidentials:

(184) ◇ ཁ་ སུགག ’DUG

person EXV+VIS
‘There is somebody.’ (Sp)

The use of ’DUG indicates that the speaker’s statement is made on visual access. The speaker sees or has seen that there is somebody. At least that’s what he or she pretends. Of course, one should bear in mind that the speaker may lie, and has not necessarily seen the event.

(185) ◇ ཁ་ ང་ཟིལ་ GRAG

person EXV+NVIS
‘There is somebody.’ (Sp)
The use of *GRAG* indicates that the speaker hears or has heard voices, but it may also be used if the speaker had access to information through other channels such as touching, smelling...

Here are three examples from *SKAI,BZANG 'GYUR.MED & SKAI,BZANG DBYANGS,CAN* (2002):

(186) ◊ ཨ་ང་ེན། རིན་པོ་ཆེས་བུ།

*DBANG.CHEN SLEBS-'ONG-GRAG*

Wangchen arrive-DIR-PERF+NVIS

‘Wangchen has arrived.’ [based on hearing] (Kh: Bathang)

(187) ◊ ཨ་ཿ་ིོ། སྤེལ་འོང་གྲག

*KHYOD-GI RUM-NANG STAR.GA YOD.GRAG*

you-GEN pocket-LOC walnuts EXV-NVIS

‘In your pocket, there are walnuts.’ [based on touching] (Kh: Bathang)

(188) ◊ ཨ་ཿ་ིོ། སྤེལ་འོང་གྲག

*SPOS-KYI DRLMA DRO-'ONG.GRAG*

incense-GEN flavor exhale-DIR+NVIS

‘The flavor of the incense is coming.’ [based on smelling] (Kh: Bathang)

Zeisler (2018a: 93) provides a good example of the differences between the visual and non-visual markers in Ladaks (Leh):

(189) ◊ ཨ་ཿ་ིོ། སྤེལ་འོང་གྲག

*THAR.MOSI NANG-NGA JA DA.RUNG RAG-GA MI-RAG*

/'tarmos-i nang-a ça darung rag-a mi-rak/

thermos-GEN in-LOC tea still EXV+NVIS-Q NEG-EXV+NVIS

‘Is there still [some] tea in the thermos flask or not?’ (La)

She gives the following comment: “While uttering this sentence, the speaker might take up the flask and shake it to feel whether there is some liquid left. S/he might also expect the addressee to do so or to have done so a moment before. If s/he would take out the cork and peep through the opening or if s/he expects the addressee to do so, s/he would use the existential verb for visual experience *ḥdug*”:
It is interesting to note that in the case of perceptions involving both visual and other senses such as auditory, etc., the visual marker is generally used.

In the first example above, the child is present in the room, and the visual auxiliary is used although the speaker also hears the child crying, whereas in the second sentence, the non-visual is used because the speaker hears the child crying in the next room but does not see it. It should be noticed that although the sound is certainly a salient feature of cries, it is nevertheless the sight that is cognitively dominant. The same situation appears with coughing as illustrated below:

These examples also show that the visual marker does not entail that other senses are not involved. Indeed when using the visual marker ("seeing that somebody cries or
coughs”), the speaker also perceives the sounds, but the visual perception is considered as primary.

This appears even more clearly in the following examples:

(195) ◊ གླུ་རྒྱལ་ལ་བཏང་འདུག
GLU RGYAL.LA BTANG-'DUG
song well V*-PRES+VIS
'(She) is singing well.' (I can see her performing on the stage) (La)

(196) ◊ གླུ་རྒྱལ་ལ་བཏང་ང་རག
GLU RGYAL.LA BTANG-NGA.RAG
song well V*-PRES+NVIS
'(She) is singing well.' (I hear her singing on a recording.) (La)

Even, in this case where the sound is determinant for the evaluation, the visual marker is normally preferred (see the evidential accessibility hierarchy in 8.4.3.10).

In systems with visual and non-visual sensory, as one could expect, the endopathic (or "inner sensory") function is always marked by the non-visual evidential.

(197) ◊ གླུ་ལྟོགས་རེ་གྲག
NGA-LTOGS-RE.GRAG
1SG-DAT hungry-PRES+ENDO
'I am hungry.' (Garzha)

(198) ◊ གླུ་ལྟོགས་ས་རག
NGA-LTOGS-SA.RAG
1SG hungry-PRES+ENDO
'I am hungry.' (La)

There are however some rare exceptions such as Dolpo which uses a visual sensory and not the non-visual as excepted.

8.4.3.3. Participatory-sensory

Some Tibetic languages such as Dzongkha or Choča-ngača have a special marker to convey the fact that the speaker has either witnessed or consciously taken part in an action or a situation. This marker has been called "witnessed past" by van Driem (1992, 1998). However, as we have seen earlier in the Tibetic languages "sensory" markers normally do
not occur with first person (except for the endopathic function and other specific situations, see 8.3.11, Oisel 2017a). In order to designate the markers that normally occur both with 1st person (the speaker is a participant of the event) and with 2nd or 3rd person (the speaker is a witness of the event), we propose the term participatory-sensory. The term “participatory” is used to describe some evidential markers of Papua New Guinea and are related to the participation of the 1st person (see San Roque and Loughnane, 2012a-b).

This is, for example, the case of the marker yi in Dzongkha.

(199) ◊ ང་ལྟོ་ད་ཅི་ལས་ཟ་ད་ཡི།
   NGA LTO DA.CI-LAS ZA-DA-YI
   1SG meal before eat-SEC-PAST+PART
   'I have already had my meal.' (Dz; van Driem 1992: 243)

(200) ◊ ད་མོ་གིས་ཨོམ་འབོ་ད་ཡི།
   DA. MO-GIS ṬOM 'BO-DA-YI
   now she-ERG milk split-SEC-PAST+PART
   'Now she’s spilt the milk.' (Dz; van Driem 1992: 243)

This is also true of the marker -DO in Choča-ngača.

(201) ◊ ད་ལྟ་ང་མོང་སྒར་སོང་དོ།
   DA. LTA NGA MONG.SGAR SONG-DO
   now 1SG Mongar go-PROG+PART
   'Now I am going to Mongar.' (Cho; Tournadre & Karma Rigzin 2015)

(202) ◊ གར་པ་བཏང་དོ།
   CHAR.PA BTANG-DO
   rain fall-PROG+PART
   'It is raining!' [observing the rain falling] (Cho; Tournadre & Karma Rigzin 2015)

The progressive -DO in Dzongkha has a similar reading. According to van Driem (1998): it expresses “an activity which the speaker by his own observation knows to be going on in the present.” Thus, the suffix -DO is clearly a sensory marker. However, -DO also frequently occurs with the first person (van Driem 1998):
The sun is shining.' [the speaker looks at the sun] (Dz)

'I am eating.' (ibid.)

Thus, this suffix fits well the above definition of sensory-participatory that we proposed. If it is the case, we can see this suffix has a somewhat similar function (but not entirely equivalent) to the cognate form in Choča-ngača. Note that Hyslop and Karma Tshering (2017) have a different approach and analyze the form -DO in Dzongkha as an "egophoric".

8.4.3.4. Inferential

The inferential mood indicates that the basis of the speaker’s assertion is an inference or a conclusion that is being drawn from the traces or the present results of a past action, or from a logical calculation. The speaker may also make an inference drawn from the present situation to predict a future event. Inference may be considered as certain by the speaker or it may bear various degrees of uncertainty. In this section, we only deal with the inferential conveying certainty. For the epistemic inferential, see 8.4.4.

The inferential may be essentially of two types: sensory inferential and logical inferential. Let us first examine the former type. In the Tibetic languages, sensory inferential markers are often derived from the following auxiliaries: ཀྲུང་ DUG (Ü, Ts, La, Ba, La, Pur, Kh, Dz), མཉམ་ SNANG (Kh, Ba: Turtuk), ཞོང་ GDA’ (Hor, Kh), ལུང་ GRAG (La, LJ, Sp, Tö, Kh), འཁུལ་ BZHAG (Ü), མུ་ ZUG (Kh, Am). Some of these auxiliaries such as ཀྲུང་ DUG, ཞོང་ GDA’. མཉམ་ SNANG indicate a sensory perception (see 8.4.3.1-2, 8.3.10) when they are used in the present and progressive or uncomplete past, whereas they indicate a sensory inference when they occur in the perfect aspect.

It is important to emphasize upon the fact that sensory inferences are not always conveyed by "sensory inferential markers." They may often be realized by simple
sensory markers. The reason is obvious. From a cognitive point of view, perceptions often imply a certain degree of inference as shows the following Ladaks sentence:

\[(206) \quad \text{CHU} \quad \text{DRON.MO} \quad \text{DUG} \]

today water warm VIS

‘This water is warm.’ [the speaker sees the steam over the water, and infers that the water is warm] (La)

Although the sensory marker ‘DUG’ is used, the speaker may not directly see that the water is warm and only infers the warm temperature from looking at the steam.

The same sentence with cold water would be inappropriate:

\[(207) \quad \text{CHU} \quad \text{GRANG.MO} \quad \text{DUG} \]

this water cold VIS

‘Intended meaning: This water is cold.’ (La)

The speaker would not have sufficient clue to “see” that the water is cold.

The only way to know would be to touch the liquid and then the non-visual sensory marker ‘RAG’/rak/ would be appropriate:

\[(208) \quad \text{CHU} \quad \text{GRANG.MO} \quad \text{RAG} \quad \text{DRON.MO} \quad \text{MI-RAG} \]

this water cold NVIS warm NEG-NVIS

‘This water is cold, it is not warm!’ [the speaker touches it or tastes it] (La)

The following frequent sentence in Ladaks illustrates the same phenomenon:

\[(209) \quad \text{DI.RING} \quad \text{NAM.LA} \quad \text{GRANG.MO} \quad \text{DUG} \]

today weather cold VIS

‘Today, the weather is cold!’

The speaker uses a visual sensory marker, looking at the landscape from the window. Although, the sentence implies an evaluation of the temperature, the speaker may have clues about the outside temperature. For example, in Ladakh, in winter, a cloudy weather implies a cold temperature. Of course, if one goes outside and feels directly the temperature, one will naturally say:
This being said, when one does not have a direct perception and infers a situation from various sensory evidences.

Here are some examples of sensory inferential markers:

(210) ‘DI.RING NAM.LA GRANG.MO RAG
today weather cold NVIS
'Today, the weather is cold!' (La)

(211) CHAR.PA BABS-NAS.SNANG
rain fall-PTT+SENS+INFR
'It has rained.' [The speaker sees the ground wet, sensory inference]. (Kham, Budy dialect)

(212) KHO NA-BSDAD.SNANG
3SG be sick-CONT-INF+SENS
'He is sick.' [The speaker has seen him or is looking at him and concludes he is sick] (Kham, Tormarong [alt. Dongwang] dialect; adapted from Bartee 2007)

(213) KHA BTANG-TOG
snow LV-CMP+INF+SENS
'Oh, it has snowed!' [looking at the white mountains around] (La)

The same sentence in Common Tibetan would be:

(214) GANGS BTANG-BZHAG
snow LV-CMP+INF+SENS
'Oh, it has snowed!' [looking at the white mountains around] (Ü, ComTib)

The sensory inferential contrasts with the sensory:

(215) GANGS BTANG-SONG
snow LV-CMP+SENS
'It snowed.' [yesterday, I saw the snow falling] (Ü, ComTib)
It is important to clarify that sentence (198) and (199) do not bear any uncertainty and can not be translated by: ‘Oh it must have snowed!’. The Tibetic languages usually have grammatical means to distinguish these two interpretations.

Even with inferential related to the future, the prediction is considered as certain.

(216) ◊ གོ་རེ་ཁ་བཏང་ངོག
THO.RE KHA BTANG-NGOG
'tomorrow snow LV-FUT+INF
'Tomorrow, it will snow.' [looking at the sky complety covered] (La)

Even, if meteorological prediction is far from being certain, the speaker by using this inferential does not have doubts about his prediction. Otherwise he can use various epistemic inferences (see 8.4.4).

For example the sentence below would convey uncertainty:

(217) ◊ ལ་འ་ཁ་བཏང་མི་བཏང་ང་ཧེ།
LA-A KHA BTANG-MI-BTANG-NGA.HE
pass-LOC snow fall-NEG-fall-FUT+EPI
'On the pass, it will probably snow.' (La)

In Ladaks, the opposition between visual and non-visual sensory inferences is very common. Here are some other examples:

(218) ◊ རོ་chang ནིན་ནོག
ʃO CHANG YIN.NOG
INTJ chang be+SENS
'Oh, (it) is chang!' [by looking at the liquid in the glass] (La)

(219) ◊ རོ་chang ནིན་གྲག
ʃO CHANG YIN.GRAG
INTJ chang be+SENS
'Oh, (it) is chang!' [by smelling or tasting liquid in the glass] (La)

In Balti, Purik and Ladaks, the inferential may be conveyed by the auxiliary ◊ །TSUG or ◊ །SUG (Ba, La, Pur) < CT །DUG (Zeisler 2012b, 2017; Zemp 2018), whereas in Amdo and Kham, the auxiliary །ZUG (Kh, Am) is attested. The latter is also attested in Kham Derge dialect (see Häsl 1999) and Amdo.
The water came up to here (pointing to the chest).’ (Pur; Zemp 2018)

‘They are drinking alcohol.’ (Am)

‘He turned out to be a policeman [dressed in civil clothes].’ (La)

‘If he ate all the sweets he would have a stomach ache.’ (Ü, ComTib)

In this example, the logical inference is indicated by the use of a factual marker.

For other discussion on logical inferential, see Garrett (2001).

8.4.3.5. Authoritative

Authoritative stance is one of the 5 the E-E macro-categories found in all the Tibetic languages.66 Stance marking is not only related to the speaker’s commitment but also has an interactional function (see Zeisler 2018a-b). Some languages such as Amdo make a distinction between markers indicating common ground, shared experience and, on the other hand, markers claiming the speaker’s authority and emphasizing

66. The term *authoritative* was used by Zeisler (2018a-b) with a slightly different meaning. This author used the term in a more restricted sense similar to our “egophoric” meaning: ‘I know that’, ‘I know well that’, ‘I personally know that’, ‘I am convinced that’, conveyed by the auxiliaries *YIN* and *YOD*.
upon his/her knowledge that the hearer or “co-speaker” is not supposed to know (Simon 2019, Tübingen).

The use of an authoritative implies that the speaker has/had access to information not via senses or inferences, neither via reported speech but either through his/her own knowledge or through general knowledge but also objective information available to her/him. In many languages, the authoritative category is subdivided into “factual” (general knowledge and objective information) and “egophoric” (personal knowledge). The former category is presented as “objective” whereas the latter is “subjective.”

Some rare Tibetic languages lack one of these two categories. For example, Yolmo (Gawne 2017) does not have a factual marker, but has an egophoric type. Balti and Purik possess a general authoritative marker but lack an egophoric category. These two closely related languages have preserved the Old Tibetan verb ཨཡིན། YIN and ཨཡོད། YOD in their original function.

(224) གཉིས། ཀི། རིན།
‘DI’U CI YIN
this what be
‘What is this?’ [neutral question] (Ba)

(225) གཉིས། མིན་ིན།
‘DI’U SKYIN YIN
this ibex be+AUTH
‘This is an ibex.’ (Ba)

(226) དེ་ལས་སུན་ཆོད་པ་ན་ཡོང་ངེད།<YONG.BAYOD>
DE LAS-SUN CHOD.PA-NA YONG-NGED
that work-PL finish-CONJ come-UNCMP+AUTH
‘As soon as these works are done I will come.’ (Pur; adapted from Zemp 2018)

(227) མཉན་ལ་དམུལ་བརྒྱ་ཁྱེར་རེད།<KHYER.BAYOD>
MAN-LA DMUL BRGYA KHYER-RED
man-DAT rupee 100 charge-UNCMP+AUTH
‘They charge 100 rupees for one man[maund].’ (ibid.)
In Purik the simple bare past stem (without auxiliary), usually marked by an /-s/ “tends to be evidentially and epistemically neutral, except that they imply full confidence about the truth of the statement” (Zemp 2018). It corresponds to our "authoritative" function.

In most cases, the use of the simple past in Purik as well as Ladaks implies that the speaker directly experiences, controls or witnesses an event. However, this direct observation or participation is only a default interpretation and is not encoded in the grammar of the simple past. Here are some examples:

(228) འོ་མ་འོ་རྗེན་ལ་བཅོ་འད། (བཅོ་[བ་ཡོད་]<<BO.BA YOD>)

‘Yoghurt is made from milk.’ (ibid.)

(229) མདང་ཀར་གིལ་ལ་ཁ་ཙ་པིག་བཏངས། (MDANG.KAR.GIL-LOC KHA.TSA.PIG.BTANGS)

‘Yesterday it snowed a little in Kargil.’ (La)

The default interpretation is that the speaker saw the snow falling, but it does not need to be the case. He might have received a phone call of somebody he trusts and then convey the information with this authoritative mood. However, in Ladaks, many speakers will prefer to use the inferential form in TOG:

(231) མདང་ཀར་གིལ་ལ་ཁ་ཙ་པིག་བཏངས-ཕོག (MDANG.KAR.GIL-LOC KHA.TSA.PIG.BTANGS-TOG)

‘Yesterday it snowed a little in Kargil.’ (La)

In Ladaks, the authoritative is used for sentences such as ’I was born in X’ which clearly indicates that in that case, no observation can be involved...
PART 2 – CHAP 8. Grammatical outline of the Tibetic languages

(232) གཉེ་རང་ག་རུ་སྐྱེས།

NYE.RANG GA.RU SKYES
2SG where be born+CMP+AUTH
‘Where were you born?’ (La)

(233) གེ་ོི་སྐྱེས།

NGA GLE.-’A SKYES
1SG Leh-LOC be born+CMP+AUTH
‘I was born in Leh.’ (La)

The following variant is also quite frequent:

(234) གེ་ོི་སྐྱེས་ཏེ་ཡིན།

NGA GLE.-’A SKYES-TE.YIN
1SG Leh-LOC be born-CMP+EGO
‘I was born in Leh.’ (La)

The simple past is also used with a 1st person agent (with controllable verbs) when the speaker wants to make a strong emphasis.

(235) གེ་ོི་ིརང་།

NGA-S ZERS
1SG-ERG say
‘I did tell (you)!’ (La)

(236) གེ་ོི་ིརང་མི་ིརང་།

NGA-S ZERS-SA MA-ZERS
1SG-ERG say-NMLZ NEG-say
‘I told (you), didn’t I!’ (La)

Here is an example for Purik, by Zemp (2018):

(237) ཁོ་འརྒྱལ་ལ་སོང་།

KHO.-’A RGYAL.LA SONG
3SG-DAT well go-PAST
‘It went well for her/him.’ [S/he benefits now from it] (Pur)

8.4.3.6. Factual

Many Tibetic languages have developed a factual marker (see e.g. Tournadre 2008; Vokurková 2008; Oisel 2013). The speaker simply presents the information as a fact.
The marker is used for specific facts, known by the speaker as well as gnomic and historical information. For the equative copulative verb, the factual is marked by རེད་ (Ü, Kh, Am, E), ལྡན་ (Ts, Lho), འདག་ (Tö Ngari), འགི་ (E: Thewo mä). In some regions a compound form is used, e.g. འོད་ དར་པོ (La) and འོད་ ལེ་ དར་པོ (Sh). Alternative terms have also been used such as “assertive” (Tournadre 1996a) for Common Tibetan and Lende Kyirong (Huber 2002), or “neutral” (Yliniemi 2019) for Lhoke.

Note that Suzuki and Sonam Wangmo (2018b) have proposed an alternative term “statemental” to describe an auxiliary use in some Kham dialects. This term indicates that the speaker does not specify any access to information whereas the term ‘factual’ may also indicate a logical inference or an authoritative stance.

The existential verbs that indicate the factual are nearly always compound forms usually containing the auxiliary འདག་ (Ü), འོད་ (Kh, Hor), འོད་ལེ་ (Kh, E), འོད་ཅི་ (Am), འོད་བསད་ (Ts), འོད་པརསད་ (Lho), འོད་དཔད་ (Lho), འོད་ ལྡན་ (Tö), འོད་ ལེ་ དར་པོ (E: Thewo mä). The factual auxiliaries are usually made of the above copulative and existential factual verbs.

The neutral question ‘what is (this)’ is conveyed by a factual in most languages. The speaker wants to get a factual information or neutral information and thus makes use of a factual by anticipation: ཁ་ རེ་ རེད། (Ü), ཁི་ རེད (Kh, Hor), ཁི་ རེད་ དགེ་ རྒན་ (Am), ཁང་ བསད་ (Ts, Lho), ཁང་ བཙན་ (Th-m). The speaker presents this information as a fact. S/he does not claim any personal knowledge. [The speaker presents this information as a fact. S/he does not claim any personal knowledge.]

Here is another example of factual:

(238) ཕིསྐྱེ་ བསྐེ་ མདོ་ རེད།

Tshering teacher be+FACT

‘Tshering is a teacher.’ (ComTib, Ü, Kh, Am)
The same meaning is conveyed by the following Sherpa, Lhoke, Tsang, Tö, Jangpa and C. Ladak equivalents:

(239) ◊ ལེགས་རིང་དགེ་རྒན་ཡིན་ཛ།
TSHE.RING DGE.RGAN YIN.DZA
Tshering teacher be+FACT
‘Tshering is a teacher.’ (Sh)

(240) ◊ ལེགས་རིང་དགེ་རྒན་སྦད།
TSHE.RING DGE.RGAN SBAD
Tshering teacher be+FACT
‘Tshering is a teacher.’ (Ts, Lho)

(241) ◊ ལེགས་རིང་དགེ་རྒན་འདག
TSHE.RING DGE.RGAN ’DAG
Tshering teacher be+FACT
‘Tshering is a teacher.’ (Tö)

(242) ◊ ལེགས་རིང་དགེ་རྒན་ཡིན་དག
TSHE.RING DGE.RGAN YIN.’DAG
Tshering teacher be+FACT
‘Tshering is a teacher.’ (LJ)

(243) ◊ ལེགས་རིང་དགེ་རྒན་ཡིན་ནོག
TSHE.RING DGE.RGAN YIN.NOG
Tshering teacher be+FACT
‘Tshering is a teacher.’ (La)

Despite the diversity of forms, all the above languages have developed a factual marker.

8.4.3.7. Egophoric

The notion of egophoric may be defined in the following way: egophoric expresses personal knowledge or intention of the speaker (Tournadre 2008). In other words, “Egophoric evidentiality is therefore about a speaker’s access to her own knowledge” (Gawne 2017).

Let’s note here that some authors, such as DeLancey (2018) and Tribur (2019), differentiate egophoricity from evidentiality.
The notion of egophoric is very similar to "personal knowledge" (van Driem 1998; DeLancey 1990; Caplow 2017; Yliniemi 2017, 2019), "self-person" (Sun 1993), "personal experience" (Huber 2002), "ego evidentiality" (Garrett 2001), "speaker's involvement" (Hein 2007). Here are some examples (244-247) in Common Tibetan from Tournadre & Sangda Dorje (2003).

(244) འདི་མིག་འདི་ངའི་ཡིན།
LDE.MIG 'DI NGA'I YIN
key this I+GEN be
'This key is mine!'

(245) ལྡེ་མིག་འདི་ངའི་བུ་མོ་སློབ་གྲྭ་ལ་འགྲོ་གི་ཡོད།
NGA'I BU.MO SLOB.GRWA-LA 'GRO-GI YOD
I+GEN daughter school-LOC go-UNCMP+EGO
'My daughter goes to school.'

(246) ཕྱོང་གིས་ངར་ཡི་གེ་བཏང་བྱུང་།
KHONG-GIS NGA-R YLGE BTANG-BYUNG
he-ERG I-DAT Letter Send-EGOREC
'He sent me a letter.'

(247) འདི་ཁྱེད་རང་གི་གསོལ་ཇ་ཡིན།
'DI KHYED.RANG-GIS GSOL.JA YIN
this you-GEN tea(H) be+EGO
'This is your tea [The tea I made for you].'

It is important to note that in the above example, it is always possible to replace the egophoric by a factual marker.

Egophoric markers are found in nearly all the Tibetic languages with some rare exceptions such as Balti and Purik. However, as we will see below their semantic extension may greatly differ from a language to the other.

Although the category of egophoric was first applied to Common Tibetan, this category has recently attracted a growing attention in the community of linguists who describe languages belonging to other language families (Hyslop 2011; Daudey 2014; San Roque et al. 2018, etc.). The term "égophore" was coined by Claude Hagège in
PART 2 – CHAP 8. Grammatical outline of the Tibetic languages

1982, and meant that "ego" as a deictic center is a fundamental property of linguistic systems. Égophore is a hyperonym for a few other notions related to the deictic center: chronophore, exophore, and endophore, which included the subcategories of autophorique, anaphorique, cataphorique and logophorique. The term "égohphore" used by Hagège (1982) did not refer to the grammatical phenomenon now known as "egophoric" in the Tibetic languages and did not apply to any particular language or language group. In 1991, N. Tournadre first applied the term "egophoric" to Common Tibetan with an entirely different definition and thus coined a new concept to describe a specific phenomenon in this language. For a discussion on terminological issues related to egophoric and egophoricity, see Gawne & Hill (2017) and Gawne (2017).

Origin of the egophoric markers in the Tibetic languages

In the Tibetic languages which have grammaticalized egophoricity, the main egophoric auxiliaries are derived from the copulative verb ཡིན་ YIN ‘to be’ and the existential verb ཡོད་ YOD ‘to exist’, ‘to have’. More marginally, in some languages such as Ü and Tsang, other verbs which have acquired egophoric meanings include བྱུང་ BYUNG, ལྟོང་ MYONG, ོགས་ DGOS, གོག་ CHOG and དབང་ YONG. In Amdo an egophoric marker -a, of unknown origin is attested.

Egophoricity and access to information

Egophoric may also be described in terms of access to information. The access is not sensory nor inferential but is related to "self-awareness" (Tournadre & LaPolla 2014). This type of access has not yet received sufficient attention. To explain this type of access, let’s us provide a simple example: if a person is sitting on her bed in the dark in the middle of the night and asked by her partner what she is doing, she may answer: ‘I am thinking about my project’. The speaker’s access is not sensory (since it is in the dark and it is a mental activity) and only possible through the speaker’s “self awareness.”

67. “Le système de l’égophore est une propriété capitale des énoncés linguistiques.”
In some Tibetic languages, this access is grammaticalized as an egophoric (see Tournadre 2008; Gawne & Hill 2017).

\[
\text{NGA \hspace{0.5cm} BSAM.BLO \hspace{0.5cm} BTANG-GI.YOD} \\
\text{1SG \hspace{0.5cm} reflexion \hspace{0.5cm} LV-UNCMP\#EGO (INT)}
\]

'I am thinking (about it).’ (Ü, ComTib)

**General characteristics of egophoric markers**

As mentioned earlier, evidential and epistemic categories feature essentially in the main clause. In non-finite clauses, one encounters essentially egophoric markers (\textit{YIN, YOD, BYUNG}, etc.), but they are used in a neutral way and do not convey any egophoric meaning.

Egophoric markers do not only insist on the personal knowledge, self-awareness and intention, they may also convey a more emotional load, whereas by contrast the factual marker presents the information as "objective" (see also Yliniemi 2019).

There are several types of egophoric both within a single language and across the different languages. For example, Common Tibetan includes several kinds of egophoric depending on semantic parameters such as intentionality or aspect: \textit{intentional, receptive, habitual, experiential} and \textit{benefactive} (see 8.3.5 and Tournadre and Sangda Dorje 2003).

Moreover, it is important to make a distinction between \textit{loose egophoric or wide scope egophoric} and \textit{strict egophoric or narrow scope egophoric}. The alternative terms of \textit{weak egophoric} and \textit{strong egophoric} have also been used (Gawne 2017; Tournadre 2008, 2017). The idea of a difference in scope was suggested by Garrett (2001: 107) who used the terms "strong ego" and "weak ego."

**Strict egophoric**

Egophoric auxiliaries in the strict sense are used in declarative sentences (for interrogative sentences, see the anticipation strategy, 8.4.8.3) normally only with the first person singular or plural occurring overtly or covertly, regardless of its function in the given clause – “subject”, “object”, “indirect object”, “genitive complement”, “locative complement” (see Tournadre 2008, 2017). So, in the case of strict egophoricity, the
correlation with the first person is very strong. There are of course some marginal exceptions and the use of egophoric with the 2nd and 3rd persons in declarative sentence is not impossible but requires very specific situations in which the speaker insists strongly on his personal or intimate knowledge. In unmarked situations, the egophoric is usually not compatible with 3rd persons in declarative or interrogative sentences. For example, in Common Tibetan:

(249) སྐྱེ་ རུལ་བའི་མིན།  
NGA PHYUG.PO MIN  
1SG rich EGO+NEG  
'I am not rich.'

(250) བློ་ རུལ་བའི་ཨུད་པས།  
KHONG PHYUG.PO MIN-PAS  
3SG rich EGO+NEG-Q  
'Is he not rich?'

As we will see in the case of loose egophoricity, the correlation with the 1st person is less strong and the auxiliaries are also used with 3rd and 2nd persons, and the above sentence would be completely acceptable.

The strict egophoric markers are found essentially in Tibet, e.g. in Amdo, Northern Kham, Hor, Ngari, Ú, Tsang, and more rarely outside, e.g. Sherpa. (See for example Bartee 2007; Robin & Simon, forthcoming; Denwood 1999; Tournadre & Sangda Dorje 2003; Häsl er 1999; Huber 2002, etc.)

**Loose egophoric**

Loose (or weak) egophoric auxiliaries are found in languages spoken outside Tibet in the southern and western Himalayas such as Dzongkha, Lhoke, Choća-ngaća, Yolmo, etc. They have been described as conveying "old knowledge", or "personal knowledge" by van Driem (1998) or Yliniemi (2019). For Dzongkha, van Driem (1998) provides the following explanation:

69. Tournadre (2017) claimed that "Yolmo, just as Ladakhi and Dzongkha do not have an egophoric category per se." The statement was too strong. The intended meaning was that these languages do not have a strict egophoric contrary to many languages and dialects of Tibet.
“the form ing expresses old, ingrained background knowledge or has become a firmly integrated part of one’s conception of reality […].”

Lhoke:

(251) ◊ མིན་པོ་ཁྲིམས། KHONG SMAN.PO ʔIN
3SG doctor be
‘He is a doctor.’ (Yliniemi 2019)

(252) ◊ མིན་པོ་ཁྲིམས། LHAN.RGYAS SMAN.PO ʔIN
2SG(h) doctor be
‘You are a doctor.’ (ibid.)

(253) ◊ མིན་པོ་ཁྲིམས། NGA SMAN.PO ʔIN
SG doctor be
‘I am a doctor.’

(254) ◊ མིན་པོ ཕྲིམས། SMAN.KHANG-NA SMAN.PO KARMA YOD
hospital-LOC doctor Karma Exist
‘In the hospital there is Doctor Karma.’

Dzongkha:

(255) ◊ རིང་ཐུང་ཤེས་ DRUNG.YIG ʔIN
DEM monk-DEF clerk be
‘That man is a clerk.’ (van Driem 1998)

(256) ◊ མིན་པོ་ཁྲིམས། NGA DGE.SLONG ʔIN
1SG monk be
‘I am a monk.’ (ibid.)
Ladaks:

(257) དབྱར་ལ་ མ་ན་ལི་ཡ་ ལཱི་རྒྱལ་ མང་པོ་ ཡོད།
DBYAR.LA MANALI.YA PHYL.RGYAL.PA MANG.PO YOD
summer Manali-LOC tourist many exist

‘In summer, there are a lot of foreign tourists (Lit. ‘foreigners’) in Manali.’
[I know well, I have personal information] (La)

KHONG DGE.RGAN YIN
3sg(h) teacher be
‘He is a teacher.’ [I know very well] (marked sentence contrasting with DGE.RGAN YIN.NO)

(259) ◊ བླ་མ་ཀུན་ནིས་ སྐུ་རིམ་ སལ་ལ་ཡོད།
BLA.MA-KUN.NIS SKU.RIM SAL.LA.YOD
lama-PL-GEN ritual give(h)-UNCMP+EGO
‘The monks are performing a ritual [in my home].’ (Example adapted from Koshal, 1979.)

The example below is a little awkward since both the factual YIN.NO and the reportive-inferential YIN.KYAG would both be preferable, but in a very marked utterance, it is not impossible:

(261) ◊ སེང་གེ་རྣམ་རྒྱལ་ ལ་དྭགས་སི་ རྒྱལ་པོ་ ཡོན།
SENG.GE RNAM.RGYAL LA.DWAGS-SI RGYAL.PO YIN
Sengge Namgyal Ladakh-GEN king be
‘Sengge Namgyal was a king of Ladakh.’ [intended meaning: I know personally very well.] (La)

Yolmo:

(262) མ་གི་ ཡིན་པ།
MA.GI YIN.PA
corn COP+EGO
‘It is corn.’ (Gawne 2017)
Even within the realm of loose egophoric, we can observe various degrees of grammaticalizations and restrictions. As pointed out by Gawne (2017):

“Yolmo does not have the same ‘general fact’ [factual] category that is found in Common Tibetan and Tibetic varieties. This is possibly a factor in why the Yolmo egophoric form has a broader distribution than the Common Tibetan cognate. Of all the languages with an egophoric, in Yolmo, the personal knowledge’ component is much weaker than in varieties where it is in contrast with a clear factual, gnomic or ‘non-egophoric’ category.”

Concerning Lhoke, one of the official languages of Sikkim, she adds:

“the Denjongke [Lhoke] personal forms, equative ĩʹ [IN] and existential ĭʹ [IN] [YOD] are constricted with the familiar sensorial duʔ [DU] but also with a neutral [factual] copula beʔ [SBAD]. This means that while the forms have a quite broad distribution like other Southern Tibetic languages, the neutral form covers some of the semantic space that the egophoric covers in Yolmo. For example, in Lamjung Yolmo, a speaker can use the egophoric to talk about historical events as there is no other evidential form that is preferred. In Denjong, however, ‘it seems impossible to gain personal knowledge of distant historical events’ (Yliniemi 2017: 317-318) and instead the neutral form is used.”

Thus the precise semantic extensions of the egophoric largely depend on each language.

8.4.3.8. Reported speech and hearsay

We use here “hearsay” to refer to reported information without mention of the source (or lacking a precise source) and “reported speech” or “quotation” when the source of information is either explicitly mentioned or clearly identifiable. In the case of hearsay as well as in the case of reported speech, a verb of speech or a reported speech marker is necessarily present.

In a few dialects, we additionally find a “reportive-inferential modality.” The speaker may rely on his own inference or a reported information but it remains implicit and does not normally entail the presence of any reported speech marker. We will first examine the reported speech and hearsay.
a) Reported speech

Since Tibetic languages are all verb final and normally manifest a neutral order SOV, the reported clause is usually embedded between the author reporting the quotation and the verb of speech which occurs in the sentence final position. From a syntactic point of view, the reporting author is marked by a case (usually the ergative). The reported clause is not introduced by any marker (as expected for a verb final language), but is closed by a marker indicating the end of the quotation or reported speech final marker (hence RSF). This RSF is normally followed by a verb of speech, but in some cases, the verb may be used alone. Conversely the verb of speech is often dropped and the RSF may appear in the sentence final position. Let us summarize here the main reported speech constructions frequently attested in the Tibetic languages:

(Source-ERG) – (Goal-DAT) – “Quotation” – RSF – Verba dicendi
(Source-ERG) – (Goal-DAT) – “Quotation” – Verba dicendi
(Source-ERG) – (Goal-DAT) – “Quotation” – RSF
“Hearsay” – RSF

Here are some examples that illustrate the above structures:

(263) ◊ བདེ་ མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ སྦྱངས་ན་ ཡག་པོ་ ཡོད་པས་ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ་ ཟེ
people all-ERG English learn-CONJ good exist+FACT-RSF say-SENS
‘Everybody says that it is good to learn English.’ (Ü, ComTib; Mélac 2014)

(264) ◊ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི་ རིག་མི
3SG+ERG what think-ENDOR-RSF
‘What did he say he was thinking?’ (La)

(265) ◊ རིག་མི
3SG+ERG what-RSF-H
‘What did you say? / What was said?’ (La)

In the above sentence, the RSF functions as a main verb but unlike the latter may not be followed by TAM markers.
The following example illustrates the fact that there may be a recursivity in the reported speech. Somebody said that X said that.

(266) (He) says that the abominable snowman says, “bye bye”. (Lit. ‘go slowly’) *(ibid.*) *(Ü, ComTib)*

In Melac’s example above, we can see that there are two reported speakers: the “abominable snowman” and another person (he) who reported what was said by the snowman. It is worth noting that both the source(s) and the goal(s) may be deleted.

In some cases, the quotative marker may be used alone without a verb of speech, whereas in other cases, the verb of speech can occur without any quotative marker.

In CT, the RSF is the morpheme ཇེས ZRES, which has two allomorphs ཁེས CES and གེས SHES. However, this marker does not seem to be attested in the modern languages. In a number of languages (Ü, Am, Dz, etc.), the RSF /se/, /sa/ (Am, Kh), /s/, /sa/ (Ü, Ts) is derived from the verb of speech རྐྱེད ZER ‘to say’ found in CT and in many languages. In Amdo, for the past, a suppletive form derived from the verb འབིགས BZLAS. In languages such as Ladaks, Zanhar, Purik, Kyirong, Lhoke or Dzongkha, another form derived from the CT form བོ་ LO is used to indicate reported speech and/or hearsay. (For examples of བོ LO in CT, see Kesang Gyurmé 1992). Another RSF marker /tʂa/ (Kh: Tormarong), is derived from the CT verb གྲག་ GRAG ‘to sound’ (see Bartee 2007), so are /tɕɑʔ/ and /cɑʔ/ found in Gyalthang Kham (Suzuki 2014a).

As noted by Driem (1998: 400), the RSF may be used not only with verbs of speech (such as བོ་ LAB, རྐྱེད ZER, བཤད BSHAD, འབིགས BZLO) but also with verbs of perception (such as དུས GO ‘hear’) and psychological verbs (such as མོའོ MNO ‘think’, བསམ་ BSAM ‘think’).

In many languages of the world, utterance complement clause may be expressed either by direct or indirect speech constructions. In “direct speech constructions” the quotation is reported by the actual speaker verbatim, i.e. using original wording (or pretending to use it), whereas in “indirect speech construction” the reported utterance
is rephrased by the actual speaker and the deictic words are recalculated with reference
to the speaker’s current situation. Cristofaro (2013) notes that “in a number of
languages direct speech is the only means available to convey reported speech.”

In the Tibetic languages, two types of reported speech constructions are attested.
However, the distinction between direct speech and indirect speech does not exactly
match the equivalent in the European languages. This has been noted by various authors.
Van Driem (1998: 404) has shown that “sometimes, it is only the context which dis-
ambiguates between direct and indirect speech.” According to Koshal (1982) “Ladakhi
does not distinguish between direct speech and indirect speech the way English or
Hindi does.”

In many cases, there is no alternative between direct and indirect speech as shown
in the examples below:

(267) ◊ བོལ་བཟང་ལགས་ཀྱིས་ཚེ་རིང་ཨེམ་ཆི་རེད་ཟེ་གསུང་སོང་།
BLO.BZANG-LAGS-KYIS TSHE.RING ʔEM.CHI RED-ZE GSUNG-SONG
Lobzang-H-ERG Tshering doctor be(FACT)-QUOT tell-CMP+SENS
‘Lobzang said: “Tshering is a doctor.”’ [direct speech] or
‘Lobzang said that Tshering was a doctor.’ [indirect speech] (Ü, ComTib)

From a syntactic point of view, the above Tibetan sentence corresponds to direct
speech.70 However, there is no indirect speech equivalent.

An alternative reported speech construction is available only when there is a corefe-
rence between the author of the reporting speech and a participant of the reported
speech. For example, in the sentence below, the pronoun KHONG’s/he’ and the
pronoun NGA’I’ refer to the same person.

(268) ◊ གོང་གིས་ང་ཨེམ་ཆི་ཡིན་ཟེ་གསུང་སོང་།
KHONG-GIS NGA ʔEM.CHI YIN-ZE GSUNG-SONG
3SG(H)-ERG ISG doctor be(EGO)-QUOT tell-CMP+SENS
‘S/he, said: “I, am a doctor.”’ [direct speech] (Ü, ComTib)

70. Although the sentence has a quotative marker (ZE) that could be interpreted as a comple-
mentizer, there are functional reasons (related to deixis) to consider that it is rather a form of direct
speech.
However, Tibetic languages tend to avoid coreferential pronouns or NPs. Thus, another construction is often preferred to convey the same information. This construction has been labeled 'hybrid indirect speech' (see Tournadre 1992; Tournadre & Sangla Dorje 1998, 2003; Zemp 2018). Some authors have described this phenomenon in the framework of the conjunct/disjunct model (see e.g. DeLancey 1990, 1992, see 8.4.4.2).

Thus, the sentence above is more often rephrased as:

(269) "S/he said that s/he was a doctor." [hybrid indirect speech] (Ü, ComTib)

The literal translation of this sentence would be: "S/he am doctor, said."

As we see, the coreferential pronoun NGA.T has been deleted and the ergative case marking the pronoun KHONG has also been deleted since it is now governed by the verb of the embedded clause, here the copulative verb YIN.

In such constructions, the personal pronouns and other deictic markers are reformulated according to the speaker's current situation, as expected in the case of indirect speech. The honorific markers are also reformulated according to the actual speaker's perspective (see ex. below). However, one of the distinctive features of the 'hybrid indirect speech' is that it preserves the original evidential and epistemic auxiliaries used in the quotation. This is true even for egophoric auxiliaries that are normally related to the 1st person.

Here are some additional examples:

In Common Tibetan:

(270) 'Lobzang, says: "I, will come."

(271) 'Lobzang, says that he, will come.'
In Amdo:

(272) ◊ གུར་གེ་བོད་ཟིག་མིན་ཟེར་གི KHUR GE BOD-ZIG MIN ZER-GI
/k’arge wo’-zaq man ser-ka/ 3SG+ERG Tibetan-IND NEG+COP+EGO say-PROG+SENS
‘He says that he is not Tibetan.’

(273) ◊ གུར་གེ་བོད་ཟིག་མ་རེད་ཟེར་གི KHUR GE BOD-ZIG MA-RED ZER-GI
/k’arge wo’-zaq ma-ra ser-ka/ 3SG+ERG Tibetan-IND NEG+COP+FACT say-PROG+SENS
‘He says that he is not Tibetan.’

b) Hearsay

Hearsay constructions are made with the same RSF, མིན་ZER, བྲུས་BZLAS, ལོ་LO, དབུག་GRAG ‘to sound’ (see above) as the quotative markers. The main difference in the construction between hearsay and quotative is that the argument corresponding to the reported speaker is dropped.

However, since the reported speaker can also be dropped in the case of quotation, the two constructions are syntactically equivalent and it is not always easy to differentiate between hearsay and quotation.

Here are some examples of hearsay in various Tibetic languages:

In Common Tibetan:

(274) ◊ བལ་ཡུལ་ལ་ས་ཡོམ་མང་པོ་ཡོད་རེད་ཟ། BAL YUL LA SA YOM MANG PO YOD RED ZA
Nepal-LOC earthquake many exist-HS
‘There are reportedly many earthquakes in Nepal.’

In Kyirong:

(275) ◊ དུད་པ་སྒམ་མཁན་ལོ DUD PA SGAM MKHAN LO
smoke inhale-UNCMP-HS
‘He inhales smoke, it is said.’ (Huber 2002) [‘They say he smokes.’]
In Dzongkha:

(276) ◊ ཡོན་ཏན་དགའ་བས་ལོ།

KHO-GIS NGA-LU DGA'-BAS-LO
3SG-ERG 1-DAT love-INF-HS
'I have been told that he loves me.' (van Driem 1998: 196) ['I heard that he loves me.]

In Purik:

(277) ◊ མ་ཆེན་འདུག་ལོ།

KHO KHANG.MA CHEN'-DUG-LO
3SG home go+PROG-SENS-HS
'I hear she is going home.' [Somebody who saw her going home said it to me] (Zemp 2018)

In Minyak Rabgang Kham, the hearsay marker is ZER.RED. Here is an example from Lhagang Kham (Suzuki & Sonam Wangmo 2015):

(278) ◊ བོ་ང༌ང་ཅིག་ལེན་རེད་རེད།

JO.BO-"U KHA GRAGS-ZIN.RED-ZER.RED
Jowo-ERG speak-AOR-HS
'It is said that the Jowo spoke words.'

In the Amdo sentence below, the hearsay marker is BZLAS 'to tell' followed by a secondary verb (BTANG) and the sensory suffix THAL which indicates that the actual speaker heard it directly.

(279) ◊ དཔོན་པོ་ཁ་རྩང་ཟི་ལིང་ང་བུད་ཐལ་བཟླས་བཏང་ཐལ།

TSHE.RING KHAL.RANG ZILLING-NGA BUD-THAL BZLAS-BTANG-THAL
/ts'erang k'artsa nga wa-t'a zl[PAST]-tang-t'a/
Tshering+ABS yesterday Xining-DAT go[PAST]-CMP+SENS QUOT-SEC-SENS
'S/he said that Tsering went to Xining yesterday.' (Am; Tournadre & Shao forthcoming)

71. CHEN'-DUG-LO is the contraction of དཔོན་ལིང་རེད་དེ། CHAYIN.'DUG-LO.
Compatibility of hearsay and quotative with evidential markers

In the Tibetic E-E systems, quotative as well as hearsay have a special status, since they are compatible with other evidential categories such as sensory, inferential, egophoric, etc. (see Sun 1993; Tournadre 1996a-b; Tournadre & Konchok Jiatso 2001; Mélac 2014). Sun (1993: 991) noted for Amdo that "the quotative morpheme [...] is on both categorial and distributional count, at variance with the other three evidential markers [direct, indirect and immediate evidential (in our terminology, respectively to sensory, inferential and endopathic)]."

Indeed, reported speech and hearsay provide important arguments to distinguish “access to information” and “source of information” (see Tournadre & LaPolla 2014). As mentioned earlier, quotative markers and hearsay refer to a verbal source of information, whereas sensory, inferential and egophoric markers essentially convey access to information.

The examples below from Sun (1993) illustrate the compatibility of the hearsay (as well as quotation) with other evidential markers:

In Amdo:

(280) ◊ མྱེ་ཤོར་བུད་ཐལ་ཟེར།
KHAR NUB MYE SHOR-BUD-THAL-ZER last night fire slip-away-SENS-HS
'I heard (from someone who saw it happen) that a fire broke out last night.' (adapted from Sun 1993)

(281) ◊ མྱེ་ཤོར་སོང་ཟིག་ཟེར།
KHAR NUB MYE SHOR-SONG.ZIG-ZER last night fire slip-away-INF-HS
'I heard (from someone who did not see it happen but saw the traces of the fire) that a fire broke out last night.' (ibid.)

As we have seen above, the hearsay or quotative marker is normally preceded by an evidential marker, which specifies the access to information of the reported source. Note that it does not tell anything about the access to information of the actual speaker, i.e. how she had access to the hearsay. The actual speaker may also specify how she had access to information.
Lobzang, says that he, will come.' (Ü, ComTib)

Again the Purik ex. 277 mentioned above illustrates the same phenomenon:

'I hear she is going home.' [Somebody who saw her going home said it to me] (Pur; adapted from Zemp 2018)

The disclaimer 's'

Finally, in some languages such as Common Tibetan, the RSF is also used after single words, expressions or whole sentences as a “disclaimer” to indicate that the reader or the speaker does not take the responsibility for the utterance. This habit is so deeply rooted that speakers in Central Tibet would systematically add an ལེར་/s/ < CT རེ་ ZER 'to say' when answering question about elicitated words or expressions. In Ladakh, students of English will usually add the marker ཨ་/lo/ even when reading an English sentence (Rebecca Norman, pers. comm. 2017).

c) The reportive-inferential modality

This category is not widespread within the Tibetic languages, but it is interesting from a typological point of view. It is found in some western languages such as Ladaks. Unlike the reported speech and hearsay, the reportive-inferential modality does not explicitly mention reported speech, nor does it include any verb of speech, but rather implies that the speaker had access to the information through media or, in some cases, through his own logical inference. It may be rendered in English as 'According to what I heard or read, or as far I know'. In a way, from a cognitive point of view, the reportive modality functions more like an access marker than a source marker since it is not concerned with a precise source of information but rather with the cognitive access to
information, which is not sensory but mediated. As we will see, a strong confirmation of the fact that reportive-inferential modality functions as an access marker will appear clearly from the morphological distribution.

Here are some examples:

(284) ◊ སངས་རྒྱ་སི་ས་དུས་འཁོར་དབང་ཆེན་སལ་ཀྱག SANGS.RGYAS-SIS DUS.'KHOR DBANG.CHEN SAL-KYAG Kālacakra give(h)-CMP+REPINF

’[According to tradition] the Buddha has taught the Kālacakra initiation.’

(285) ◊ སེང་གེ་རྣམ་རྒྱལ་ལ་དྭགས་སི་རྒྱལ་པོ་ཡིན་ཀྱག SENG.GE RNAM.RGYAL LA.DWAGS-SI RGYAL.PO YIN.KYAG Sengge Namgyal Ladakh-GEN king be+REPINF

’Sengge Namgyal was a king of Ladakh.’ [according to what I read in history books] (La)

These two examples are adapted from Koshal (1979: 191, 207). This author describes these as "narrative forms." She does not use the term "reportive-inferential" for this function nor does she mention the fact that it is based on second hand information, but her comments are useful: “The suffix -kək[kyak] is used in narrations. In such cases, -kək[kyak] implies a certain degree of uncertainty about the veracity of the statement as the speaker cannot himself vouch for it. -kək[kyak] forms are really indifferent to the temporal distinction of present and past as they express uncertainty about an event. -yinkək[yinkyak] expresses a higher degree of uncertainty than -yotkək[yotkyak].”

There is indeed often some uncertainty in the reportive-inferential forms which is inherent to second hand information or to some types of inference. The degree of certainty completely depends on the credibility that the speaker gives to the sources on which s/he relies.

---

72. Strangely enough she describes the simple authoritative form yod as reportive which is not at all suitable. However, her study at the end of the 1970s was a pioneer work in the field of Ladakhs linguistics.

73. It is not clear whether S. Koshal used a dialectal pronunciation or if it is due to an evolution of the language, but nowadays, people in Leh pronounce this marker /-kyak/ and not /kək/. Many dialects of the area however still pronounce /kak/ (but not /kək/).
For example:

(286) ཆོས་ལྡན་གཅིག་ ལག་ཡོངས་ དབྱ་རིའི་ མང་པོ་ འྱི་རྒྱལ་པ་ མོང་པོ་ འོད་ཀྱི་

summer Manali-LOC tourist many exist+REPINF

ʼIn summer, there are a lot of foreign tourists (Lit. 'foreigners') in Manali.’
[from what I hear] (La)

In front of such sentences in the reportive-inferential modality, the speaker can add:

(287) ང་འཚོར་ སེས་ལ་ དབྱར་ལ་ མ་ན་ལི་ཡ་

summer Manali-LOC tourist many exist+REPINF

ʼAccording to what I hear, in summer there are a lot of foreign tourists
(Lit. 'foreigners') in Manali.’ (La)

In some cases, the second-hand information or inferential meaning conveyed by

YOD.KYAG may be weakened and interpreted as an equivalent of འདུག་

YOD.KANO.G or འདུག་

YOD.DE.YINO.G which convey a factual or gnomic meaning. This

latter form, YOD.KANO, is rarely used in Leh. It has a dialectal flavor. Thus,

somebody living in Leh may perfectly say:

(288) ཆོས་ལྡན་གྲེ་བའམ་ ལག་ཡོངས་ དབྱ་རིའི་ མང་པོ་ འདུག

summer Leh-LOC tourist many exist+REPINF

ʼIn summer, there are many tourists in Leh!’ [general statement] (La)

Here, of course the speaker has a complete certainty about his/her statement. It

would still possible for this sentence to add 'as people say' or 'as far as I know’, there are

many tourists in Leh. The speaker could also emphasize her own visual knowledge:

(289) ཆོས་ལྡན་གྲེ་བའམ་ ལག་ཡོངས་ དབྱ་རིའི་ མང་པོ་ འདུག

summer Leh-LOC tourist many exist+VIS

ʼIn summer, there are many tourists in Leh!’ [I see it all the time] (La)
This weakening of the "second hand" or inferential information meaning is not present with the form རྩེ་བུག་ YIN.KYAG, which is opposed to the factual form རྩེ་ནོག་ YIN.NOG.

Compare for example:

(290) ◊ བླ་མ་ཀུན་ནིས་སྐུ་རིམ་སལ་ལ་ཡིན་ཀྱག་ KHONG DGE.RGAN YIN.KYAG
3SG teacher be+REPINF
'He is a teacher.' [from what I heard, as far as I know] (La)

(291) ◊ བླ་མ་ཀུན་ནིས་སྐུ་རིམ་སལ་ལ་ནོག KHONG DGE.RGAN YIN.NOG
3SG teacher be+FACT
'He is a teacher.' [It is a fact, everybody knows it] (La)

Let's give another example:

(292) ◊ བླ་མ་ཀུན་ནིས་སྐུ་རིམ་སལ་ལ་ནོག BLA.MA-KUN-NIS SKU.RIM SAL.LA.YOD.KYAG
lama-COL-ERG ritual give(H)-UNCMP+REPINF
'lama-kun-nis kurim sal-la-yod-kyak/
'The monks perform rituals.' [in the neighbor's house, I have been told] (La)

(293) ◊ བླ་མ་ཀུན་ནིས་སྐུ་རིམ་སལ་ལ་ནོག BLA.MA-KUN-NIS SKU.RIM SAL.LA.NOG
lama-kun-i kurim sal-anok/
larma-COL-GEN ritual give(H)-UNCMP+FACT
'The monks perform rituals.' [it is their profession, gnomic] (La)

Finally, note the reportive-inferential modality as other evidential categories is perfectly compatible with a mark of reported speech:

(294) ◊ བླ་མ་ཀུན་ནིས་སྐུ་རིམ་སལ་ལ་ནོག KHONG DGE.RGAN YIN.KYAG-LO
3SG(H) teacher be+REPINF-HS
'She told me that, as far as she knows, he is a teacher.' (La)
8.4.3.9. Mnemic, self-corrective and other marginal categories

The richness and complexity of the Tibetic E-E systems is not restricted to the categories that we have examined so far. In some languages, one finds other categories that play a marginal role but tell a lot about the semantico-cognitive and pragmatic principles that govern these systems (see e.g. Oisel 2017a). Most of these categories have not received sufficient attention. Yet they are an integral part of the E-E system and could shed light on the functioning of these systems.

For example, some languages have developed a form called "self-corrective" (Tournadre & Sangda Dorje 1998, 2003), "erroneous belief" (Huber 2002) or "counter-expectation" (Zeisler pers. comm. 2019). This author provides the following explanation: "I thought X was Y but it isn’t." Here is a Kyirong example (Huber 2002: 143).

(295)  དཔོན་དགེ་རྒན་ཡིན་ནོག་ལོ།
KHONG DGE.RGAN YIN.NOG-LO
3SG(H) teacher be+FACT-HS
‘He is a teacher, I heard.’ (La)

Another interesting category is the mnemic (Tournadre & Sangda Dorje 1998, 2003; Vokurková 2008, 2018). In Common Tibetan, it is marked by a series of
endings such asཡིན་པ་ཡོད་YIN.PA.YOD andཡོད་པ་ཡོད་YOD.PA.YOD (not to be confused with the counter expectative or self-corrective, see above). These markers indicate that the speaker has only a vague recollection of what he is saying. Vokurková (2008: 198) describes the mnemonic in the following way:

“The speaker remembers something but he is not absolutely sure because, often, some time has elapsed since it happened, therefore, they can be translated in English by such expressions as ‘I remember that (perhaps)’ or ‘I think that it is like this (but do not remember it well)’.”

This type is quite common in the spoken language of Lhasa though less frequent than some other types of epistemic endings (e.g. YOD, ‘GRO, YOD.PA.’.

Bartee (2007) mentions the existence of the marker /dʒã/ in gTormarong (alias Dongwang) Kham with a similar meaning: “the validational dʒã indicates that the speaker has a vague recollection regarding the statement s/he is making.” The following example was recorded by Vokurková (2008).

(299) གྲི་ངས་འཁྱེར་ཡོད་པ་ཡོད། GRI.NGA-S.KHYER-YOD.PA.YOD
knife 1SG-ERG bring-PERF+EPII+EGO
‘I’m pretty sure I brought that knife.’ (Ü, ComTib)

The author uses the gloss: EPII+EGO (egophoric + epistemic) to render the notion of mnemonic.

She provides the following comment: “For a picnic, the speaker has brought a lot of things but he is not absolutely sure whether he has the knife” and adds (ibid.): “Although the speaker is quite sure when uttering the above sentence, he may follow: “Oh, I haven’t, I am wearing another jacket today. It’s in the other one.”

Similar effects are obtained in Ladaks with the marker ཉི་-PIN < CT PAYIN.

(300) ◊ གི་ཞག་ཞིག་དེ་ཞག་ཞིག་MTHONG-PIN
DE.ZHAG.ZHIG see-PAST+MNEM
‘I think I saw it a while ago.’ (La; Norman, pers. comm. 2019)
(Note that བིན་ -PIN in its normal intentional use is not compatible with the uncontrollable verb ‘to see’.)

Various marginal categories as the mnemonic or the self-corrective are attested in the Tibetic languages, but they are hard to detect because their frequency is a lot lower than major evidential and epistemic markers. Yet, such categories tell a lot about the grammatical inventivity and sensitivity of the communities speaking Tibetic languages.

For example, people would generally not ask questions to themselves as they would ask to other people. In European languages, the question ‘Where did I put my key?’ would not necessarily have a different morphosyntactic treatment depending on whether the question is asked by the speaker to the hearer or to himself. The intonation might be enough. However, in Tibetic languages, the two situations will yield different sentences. In the former situation, the speaker has to anticipate the hearer’s access to this information (sensory, inferential, etc.). In the latter situation, the speaker just wonders and has no access to this information:

(301) བིན་ ལེ་ལེའི་ སྐྱར་ རབ་མོར་ བཞག་སོང་། /འདུག

\[
\text{NGA-S LDE.MIG G.A.PAR BZHAG-SONG -DUG} \\
\text{1SG-ERG key where put-CMP+SENS -PERF+SENS}
\]

‘Where did I put my key?’ (Ü, ComTib)

Using ལང་ SONG the speaker supposes that the hearer has witnessed the scene but འདུག་ DUG implies that the hearer has seen the key or may infer the location of the key on the basis of visual clues.

In the next example, the speaker may just talk to himself.

(302) བིན་ ལེ་ལེའི་ སྐྱར་ རབ་མོར་ བཞག་ཡོད་འགྲོ

\[
\text{NGA-S LDE.MIG G.A.PAR BZHAG-YOD.'GRO} \\
\text{1SG-ERG key where put-PERF+EPI}
\]

‘Where did I put my key?’ (Ü, ComTib)

Again, asking a question as simple as ‘who is s/he?’ will trigger a different formulation depending on whether it is addressed to a hearer or to one self.
The pragmatic and cognitive parameters thus play a fundamental role in the verb morphosyntax. Let’s give a final example, which corresponds to a very specific form of future described by van Driem (1998: 363) for Dzongkha:

“"There is a special future form, the autolalic future, which expresses the intent of the first person subject. The autolalic future is only used when thinking to oneself in Dzongkha about what one intends to do. The form is never uttered, unless one is talking to oneself, and is always in the first person singular. The autolalic future also occurs in narrative, in direct quotation of someone’s thoughts, and is marked by the ending –ge-no, which originally derives from the adhortative suffix.”

The main differences in the evidential systems

Although the various Tibetic languages share fundamental features in their evidential systems, they also differ in a number of ways. For example intentionality interacts in a significant way with the evidential system of Central Tibet, Amdo or Ladaks whereas it is not grammaticalized in some languages such as Dzongkha.
If we only consider the core evidential categories of the CEV, we find significant discrepancies. As shown, for example in the chart of section 8.3.3.4, we can see that from Central Tibet to Baltistan, we have only four types of core evidential systems:

(a) a threefold system with: sensory vs. egophoric vs. factual (Phânpo, Lhasa, Zhikatse, etc.) expressed by different morphemes but functionally similar;

(b) a fourfold system with: visual sensory vs. non-visual sensory/endopathic vs. egophoric vs. factual (Tô, Spiti, Ladakh Janthang, Ladaks, Zanhar). It is interesting to note that these dialects have not only a similar functional system but are very similar in their morphology.

(c) Then further west, we find a threefold system with: visual sensory vs. non-visual sensory/endopathic vs. authoritative (East Purik). This system mainly differs from the Ladaks system by the absence of egophoric.

(d) Finally in the west, we find a twofold system with external sensory vs. endopathic/authoritative (West Purik and Balti).

We remind that all the above categories in a), b), c) and d) are compatible with hearsay markers.

Additionally, as we have seen above, the evidential systems also greatly differ in their authoritative and egophoric categories. (See 8.4.7.)

8.4.3.11. Source and accessibility hierarchies
In rich evidential-epistemic systems such as those found in the Tibetic languages, the speaker often has various types of access to information related to various sources (reported speech) or to his/her own sensory perceptions or inferences.

There has been some discussions about the speaker’s preferences concerning the choice between inference or reported speech.

Some authors have proposed the hierarchies (see Faller 2002)

INF > QUOT / REPORT

Whereas others favor:

QUOT / REPORT > INF
In the Tibetic languages, things appear rather clearly since the quotative is not in the same slot as other evidential and combines with all the evidentials! (see Tournadre & LaPolla 2014), which is not the case in many other world systems (Aikhenvald 2004). As we have seen in 8.4.3.8, there is often a double marking of the access related to two sources (the present speaker and the reported speaker).

\[ S'(A:x): S'(A:y). \]

From a syntactic point of view, the reported speech occurs in the first position and is followed by the reporting verb which is placed in the final position of the sentence.


For the Tibetic languages, we thus propose below two types of hierarchies (Tournadre, Tübingen 2019). The first hierarchy is related to the source (S) and the latter is related to the access (A). These hierarchies could also well apply to other languages with rich E-E systems.

\[ S'(A:x) > S'(A:x) \]

When the actual speaker (S') has the same access to information as another reporting speaker (S'), s/he will normally choose his own access to information, rather than quote somebody else.

Concerning the accessibility hierarchy (A) we propose the following ranking.

**A: END/EGO (intent) > VIS SENS > NVIS SENS > SENS INF > NSENS INF**

We will discuss this ranking hereafter but let’s first turn back to Faller’s question (2002) about the choices between inference and reported speech. Generally the preference of inferential over quotative or the reverse depends on the types of clues available to the actual speaker and on the reliability or trustworthiness of the reporting speaker.

If a person sees snow on the ground in the morning, he is likely to use this evidence to tell his children: ‘It has snowed’ (during the night) even if a neighbor has already told him ‘it (has) snowed’ (with a sensory marker) indicating that the reporting speaker saw the snow falling.

\[ S'(A:INF) > S'(A:VIS SENS) \]
However, when the speaker’s statement depends on an inference based on a weak evidence, the speaker is likely to refer to a reporting speaker who has directly witnessed the event, if the latter is trustworthy. For example, if the speaker’s neighbor has seen a wolf attacking a small donkey and reported it, the speaker is likely to quote his neighbor than say: ‘he said that the donkey was killed by a wolf’ (he saw it), rather than base its statements on the traces left by the wolf: ‘the donkey was killed by a wolf’ (according to traces left on the body).

S1 (A: VIS SENS) > S° (A: INF)

Let us now examine the accessibility hierarchy. When describing a situation, the speaker has often various types of access to information. For example, if someone who is shivering says: “I am afraid!”, the person has access to information through two distinct ways: the inner feeling of fear (endopathic) and the vision of her hands shaking (sensory inference). If this person tells her friend that she is afraid, the addressee will have at least two types of access to information: the vision of his friend shivering (sensory inference) and the reported speech.

Let’s take another example, if a child is crying nearby, one may both hear and see the child crying. In languages which have an opposition between visual and non-visual sensory (such as Ladaks, Zanhar, Spiti, Ngari, etc.), the speaker could thus choose between the two. However, s/he is more likely to select the visual sensory marker than the non-visual sensory marker.

Thus it seems that there is a default preference for the selection of markers which seems to follow general cognitive principles of accessibility.

The two markers that have the highest accessibility in the hierarchy are the endopathic and egophoric intentional markers. If the speaker has access to this information, he is likely to use it. Both types of markers can normally be used only in the first person referring to the actual speaker or a reported speaker (in case of reported speech). In interrogative sentences, they are used with the second person (see the anticipation strategy below 8.4.7.3, see also the person correlation in 8.4.7.1). The reason of this restriction is purely semantico-cognitive. The speaker is the only one to have access to
his own intentions and to his endopathic feelings or sensations, but s/he has no access to other’s intentions and endopathic feelings or sensations.

The next highest access in the hierarchy are marked by other sensory markers (non endophatic): visual sensory then non-visual sensory, then sensory inferential (see Tournadre 2022). The most indirect and non deictic types of access are the non-sensory inferential and the quotative or hearsay markers.

A somewhat similar notion of “evidential hierarchies” has been proposed, independently, by other authors (see e.g. Faller 2002; Zeisler 2018; De Haan 1998).

### 8.4.3.12. Evidentiality and discursive types

Among the characteristics of evidentiality one specificity which has largely been overlooked is the significance of discourse types. In the Tibetic languages, at least, this factor plays a crucial role in the selection of auxiliaries. For example, when one is dealing with food recipes, proverbs, dream narratives, dialogue interactions, historical narratives, biographies, personal life stories, reporting news, etc. the set of evidential auxiliaries is likely to be different.

For example, dream narratives in the Tibetic languages are told with sensory markers (e.g. in Common Tibetan ‘DUG, SONG, BZHAG) and sensory inferential markers, including when the speaker is describing his/her own actions. Some historical or religious stories are narrated with the help of authoritative or factual auxiliaries. For example, Konchok Phanday (2017) wrote in Ladaks language the ‘Twelve deeds of the Buddha’ (MDZAD.PA.BCU.GNYIS) using all the time the factual auxiliary KAG which presents the events as facts (see also Bakula 2014). In the text (2017: 2-9), this auxiliary which is the only auxiliary used in the past occurs more than sixty times. In a sharp contrast, a tale is more likely to be told with the inferential marker TSUG which presents the event as reported. For example, in a story called ‘The foolish grandfather’ collected in Nyoma (Ladakh Janghang, Tournadre’s unpublished corpus 2018), all the events and states are reported with the auxiliary TSUG/SUG.

### 8.4.4. Epistemic modalities

Nuys (2006) proposes the following definition for this notion: “[Epistemic modality corresponds] to the marking by S of the degree of likelihood s/he attaches to a state of
affairs or the degree of certainty s/he attaches to a certain thought” or in a more detailed formulation (Nuyts 2001: 21):

“Epistemic modality is [...] an evaluation of the chances that a certain hypothetical state of affairs under consideration (or some aspect of it) will occur, is occurring, or has occurred in a possible world which serves as the universe of interpretation for the evaluation process, and which, in the default case, is the real world (or rather, the evaluator’s interpretation of it) [...].”

The epistemic modalities have not received much attention in the Tibetic languages. In comparison with the evidential modalities, there are only a few studies and the notable work by Vokurková (2008) on epistemic modalities in Common Tibetan. A possible explanation for this scarcity of academic studies is the fact that they have a much lower frequency than the evidential modalities (Mélac 2014).

Some authors (e.g. Aikhenvald 2004; de Haan 1999) insist on maintaining a clear-cut distinction between epistemic and evidential markers. However, in the Tibetic languages evidential and epistemic markers are sometimes fused together and often form a single paradigm historically made up of the same lexical source forms.

As shown by Tournadre & LaPolla (2014), “Simple evidential copulas and auxiliaries such as འིན་ YIN or རེད་ RED ‘to be’, འོད་ YOD or ཆ་ DUG ‘to be, there is’ (location, existence) may combine together with the help of connectives or nominalizers (such as sa, pa or gi) to yield compound forms (used as copulas or auxiliaries) which bear either an evidential or an epistemic meaning or both. For example, the epistemic auxiliaries འོད་ཀྱི་རེད་ YOD.KYL.RED, འོད་གྱི་རེད་ YINGY.L.RED, འོད་པ་འདྲ་ YOD.PA.DRA, འོད་གྱི་རེད་ YOD.SARED, འོད་པ་འདྲ་ YOD.SARED and the evidential auxiliaries འོད་རེད་ YOD.RED, འོད་པ་རེད་ YOD.PARED, འོད་ཀྱི་རེད་ YIN.PARED are both made of འོད་ YOD and འིན་ YIN. The simple evidential forms འིན་ YIN and འོད་ YOD convey an epistemic meaning when they occur in combination with the interrogative marker ཉ་ PA [The CT form is ཉེ་ PE however the modern languages usually have a reflex of ཉ་ PA], as in ཉ་འིན་ PA.YIN or ཉ་འོད་ PA.YOD. They convey both evidential and epistemic meanings when they occur preceded by the nominalizer PA, as in རྩ་པ་ PA.DUG.” (see also Vokurková 2008; Tournadre & Sangda Dorje 2003).
The copulative verb རྡེ་ RED (in combination with various relators) suffixed to the copulative verb འཡིན་ YIN or to the existential verb འཡོད་ YOD conveys an epistemic meaning: འཡིན་ རྡེ་ YIN.GYI.RED (Ü), འཡིན་ བཀྲ་ YIN.BK.R ED(Am, Kh), འཡོད་ རྡེ་ YOD.R ED (Am). In Common Tibetan, a dubitative marker indicating a very low probability is composed of the copulative verb and the prefixed interrogation: འཁ འཡིན་ ཡིག་ རྡེ་ A.YIN.GYI.RED (Ü). In Common Tibetan, a dubitative marker indicating a very low probability is composed of the copulative verb and the prefixed interrogation: འཁ འཡོད་ A.YOD (Ü) ’I doubt X is’, འཁ འཡོད་ (Ü) ’I doubt X has’.

The other main markers used to convey epistemic modalities in the Tibetan languages are derived from the verb ཁྲ་ DRA ‘to seem’ (Ü, Ts) and from motion verbs: འགྲི་ GRO ‘to go’ (Ü, Ts, La), བོད་ CHA.PO ‘to go+ suffix’ (Pur). In Common Tibetan, the compound made of the verb YONG is also attested for the dubitative: འཁ འཡོང་ A.YONG (Ü) ’I doubt X’ (see Tournadre & Sangda Dorje 2003: 313).

Apart from these frequent constructions, one also encounters specific constructions attested only in some areas. That is the case of the Amdo suffix མ་ཐང་ NA.THANG or the Ladaks forms གཞིག་འདུག་ THIG.DUG and གཞིག་རག་ THIG.RAG derived from THIG (Pa) ’prediction’.

In many languages, there is a grammatical distinction between sensory inference and logical inference. Let us examine some examples:

(307) མིན་ལ་ དགའ་རོགས་ འདྲ། KHONG-LA DGA’ROGS YOD-P.A.DRA 3sg.[H]-DAT lover exist-EPI+SENS  ’She probably has a boyfriend.’ [Sensory inference: the speaker often sees her with the same boy.] (U, ComTib; Vokurková 2008)

(308) མིན་ལ་ དགའ་རོགས་ འདྲ། KHONG-LA DGA’ROGS YOD-‘GRO 3sg.[H]-DAT lover exist-EPI+FACT  ’She probably has a boyfriend.’ [Logical inference: she is twenty, so the speaker guesses she has a boy friend.] (ibid.)

The same examples have direct correspondences in Amdo Tibetan (Rebgong):

(309) ནུས་ ལ་ ནི་ བཀྲ་ Roks. SA YOD-SA.YOD.GI marge roksa yo-soga 3sg.[F]-DAT lover exist-EPI+SENS  ’It seems she has a boyfriend.’ [Sensory inference: the speaker can often see her with the same person.] (Am; Tournadre & Shao, forthcoming)
‘She probably has a boyfriend.’ [Logical inference: she is twenty, so the speaker guesses she has a boyfriend.] (ibid.)

‘She may have a boyfriend.’ [Sensory inference: the speaker sometimes sees her with the same person.] (ibid.)

‘She may have a boyfriend.’ [Logical inference: she is fifteen, so the speaker guesses she may have a boyfriend]. (ibid.)

The following examples illustrate sensory epistemics:

‘She seems to be afraid of water.’ [Looking at her behaviour.] (adapted from Vokurková 2008)

‘Our black cat seems to have mated with a white cat.’ [From looking at the color of the cat.] (La; Zeisler 2012b)
In Ladaks, there is also a distinction between two types of sensory inference: visual and non-visual inferences. Here are some examples that we collected:

(316) སྤོས་ཐེ་ཡིན་ཐིག་འདུག

LTOGS-TE YIN.THIG.'DUG
hungry-PTCP be+EPI+VIS

'(Oh the baby) must be hungry!' [Looking at the baby crying.] (La)

(317) སྤོས་ཐེ་ཡིན་ཐིག་རག

LTOGS-TE YIN.THIG.RAG
hungry-PTCP be+EPI+NVIS

'(Oh the baby) must be hungry!' [Hearing the baby crying in the next room.] (La)

Ladaks has also a specific marker for logical inference:

(318) སྤུ་གུ་ཡོད་ཐིག་ཡོད།

PHRU.GU YOD.THIG.YOD
child exist+EPI+EGO

'They probably have children' ['They have been married for a long time now.'] (La)

Some other constructions are available to convey an epistemic meaning. Examples from Ladaks:

(319) སྤུ་ཡོད་མེད་ད་ཧེ།

CHU YOD-MED-DA.HE
water exist-NEG-TAG

'I think there is water' [don’t you think so?].’ (Norman, p.c., Leh, 2017)

(320) སྤུ་ཐོང་མི་ཡོང་ང་ཧེ།

YONG MI-YONG-NGA.HE
come NEG-come-TAG

'I think (they) will come [don’t you?].' (ibid.)
8.4.5. Deontic modalities

Deontic modalities in the Tibetic languages are mainly conveyed by the following verbs: ཇོ་ གོ་ བོད་ བྱ་ ཞི་ ཉོན་ ‘to allow, to be all right’, ཁོ་ དགོས་ ‘to want, must’, སྣོ། ཞུབ་ ‘can’, ཉོན་ ‘may, can’ (Am, Kh, La, Pur), སྣོ། ནོད་ ‘to be able’ (Sh), སྣོ། དོ་ ‘to be able’ (Dz). They always occupy the position of a “secondary verb” (see 8.3.10), i.e. a postverbal position between the verb and the final auxiliary. (See e.g. van Driem 1998; Zeisler 2014; Vokurková 2018; Tournadre & Sangda Dorje 2003; Ylieniemi 2019.)

8.4.6. Intentional modalities

Intentionality (or volitionality) is sometimes grammatically encoded in the egophoric auxiliary (see e.g. Delancey 1986a, 1990; Tournadre 1994, 1996a-b; Tournadre & Konchok Jiatso 2001; Oisel 2013).

The use of intentional egophoric indicate that the speaker is performing, has performed or will perform the action intentionally. So intentional egophoric auxiliaries occur in the past as well as the present and future with controllable verbs.

Altered states of consciousness, such as experiencing a dream, being drunk, being in a trance like a medium or shaman, etc. or simply doing something unwillingly will prevent the use of egophoric.

In Ü, Tsang, Amdo, Northern Kham and Ladaks, intentionality is often marked by the auxiliary འཡོད་ in the past and future tenses as well as by the auxiliary འབོད་ in the present.

Intentional egophorics are not present in all the Tibetic languages, but they are widely attested in the languages of Tibet, such as Ü-Tsang, Amdo and Northern Kham as well as outside e.g. Ladaks. Intentional egophoric can normally occur only with the 1st person (see also below 8.4.7.1).

(321) ཕི་ གན་ལ་ བསྐད་ཀྱི་ འཡོད།

NGA NANG-LA BSDAD-KYL.YIN
1SG home-LOC stay-FUT+EGO

'I will stay at home.' (ÜN, ComTib)
(322) ཆོང་ནང་འདུག་གིན།
NGA NANG-NGA 'DUG-GIN
1SG home-LOC stay-FUT+EGO
'I will stay at home.' or 'I intend to stay at home.' (La)

(323) ཆོང་ནང་ལ་བསྡད་ཀྱི་ཡོད།
NGA NANG-LA BSDAD-KYL.YOD
1SG home-LOC stay-UNCMP+EGO
'I stay at home.' (Ü, ComTib)

(324) ཆོང་ནང་འདུག་གད།
NGA NANG-NGA 'DUG-GAD
1SG home-LOC stay-UNCMP+EGO
'I stay at home.' or 'I will stay at home.' (La)

This sentence is similar to ex. 297, but 'DUG.GAD' implies that the person both intends to stay and will stay in any case, whereas 'DUG.GIN' insists on the intention of the speaker.

(325) ཆོང་ནང་དྲུག་པ་ཡིན།
NGA NANG-LA BSDAD-PAYIN
1SG home-LOC stay-CMP+EGO
'I have stayed at home.' (Ü, ComTib)

(326) ཆོང་ནང་འདུགས་པིན།
NGA NANG-NGA 'DUGS-PIN
1SG home-LOC stay-PAST+EGO
'I stayed at home.' (La)

(327) ཆོང་མི་འགྱོ།
NGA MI-'GYO
1SG NEG-go
'I won't go. [I don't intend to]' (Am)

(328) ཆོང་མི་ཆ།
NGA MI-'CHA
1SG NEG-go
'I don't go [I don't intend to]' (La)
The interpretation of these examples was proposed by Norman (pers. comm. 2017) and confirmed by our consultants. The postponed negation is also possible with

\[NGA \text{ CHA-MED.}\]

\[(329) \quad \text{NGA 'GRO-GI. MIN}\]

\[1\text{SG} \quad \text{go-FUT+NEG}\]

‘I won’t go. [I don’t intend to]’ (Ü, ComTib)

8.4.7. Specific features of the Evidential-Epistemic systems

As mentioned in 8.4.2, the Tibetic E-E systems are associated with a number of rare typological characteristics which will be discussed below (see also e.g. Oisel 2017a). In a very marginal way, the category of “animacy” has been described for a Kham dialect: the existential verb which is a reflex of \textit{YOD} is used for inanimate objects, whereas the reflex of \textit{DUG} is used for animate beings (Bartee 2007). We will not discuss further this issue but will examine below more frequent phenomena related to the Tibetic E-E systems such as person correlations (including the issue of conjunct/disjunct, anticipation strategy and mirativity.

8.4.7.1. Person correlations

One of the striking characteristics of the E-E systems is the correlation between the access to information and the first, second and third persons. This correlation has been largely overlooked in the literature. For example, Aikhenvald (2004) does not mention the strong correlation of the external sensory markers with the 2\textsuperscript{nd}/3\textsuperscript{rd} persons and the endopathic sensory with the 1\textsuperscript{st} person. Egophoric markers occur more frequently with overt 1\textsuperscript{st} person arguments, than with the 2\textsuperscript{nd} and 3\textsuperscript{rd} persons. Another frequent person correlation related to egophoric markers links the speaker in declarative statements and the addressee in questions, i.e. respectively the 1\textsuperscript{st} person in declarative statements and the 2\textsuperscript{nd} person in questions. The identity of marking of the 1\textsuperscript{st} person (in declarative sentences) and 2\textsuperscript{nd} (in interrogative sentences) has led some scholars to develop the notion of \textit{main speech act participant (MSAP)}\textsuperscript{74}.

\textsuperscript{74} See Zeisler (2018a). \textit{Main speech act participant (MSAP) refers to both the speaker (of the statement) and the addressee (of the question).}
Person correlations are probably frequent in various fields of the world languages but they have not received much attention. For example, Méloc (2014: 438, 492-493) showed that in English, the epistemic-inferential and deontic functions of the modal verb *must* are correlated to a large extent with the person: the epistemic-inferential is four times more frequent with the 3rd person (s/he *must*) than with the 1st one (*I must*). Conversely the deontic interpretation is predominant with the 1st person while it is only marginal with the 3rd person. The same is true of the adverb *apparently* which occurs predominantly with the 3rd person, and is only marginal with the 1st person and even more rare with the 2nd person. Finally, Méloc (ibid.) also shows that the verb *guess* is used in the present tense to convey an epistemic-inferential meaning in English essentially occurs with the 1st person (*I guess*) while it is rarely used with the 3rd person (*he guesses*).

Such correlations are motivated by pragmatic and cognitive reasons. It is interesting to note that in the Tibetic languages, person correlations have either been largely unnoticed (that is the case for example of the endopathic function normal occurrence with the 1st person) or on the contrary have been magnified as in the case of egophoric markers and went considered as instances of “person marking” and agreement. The predominant use of egophoric markers with 1st person arguments has led to the elaboration of the “conjunct/disjunct model.” Within this model, egophoric or personal markers have been interpreted as “conjunct” while all other evidential markers have been interpreted as “disjunct.” The misleading notions of “conjunct/disjunct” have sometimes been used as an alternative model to the E-E system in order to describe some Tibetic languages, particularly Lhasa Tibetan and Sherpa.

8.4.7.2. The controversy about the “conjunct/disjunct” pattern

Among the various works which have used this concept we find Schötte (1980), DeLancey (1992, 2001, 2003b), Kelly (2004). Incidentally this concept has also been used in some “second-hand data” typological works such as Aikhenvald (2004) and thus have received some wider attention.

Austin Hale (1980) in his study on Kathmandu Newari proposed the notion of “conjunct” versus “disjunct”. Post (2010) summarized Hale’s idea in the following way.

The “conjunct” set is normally employed in:
“simple first person declarative sentences”
“simple second person interrogative sentences”
“complex speech report constructions in which the matrix verb subject is co-referential with the complement clause subject.”

These three behaviours were referred to respectively as (a) *declarative C/D pattern*, (b) *interrogative C/D pattern*, and (c) *quotative C/D patterns* by Tournadre (2008).

Additionally, as mentioned by Post (2010), “specifically ‘conjunct’ forms in first-person simple clauses were associated with a construal of speaker intention or volition, while ‘disjunct’ forms were associated with construals of inadvertence.” This feature appears in DeLancey’s definition (2001: 372) “conjunct forms occur with first person subjects in statements and second person subjects in question which refer to an intentional act.”

However, in later works, the opposition conjunct/disjunct was defined in various ways and some authors ignored intentionality or quotative patterns as part of the C/D patterns. For example, Aikhenvald (2004: 391) mentioned the following definition:

[it refers to] “person-marking on the verb whereby first person subject in statements is expressed in the same way as second person in questions, and all other persons are marked in a different way (also used to describe cross clausal co-reference).”

Let us illustrate the three types in Common Tibetan (330-335), from DeLancey’s examples (1990: 295); see also Tournadre (2008):

**The declarative pattern:**

(330) ག་བོད་པ་ཡིན།
NGA BOD.PA YIN
1SG Tibetan be (Conj.)
‘I am Tibetan.’ (The speaker and the subject of the sentence are coreferential.)

(331) མ་བོད་པ་རེད།
KHO BOD.PA RED
3SG Tibetan be (Disj.)
‘He is Tibetan.’ (The speaker and the subject of the sentence are not coreferential.)
The interrogative pattern:

(332) ཕྱེད་རང་བོད་པ་ཡིན་པས།
KHYED.RANG BOD.PA YIN-PAS
2SG(T) Tibetan be(Conj.)-Q
‘Are you Tibetan?’ (The addressee and the subject of the sentence are coreferential.)

(333) ལེ་ཁོ་མི་རེད་པས།
NGA RGYA.MI RED-PAS
1SG Chinese be(Disj. from the hearer)-Q
‘Am I Chinese?’ (The addressee and the subject of the sentence are not coreferential.)

The quotative pattern:

(334) བོད་པ་ཡིན་ཟེར་གྱིས།
KHO-S KHO BOD.PA YIN ZER-GYIS
3SG-ERG he Tibetan be(Conj.) say-IMPF/DISJUNCT
‘He says that he is Tibetan.’ (The subject of the embedded clause is coreferential with the subject of the matrix clause.)

(335) བོད་པ་རེད་ཟེར་གྱིས།
KHO-S KHO BOD.PA RED ZER-GYIS
3SG-ERG he Tibetan be(Disj.) say-IMPF/DISJUNCT
‘He says that he is Tibetan.’ (The subject of the embedded clause is not coreferential with the subject of the matrix clause.)

The conjunct/disjunct essentially corresponds to a syntactic approach related to coreference/non-coreference pattern of the “1st person subject.” This approach fundamentally differs from the E-E system approach for the following main reasons:

(a) It is syntactic in nature and not motivated by semantico-cognitive parameters.
(b) It is binary in nature, while E-E systems attested in the Tibetic languages comprise a fairly large paradigm of forms and functions (see 8.4.2).
(c) The use of conjunct/disjunct categories is largely automatic and compulsory unlike the concept of egophoric, sensory, inferential, etc. which may depend on the speaker’s perspective (see Tournadre & LaPolla 2014).
(d) It is based on the notion of person coreference patterns, whereas in our approach “the person agreement” is a secondary effect of semantico-cognitive concepts related to the evidential source and access to information.

(e) The “conjunct” category does not convey any specific semantic meaning unlike the category of “egophoric” (see 8.4.2).

(f) Some egophoric markers, such as the receptive marker BYUNG or the benefactive DGOS are not mentioned as conjunct. The marker BYUNG is non-intentional while DGOS is always related to the speaker’s intention to perform an act for the benefit of the hearer.

(g) The conjunct/disjunct pattern or system is a complex category that usually manifests itself in three distinct patterns: “the declarative pattern”, “the interrogative pattern” and “the quotative pattern.”

In some papers, DeLancey (2001, 2012) has implicitly proposed to use the terms “conjunct/disjunct patterns/systems” and “egophoric systems” as synonyms, despite the fact that the two notions are very distinct in their approach. The term “egophoric” used by Tournadre since (1991) did not refer to a system but to a specific category of the Evidential/Epistemic system, used along with many other categories (see 8.4.2).

Some authors such as Sun (1993), Tournadre (2008), Barthe (2007) Gawne (2017), have explicitly rejected the notion of conjunct/disjunct as relevant categories to describe Tibetic languages and most of the linguists who have worked extensively on Tibetic languages have avoided both the terms and the notions in their description (see e.g. Gawne & Hill 2017). Sun (1993) wrote the following comment:

“[the terms ‘self person’ and ‘other person’] are related to, but not identical with, the structurally-based labels ‘conjunct’ vs. ‘disjunct’. [...] The terms ‘conjunct’ and ‘disjunct’ are incidentally, utterly unrevealing because although the nomenclature appears to be based on structural co-reference of the matrix and complement clause subjects (i.e. ‘conjunct’ if they are co-referent, ‘disjunct’ if otherwise), co-reference is actually relevant only when the subject of the complement clause is portrayed as a volitional actor. [...] Since the distinction involves more than mere structural coreference, more self-evident labels should be sought, probably along the lines of such semantically-based terms as shenzhi ‘thoroughly integrated knowledge’ [...] quezhi ‘positive knowledge’ or Tournadre’s term egophoric.”
Zeisler (2018a: 86) also shares this opinion:

“This flexibility [in the choice of E-E markers], which is not only exploited for mirative meanings, but for various pragmatic effects, speaks clearly against any description in terms of conjunct/disjunct or even some kind of congruence. Such descriptions are completely misleading with respect to the Tibetan languages.”

As we will see the “interrogative conjunct/disjunct pattern” and the “quotative conjunct/disjunct pattern” can receive alternative descriptions. More accurate explanations for these phenomena are driven from the notion of “anticipation strategy” (see below) and “hybrid indirect speech.” The conjunct/disjunct model has largely been abandoned by the specialists of the Tibetic languages. (See Gawne & Hill 2017; Zeisler 2018a; DeLancey 2018.)

8.4.7.3. Anticipation strategy

The anticipation strategy is an important characteristic of the Tibetic E-E systems. From a typological point of view, one should notice that it is a phenomenon rarely attested in the World languages. This strategy implies a correlation between the speaker in positive statement, i.e. the 1st person and the addressee in questions, i.e. the 2nd person.

This phenomenon may be best explained by the pragmatic notions of “empathy” (Kuno 1987) and “perspective” (Tournadre & LaPolla 2014). In direct interrogative sentences, the speaker anticipates the addressee’s source and access to information by using the auxiliary expected in the answer. Thus, if we take again the example (332) repeated below for convenience sake, YIN is used by anticipating the answer using the egophoric auxiliary:

(336) མཁྱེད་རང་བོད་པ་ཡིན་པས།
KHYED.RANG BOD.PA YIN-PAS
2SG(H) Tibetan be(EGO)-Q
Q: ‘Are you Tibetan?’

ལགས་ཡིན་/ མིན་
LAGS YIN/MIN
H be(EGO)/ be(EGO+NEG)
A: Yes, I am / no, I am not. (Ü, ComTib)
The same is true for the following sentence where the auxiliary *BYUNG* is used by anticipation:

(337) སྣ་ གྲོད་ཁོག་ལྟོགས་ཀྱི་འདུག་གས།

*NGA-S ZHUS-BYUNG-NGAS*
1SG-ERG tell(h)-CMP+EGOREC-Q

'Did I tell you?' (Ü, ComTib)

The answer being:

(338) མི་ལྟོགས་གི་

*KHYED.RANG-GIS GSUNG-(MA).BYUNG*
2SG(h)-ERG told(h)-CMP (NEG)+EGOREC-Q

'You told me / did not tell me.' (Ü, ComTib)

Contrary to what has been suggested by the literature on conjunct/disjunct, this phenomenon of anticipation is not restricted to the so called "conjunct" markers. In other words, the anticipation strategy is a general phenomenon of Tibetic languages, which implies a perspective shift and applies to various evidential categories. For example, the "endopathic" function of the sensory marker, which is normally only used with the first person in declarative sentences, but in interrogative sentences, the endopathic is used with the 2nd person as a result of the anticipation strategy. The speaker anticipates on the addressee's access to information:

(339) གྲོད་ཁོག་ལྟོགས་ཀྱི་འདུག་གས་

*KHYED.RANG GROD.KHOG LTOGS-KYL.DUG-GAS*
2SG where be hungry-ENDO-Q

'Are you hungry?' (Ü, ComTib)

(340) རྗེས་བྱུང་

*KHYOD VA-LTOGS-GI*
2SG Q-be hungry-ENDO

'Are you hungry?' (Amdo)

In the above sentences, the sensory endopathic markers *KYL.DUG* and *GI*, which are normally restricted to the first person in declarative sentences, occur here in the interrogative sentence by anticipation of the answer, respectively:
The anticipation of the answer is implemented for the various types of access to information. For example, in order to ask somebody “where were you born?”, one has to anticipate the access to information. The usual way to ask this question in Common Tibetan is to use the factual auxiliary RED:

(343) 

\[
\text{KH} \text{YE} \text{D.RAN} \text{G} \text{A.PAR} \text{SKYE} \text{S-P.A.RED} \\
2\text{sg(h)} \text{where to be born-CMP+FACT} \\
\text{‘Where were you born?’ (U, ComTib).} \\
\text{The honorific form } \text{KH} \text{RUNGS-P.A.RED} \text{‘to be born’ (H} \\
\text{may also be used instead of } \text{KH} \text{YED.RANG SKYES-P.A.RED}. \\
\]

In Ladaks, the authoritative is used:

(344) 

\[
\text{NYE.RANG GA.RU SKYES} \\
2\text{sg(h)} \text{where to be born-CMP+AUTH} \\
\text{‘Where were you born?’} \\
\text{The choice of the factual in the question is motivated by the expected source and access to this kind of information. A person knows about her/his birth place usually through reported speech. Other types of access are either marginal or not available: since it is not possible to witness one’s own birth, one may not use the sensory } \text{SONG} \text{ in the answer. Inferring the place of one’s own birth, based on various clues is possible but not frequent, thus the use of the sensory inferential } \text{BZHAG} \text{ would be quite marked.} \\
\]
8.4.7.4. The issue of the "mirativity" and the speaker’s attitude

The notion of “mirativity” has been used to describe some markers found in the Tibetan languages or other related languages by a few authors, such as DeLancey (1997, 2001, 2012), Hein (2007), Hyslop (2011), Huber (2002), Aikhenvald (2012), Yliniemi (2017) etc., proposes that the following values are included under the “mirativity” label:

(i) sudden discovery, sudden revelation or realization (a) by the speaker, (b) by the audience (or addressee), or (c) by the main character;
(ii) surprise (a) of the speaker, (b) of the audience (or addressee), or (c) of the main character;
(iii) unprepared mind (a) of the speaker, (b) of the audience (or addressee), or (c) of the main character;
(iv) counter-expectation (a) to the speaker, (b) to the addressee, or (c) to the main character;
(v) information new (a) to the speaker, (b) to the addressee, or (c) to the main character.

Huber (2002) used the term “mirative function” to describe the copula /nukpa/ found in Kyirong Tibetan.

(347) ꜊ འཐེ གི་མོ་ ངདོམ་ ཡུག་པ།
```plaintext
?A.MO L.DEM.MIG 'DI-LA NUG.PA
/'amo ’dimi ’di:la ’'nukpa/ INTJ key demo-DAT exist-TAG

'Oh the key is here!' (and I have been looking for it everywhere)```
In fact, Huber rightly considers that the mirative is just one of the functions of the direct sensory marker /nug/ (derived from ‘DUG’). Here it is followed by a tag /pa/. Although not using explicitly the term “mirative”, van Driem (1998) has described some markers of Dzongkha as conveying “new” versus “old” information. According to DeLancey (2001: 274), “the paradigm [of the Lhasa Tibetan markers such as YOD, ‘DUG, SONG, PA, YIN] is built on a fundamentally mirative distinction, with evidentiality as a secondary and somewhat independent addition.”

Lazard (1999) proposed to subsume the notion of “mirative” under the notion of mediativity. The very existence of a “mirative” category as a valid grammatical category has recently raised a controversy (Hill 2012c; DeLancey 2012). Zeisler (2018a: 69) advocates for the broader notion of “admirative”:

“[...] it marks the speaker’s mental distance or non-commitment towards the proposition, be it because he or she has only ‘indirect’ knowledge (inference or hearsay) or because the content of the proposition is somehow awkward and (socially) unexpected.”

Except for a few authors, the category of “mirative” has not been popular to describe copulative and auxiliary markers in the Tibetic languages. We certainly do not agree with DeLancey’s description of the sensory marker ‘DUG’ (see section 8.4.2) in Lhasa Tibetan as a “mirative” marking (1997, 2001, 2012) indicating “new” or “surprising information” (see 8.4.3.1). Our position is shared by other scholars such as Hill (2012c, 2013a-b), Mélac (2014) or Zeisler (2018a). But this does mean that mirative is not a useful category. The main issue is whether “mirativity” should be considered as an independent phenomenon distinct from “evidentiality” as claimed by DeLancey. B. Zeisler (2018a: 69-70), who has described many dialects of Ladakh, gives the following answer: “Despite DeLancey’s claim to the contrary, evidentiality and mirativity seem to be closely linked in Lhasa Tibetan, and not to be separate categories and ‘mirative’ would only be one of the values of evidentiality.” Tournadre & LaPolla (2014) recognize that “‘DUG may have overtones of ‘mirative’ in some contexts, but [...], the core function of ‘DUG is to indicate sensory and endopathic access to information.”

Zeisler (2018a) also reports the “admirative connotation” for the visual sensory markers ‘DUG and the non-visual sensory marker GRAG in Ladaks.
We will provide some examples in this language which do imply a kind of mirative (or admirative) connotation:

(348) ◊ ར་ བཏངས་ཏོག

*KHA*  *BTANGS-TOG*

snow  fall-PFT+SENS+MIR

‘Oh, it has snowed!’ [looking at the white mountains around] (La)

This example often conveys an element of surprise. S/he could say this sentence when opening the window in the morning and discovering the white landscape.

However, if he reports this information to his friend who has seen it yet, s/he is likely to report it with a resultative perfect:

(349) ◊ ར་ བཏངས་ཏེ་འདུག

*KHA*  *BTANGS-TE.'DUG*

snow  fall-PFT+SENS

‘It has snowed.’ (La)

This sentence does not entail any inference nor surprise and simply insists on the resultative state.

(350) ◊ ར་ རེ་ཆག་ཏོག

*KO.RE*  *CHAG-TOG*

bowl  be broken-PFT+SENS

‘Oh, the bowl is broken!’ [Seeing the pieces scattered on the floor.] (La)

It may convey a mirative flavor. Again, the speaker could report this information using the resultative perfect which does not have this surprise connotation:

(351) ◊ ར་ རེ་ཆག་ཏེ་འདུག

*KO.RE*  *CHAGS-TE.'DUG*

bowl  be broken-PFT+SENS

‘The bowl is broken.’ (La)

Finally, a specific form *–PALA* found in some languages such as Ü, Tsang or Ladaks and CT may also convey some mirative connotation. In Common Tibetan it appears only with predicative adjectives:

(352) ལྷའི་དབྱིངས།

*SKYID-PALA*

pleasant-SFE

‘Wow, it is so nice!’
In Ladaks, it is used with verbs:

(353) ◊ ཁ་བཏང་ང་ལ།

*KHA BTANG-NGA.LA*

snow fall-SFE

‘Wow, it is snowing!’

Both these sentences express an exclamation and are often related to a surprise.

8.4.8. Directionality

Unlike other languages of the ST macrofamily (see particularly the Qiangic, rGyalrongic or Sinitic languages) directionality is not heavily grammaticalized in most Tibetic languages. One mainly encountered deictic adverbials འཕར་*PHAR* ‘there (away)’, ‘thither’ and དུས་*TSHUR* ‘here, hither’, ཡར་*YAR* ‘upward’ and ཁམ་*MAR* ‘downward’ attested in CT that are prefixed to the verb (Bartee 2007). Deictic motion verbs ཕོང་*YONG* ‘to come’ and འགྲོ་*GRO* ‘to go’ are also used to indicate the direction towards or away, from the deictic center or from the reference point (Tournadre & Konchok Jiats 2001).

Ex. in Common Tibetan:

(354) དུས་ཟོག

*TSHUR SHOG*

hither come

‘Come here!’

Usually the directional markers function as prefixes (or clitics) but in some cases, they appear as independent words in prohibitive and interrogative sentence. In the example below the form *YAR* or *YARA* (< CT *YA*’up’ + *R* ’dative’) correspond to an adverb.

(355) ཡར-མ་’གྲོ

*YAR-MA’GRO*

upwards-NEG-go

‘Don’t go up!’ (Kh)

Here are examples of directional markers with motion verbs together with secondary verbs in Common Tibetan:
The timber is falling on us.'

Water is going out.'

In Dzayül (Dagong Village, Gula Township) and some southern Kham dialects, deictic adverbials have been grammaticalized into directional prefixes. These prefixes essentially occur in the completed past of declarative sentences as well as in the imperative. They are not compatible with verb-prefixed negation. The use of directional prefixes depends on the person and the illocutionary force (declarative, interrogative or imperative).

For example, the prefix གར་ PHAR is used with imperative sentences:

Eat!' (Dza)

While the prefix ཀྲུ་ TSHUR is used with first person and by anticipation with 2nd person:

I ate.' (Dza)

But with the negative prefix, the directional marker does not occur in Dzayül:

'I did not eat.' (Dza)
Some varieties of Southern Kham such as Chagthreng and Derong-nJol use a directional marker *BA* specialised for an imperative, e.g.:

(361) ◊ བ་བཙོང་།

*BTA-BTSONG*

DIR-sell

‘Sell it (to me)!’

Aside these marginal examples, directionality does not play a major role in the Tibetic languages, unlike many other ST languages.

### 8.4.9. Interrogative and tag question

In some languages such as Nagchu Hor, questions are essentially indicated by a rising intonation. However, in most languages, one encounters question markers. As mentioned in 8.3, there are two types of interrogative markers across the Tibetic languages: sentence final interrogative markers and verb prefixed interrogative markers. The latter is always derived from the morpheme *ཨ་/ʔa* (*CT ཨེ་/ʔE*). Languages which do have such a marker are located in Eastern Tibet, in Amdo and Kham. All other regions in Central, South and Western areas make use of sentence final interrogative markers.

Several sentence final interrogative suffixes are attested in the various languages. Various forms: *པ་/-pa* (Kyi), *པས་/-pä* (*Sp, Dz*), *ངས་/-ngä* (*Ü*), *གས་/-gä* or */-kä* (*Ü*), *ག་/-ka* or */-ga* (*Sp, Dz, etc.*). More reduced forms such *a/* or *C+a/* (with *C* corresponding to a homorganic sound with the final consonant of the preceding word (La, Cho)) probably derive from the CT marker *པ་/PA*. The form */-nã:/* found in Kyirong (see Huber, 2002) is derived from the CT interrogative marker *ནམ་/NAM*.

### 8.4.10. Imperative and jussive

Many Tibetic languages have inherited special verbal inflections stems to convey the imperative and jussive meanings (see 8.3.8, see also Zeisler 2004). For example, in Common Tibetan, *LTOS* /tös/; *ZO* /so/; *NYOL* /nyö:/ are respectively the imperative forms of *བལྟ་/BLTA* /ta/ ‘to look at’, *ཟ་/ZA* /sä/ ‘to eat’, *ཉལ་/NYAL* /nyä:/ ‘to lie down, to sleep’. In some cases, one also encounters innovative inflections for the imperative. For example, in Amdo, some verbs have an aspirated initial consonant for the imperative: ◊ ཨེ་*PHRIS* ‘write!’ (versus CT ཨེ་*BRIS*), ◊ ཨེ་*PHROGS* ‘cut!’
shave!' (versus CT བྲེགས་ 'BREGS), འཕྲོས། PHERONGS 'lead!' (versus CT འཕྲོས། DRONGS), འཕྲོས། PHROS 'run away! flee!' (versus CT འཕྲོས། BROS) (Hua & KL, 'BUM RGYAL 1993).

With the exception of Balti and Purik, which use the negation with the imperative forms, the prohibitive is usually marked in the Tibetic languages by the negation simply followed by the verb in the present form (and normally not the imperative as expected) as in the following example: ཀུན་ MA-LAB 'don’t talk!' (but not ཀུན་ MA-LOB), ཀུན་ MA-LTA 'don’t look!' (but not ཀུན་ MA-LTOS). There are exceptions however. In Kham, there are two expressions for the prohibitive of ‘go’: ཀུན་ MA-ＧＲＯ ‘don’t leave!’ (with the present form) and ཀུན་ MA-ＳＯＮＧ ‘don’t leave!’ (with the past form). The former presupposes that the person will come back while the latter may imply that the person won’t come back.

However, in the Tibetic languages, the most frequent strategy to indicate the imperative and jussive as well as other related meanings such as the exhortative and the optative is the use of postverbal clitics or auxiliaries. The postverbal clitics བླ་ DANG (Ù, Ts, Cho, etc.), བླ་ ANG (La) and བླ་ SHIG (Ù, Ts, La, Dz, etc.) are attested in a number of languages.

In Common Tibetan, aside from བླ་ DANG and བླ་ SHIG, one finds the following markers: བླ་ DO, བླ་ PA, and the auxiliaries བླ་ SHOG (derived from the imperative form of ‘to come’), བླ་ ROGS GNANG (derived from the verb ROGS ‘to help’ followed by the honorific auxiliary GNANG ‘to give/ grant/ make’), བླ་ PAR BYED (derived from the verb BYED ‘to do’), etc.


Various jussive clitics and auxiliaries are also found in other languages. For example in Dzongkha we have བླ་ SHIG /-sh/, བླ་ GSUNGS-LAGS /-sh-la/ (H), བླ་ GNANG (derived from ‘to grant’, ‘to make’), བླ་ SMAS (the latter indicates urgency. It is

75. We thank B. Zeisler (pers. comm. 2020) for mentioning this specificity.
76. SHIG is often realized as /-sh/. That is the case in Ù, Tsang, Dzongkha.
probably derived from CT verb སྨྲས་ ‘to tell’), སྨ་རེ་ ‘SMA-RE which conveys authority on the part of the speaker (see van Driem 1998: 420). In Choča-ngača, also spoken in Bhutan, the postverbal clitic མོང་ ‘DANG is used as in Lhasa but it is pronounced /dang/ (see Tournadre & Karma Rigzin 2015).

Ladaks has a form LO or LE: འོང་ཀོ་ ‘YONG-LO ’remember to come, all right’, ’make sure to come’.

Common Tibetan has a familiar form with a similar: འོང་ ‘ONG-GO’ ‘please do it’.

Along with the marker དིིོག༔ ‘SHIG, Ladaks has the auxiliary གྲེ་ ‘MDZAD (H). It is derived from the verb ‘to do’ (H): བེ་རེ་ ‘KHYER-RA-MDZAD ’Please take (it)!’ (see Koshal 1979, 1982; Norman 2019).

In Amdo one encounters the postverbal clitics ཤ་, ར་, འ་, and the polite form གོ་, as well as the auxiliary གོ་ ‘THONGS, imperative form of the verb གཏོང་ ‘GTONG ‘to send’ (see Robin & Simon forthcoming). གོ་ ‘THUNGS-RA /’t’ong-ra/ ‘(please) drink’, ནུ་སྲོག་པ་ ‘MUR-GE-A SHOD-LA THONGS /mNge x’o-la-t’ong/ ’tell her!’.

8.4.11. Negation

There are only two negation clitic morphemes found across the Tibetic languages. They are reflexes of མྱི་ and མ་ found in Old Tibetan. The archaic form མྱི་ is also attested in Amdo, in some Eastern languages and in Southern Kham. The negation may be an easy test to differentiate from the Tibetic languages some closely related TB languages located in the Tibetosphere. For example, if a language has a negation in /a/, it is clearly not a Tibetic language. This is the case of Bake (བྲག་སྐད་) in Central Tibet or Prinmi (Pumi) in Kham or Tamang in Nepal.

In CT, the use of མྱི་ and མ་ with lexical verbs depends on the TAME. In CT, མྱི་ is used with the present, future and the imperfective past whereas མ་ normally occurs with the completed past or completed future (notably in subordinate clauses) and the imperative. However, in the modern languages, this rule does not apply to the auxiliaries and the use of the negation marker essentially depends on the auxiliary. Each auxiliary is associated either with མྱི་ or མ་. For example, one finds མ་པོ་ ‘MA-RED (Ü, Kh, Am) and not མ་ ‘MI-RED, མ་ ‘MA-DAG (Tö) or
the variant 0 མའདད་ ‘MA-DAD and not མི་འདད་ ‘MI-DAD, གྲུ་ ‘MA-BYUNG (Û, Kh) and not ‘གྲུ་ ‘MI-BYUNG, མི་འདད་ ‘MI-DUG (La, Ú, Ts, Kh, Dz, Am, etc.) and not ‘གྲུ་ ‘MA-DUG’, མི་འདད་ ‘MI-SNANG (Ph, Hor, Kh) and not མི་འདད་ ‘MI-DAD, མི་འདག་ ‘MI-DUG (La, Ü, Ts, Kh, Dz, Am, etc.) and not ‘མི་འདག་ ‘MI-DAG

The two pandialectal verbs, the essential and existential copulative verbs ཡིན་ ‘YIN and ཡོད་ ‘YOD have a special negation with a reduced form of མི་ ‘MI and མ་ ‘MA as /m/ fused with the copulas: the negation of ཡིན་ ‘YIN is མིན་ ‘MIN in CT and in the majority of languages. The negation མ་ ‘MA-YIN is attested in CT and in Dzayül. The negation of the existential copula ཡོད་ ‘YOD is མི་ ‘MED in the majority of languages and མི་ ‘MI-SNANG, མི་ ‘MI-GDA’ (Hor, Kh) and not མ་ ‘MA-GDA’, etc.

In some languages, such as Sherpa, the negative clitic མི་ ‘MI used for the present and future undergoes a vowel assimilation and takes the vowel color and the tone of the lexical verb (see Tournadre et al. 2009): མ་ ‘MA, མི་ ‘MI, མི་ ‘MU, མི་ ‘ME, མི་ ‘MO:

Ex. མི་ མོ་ ’GRO/mo-’do/’(s/he) does not/ won’t go’, མི་ མོ་ ’DUNG/’mut’ung/’(s/he) does not/won’t drink’, མི་ མི་ ’me-lep/’ME-SLEB’(s/he) won’t arrive.’ The negation མ་ ‘MA used with the past and imperative remains unchanged (ibid.).

This type of vowel assimilation is also attested in Dolpo with the past negation མ་ ‘MA.

Moreover, in some dialects the form of the negation may be altered by a final nasal. This is the case in Lhoke, as noted by Yliniemi (2019). The negation with a low tone verb has a final nasal: Ex. 0 མི་ ’MAM-BYIN ‘mam-bin/’did not give’, 0 མི་ ’MAN-ZA/’man-za/’did not eat’, 0 མི་ ’MAN-GO/’mang-go/’did not understand’.

Negation can be expressed in the form of a rhetorical question. In this regard, some varieties of Southern Kham use interrogative words such as ཁར་ ‘GAR’ where’ and ཆི་ ‘CHI’ what’ as a negation marker, e.g. in Choswateng:

77. In southern Kham, some dialects allow a choice of the negation clitic MA or MI with the auxiliary ‘DUG and SNANG.
PART 2 – CHAP 8. Grammatical outline of the Tibetic languages 465

(362) ◊ བོད་མྱི་གོ

HA MYI-GO
know NEG-STEM
'I do not understand [what you want to say].' (SKh)

(363) བོད་མ་གོ

HA MA-GO
know NEG-STEM
'I did not understand [what you've said].' (SKh)

(364) བོད་གར་གོ

HA GAR-GO
know NEG-STEM
'I do not understand (it at all).' (Lit. 'Where do (I) know/understand?') (SKh)

(365) ◊ བོད་ཆི་གོ

HA CHI-GO
know NEG-STEM
'You do not/ He does not understand (it at all).' (Lit. 'What do (you/he) understand?') (SKh)

In the Choswateng dialect, གར་ GAR is mainly used as an egophoric or sensory negation, and ཆི་ CHI, as a non-egophoric negation (Suzuki & Lozong Lhamo 2021). The crucial point is the position of these negation markers, i.e. the position just preceding a verb stem. In other words, they occupy the same position as the negation markers MI/MYI and MA.

In the Čangdrong dialect, གར་ GAR as a negation is mostly used with the existential verb རོང་ YOD. It expresses a strong negation for a given sentence.

(366) གར་ རོང་

GAR YOD
where EXV
'It does not exist!' (Kh: Čangdrong, SR)

In varieties as shown above, the intention of “rhetorical question” has already been lost, and the interrogative words seem to be grammaticalized as a negation prefix which behaves in the same way as MI/MYI and MA.
See also Amdo:

(367)  ཆི་ཡིས་ཡིན།

CHI-YIS YIN
what-ERG CPV

'It does not exist!' (Am)

One ought to mention two very marginal negation enclitic morphemes: ལེ་ re attested in OT and ལ་ na in Amdo. (Shao, forthcoming)

8.5. Interjection

As in other languages of the world, the Tibetic languages have a number of interjections. Some of them are found in most areas. Among the most frequent interjections we find:

- ཧ་ སྤ་CHU or ཧ་ སྤ་CHU:CHU 'An exclamation of cold.' [FFW] (Û, Tsang, Dz, La, etc.)
- ཧ་ སྤ་TSHA or ཧ་ སྤ་TSHA: TSHA 'Ouch! Exclamation of heat or acute pain.' [FFW] (Û, Tsang, La, Dz, etc.)
- ཧ་ སྤ་ARARA 'Ouch! Expression of pain or indignation.' (La, Û) (see Norman 2019)
- ཧ་ སྤ་PALALA 'Ouch! Exclamation of pain.' (Dz)
- ཧ་ སྤ་PAL 'Expression of surprise.' (Û)
- ཧ་ སྤ་ARASHI ARASHI: TSHI 'Oh! An expression of shock, dismay.' (Û) 'pain or illness' (La)
- ཧ་ སྤ་AMLAMA (La), ཡི་ སྤ་AMSAMA 'Wow! An expression of surprising or amazement.' (Û, Ts)
- ཧ་ སྤ་SPAD (Dz), ཡི་ སྤ་SPALALE 'Wow! An expression of surprising or amazement.' (Û, Ts)
- ཡ་ སྤ་YA 'An expression of disappointment.' (Û)
- ཡ་ སྤ་AKHA 'An expression of regret, pity.' (Û, Dz)
- ཡ་ སྤ་HA 'An expression of pleasure or satisfaction.' (La) (see Norman 2019)
8.6. Summary of the main grammatical differences between the Tibetic languages

We will summarize the main grammatical differences concerning the nominal and verbal domains:

a) Nominal morphosyntax

- **Noun morphology**: There are frequent differences in the affixes (essentially the suffixes). See 8.1.1.
- **Adjective morphology**: There are frequent differences in the affixes (essentially the suffixes). See 8.1.6.
- **Demonstratives**: The demonstratives are usually placed after the head noun in the eastern and central languages (Ü, Ts, Kh, Am), whereas they are usually placed before the head noun in western and southern languages (Ba, Pur, La, Sh, Dz, etc.).
- **Definite marker or article**: Most languages lack a definite article. However, some languages such as Sherpa, Kyirong, Ladaks, Purik or Balti, do have a definite marker. It is either derived from the demonstrative 'DI or DE or from the old definite marker: 'O/PO.

b) Verbal morphosyntax

- **Nominalizers**: There are differences in the forms and functions of the nominalizers. See 8.3.13.
- **Copulative and auxiliary verbs**: There are many differences in the forms and functions of copulative and auxiliary verbs.
- **Verb stems**: Some languages such as Amdo, and to lesser extent Ü, Tsang and Northern Kham have preserved the CT morphology whereas most languages (Sherpa, Dzongkha, Ladaks, Purik, Balti, Southern Kham, etc.) have largely lost the stem variations (see 8.3.6).
- **Tense-aspect**: There are a number of differences in the tense-aspect systems. For
example, some languages have various progressive aspects and a few western languages have a specific suffix /pin/ for the past tenses (La, Ba, Pur), whereas many languages do not make such differences.

- **Evidential-Epistemic**: The languages exhibit some significant differences in their evidential-epistemic systems, particularly in the subcategories of evidentials and epistemics. For example, Tö-Ngari, Spiti and Ladaks make a distinction between visual and non-visual sensory perceptions, whereas most languages do not make this distinction. Many languages have various types of egophoric but some languages such as Balti and Purik entirely lack egophoric marking.
9. Inner classification of the Tibetic languages

9.1. Previous works on the classification

Until very recently, works dealing with the linguistic classification of languages derived from Old Tibetan have not used the term “Tibetic languages” but the term “Tibetan dialects,” thus giving the wrong impression that Tibetan was a single language with an incredible amount of dialectal variety. There was however some ambiguity since the so called “Tibetan dialects” were divided into groups which may be interpreted as various closely related “languages.”


With a few exceptions such as Bielmeier (2003), Bielmeier et al. (2018), Nishi (1986), van Driem (2001), Zeisler (2004, 2005) and Tournadre (2005, 2008, 2014a), SUM.BHA DON.GRUB TSHE.RING (2011), most authors limit their analysis to Tibet and do not propose a classification of all the Tibetic areas. Usually, because of the lack of data available on languages and dialects outside of China, the Chinese scholars, whether Tibetans or Han Chinese concentrate their articles on the Tibetan dialects spoken within the People’s Republic of China, except for Jiang (2002), which also mentions Dzongkha, Zanhar and Ladakhs and the recent publications of SUM.BHA DON.GRUB TSHE.RING (2011, ed. 2013). A few studies of western authors propose mainly classifications of western languages or dialects (see e.g. Zeisler 2011).

Some authors are mainly concerned with the classification of the macro-family (Sino-Tibeto-Burman). That is for example the case of Shafer (1955; 1966) who proposed a general classification of ST. In this classification, he proposed the establishment of a Bodish
branch, which roughly corresponds to the Tibetic family (except for East Bodish) was divided in four units: West Bodish, Central Bodish, South Bodish and East Bodish. Shafer (1966) proposed that Ladakhi [Ladaks], Purik and Balti are not derived from Old Tibetan and form a separate branch (West Bodish) of the Tibetic languages ("Tibetan dialects"), which are derived from Old Tibetan. This hypothesis supported by Bidmeier (2004) was based primarily on the fact that Ladaks, Balti and Purik "have not preserved traces of the present and future stems." Zemp (2014) criticized this approach and showed it was not entirely accurate. As noted by Zhang (1996: 115), "In general, Chinese scholars divide the Tibetan language in China into three groups: Dbus-gtsang, Khams and Amdo." Zhang summarises the various propositions of the Chinese authors in a synoptic chart, which we reproduce here. Within each group, the author provides the sub-groups.

These classifications have some flaws, which we will not discuss in detail. Let us just mention the main ones:

a) Some languages/dialects listed in Chart IX.1 are clearly not Tibetic languages as defined in the present book. That is the case of Brag-gsum which is a Bodish language.

b) The groupings are not based on specific phonological or lexical common innovations but rather very general phonological features, such as the loss of final consonants, the absence of tones, the preservation of initial consonantal clusters, etc. They usually do not take into account grammatical features.

c) The classification below takes in account neither the degree of mutual intelligibility nor the geographic parameters.

d) The so called 'brog-skad' or Pastoralists' speech is both a sociolinguistic and geolinguistic category. These dialects are not homogenous, and it is difficult to speak of one pastoralist dialect for an entire region such as Amdo or Kham.
## Chart IX.1. – Zhang's synoptic chart (1996)

<table>
<thead>
<tr>
<th>dBus-skad sg</th>
<th>gTsang-skad sg</th>
<th>Southern sg</th>
<th>Central sg</th>
<th>Northern sg</th>
<th>'Brog-skad sg</th>
<th>Rong-skad sg</th>
<th>Rong-ma-'brog-skad sg</th>
<th>'Brog-skad sg</th>
<th>'Taus-skad sg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qu Aitang</td>
<td>dBus-skad sg</td>
<td>gTsang-skad sg</td>
<td>sTod-skad sg</td>
<td>Brag guum skad sg</td>
<td>Shar-pa skad sg</td>
<td>Eastern sg</td>
<td>Southern sg</td>
<td>Western sg</td>
<td>Northern sg</td>
</tr>
<tr>
<td>Hu Tan</td>
<td>dBus-skad sg</td>
<td>gTsang-skad sg</td>
<td>Southern sg</td>
<td>Northern sg</td>
<td>'Brog-skad sg</td>
<td>Rong-skad sg</td>
<td>'Brog-skad sg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang Jichuan</td>
<td>dBus-skad sg</td>
<td>gTsang-skad sg</td>
<td>sTod-skad sg</td>
<td>Central sg</td>
<td>(Eastern branch, Western branch, Northern branch, Central branch, 'Brog-skad)</td>
<td>bDe-chu skad sg</td>
<td>Co-ne and 'Brug-chu skad sg</td>
<td>Rong-skad sg</td>
<td>(Southern branch, Northern branch)</td>
</tr>
</tbody>
</table>

In the above chart, there is no mention of the dialects spoken in India, Bhutan, Nepal, Pakistan and Myanmar (Burma). It mentions three groups of dialects and a number of sub-groups ('sg').


*Zangyu fangyan gailun* has been published in 2002 by SKAL.BZANG 'GYUR.MED (Gesang Jumian), with the help of his daughter SKAL.BZANG DBYANGS.CAN (Gesang Yangjing) but is based on a manuscript written in 1964 and used since as a pedagogical tool by students in the Central University of Nationalities in Beijing (now called Minzu University of China). This book deals with the three main dialectal groups spoken in China: Ü-Tsang, Kham and Amdo. It is the result of SKAL.BZANG 'GYUR.MED's personal research but it was also conceived as a general project of research on Tibetan dialects. This large survey involved a team of nearly one hundred
linguists and was carried out between 1956 and 1958, in ninety places of the Tibet Autonomous Region and Prefectures within China (SKAL.BZANG’GYUR.MED pers. comm. 1988). However, most of the collected data have not been released.

Nishi (1986) distinguishes six major groups: Central (or Ü-Tsang), Western Innovative, Western Archaic, Southern, Kham and Amdo. This classification is very similar to the one used by Bielmeier et al. (2018) in their Comparative Dictionary of Tibetan Dialects (CTDT): 1) Western Archaic Tibetan, 2) Western Innovative Tibetan, 3) Central Tibetan, 4) Southern Tibetan, 5) Northern Kham, 6) Eastern Kham, 7) Eastern Amdo Tibetan. The main difference between the two classifications is that the Khams group is divided into Northern Khams and Eastern Kham in the CDTD.

The Comparative Dictionary of Tibetan Dialects (Bielmeier et al. 2018) is a large project carried out by a team of linguists and researchers based in the University of Berne: R. Bielmeier (project director), F. Haller, C. Haller, K. Häslar, C. Haller, V. Hein, B. Huber, Ngawang Tsering, M. Volkart, M. Zemp as well as former members of the CTDT project D. Klapproth, R. Piva, A-K Röthlisberger-Beer and K. Wymann-Jespersen. The CTDT has so far been the largest collection of data on the Tibetic languages in the five countries (China, India, Nepal, Bhutan and Pakistan). The CTDT is divided in two volumes, a noun volume and a verb volume, which present data concerning about sixty “Tibetan dialects.” As a rule, the main entries are the Written Tibetan etymological equivalents.

The CTDT is a unique work which allows the comparison of many Tibetic languages with Classical Tibetan. The entries correspond for the most to the entries of the great Tibetan Tibetan Chinese dictionary, BOD RGYA TSHIG, MDZOD CHEN, MO, but some entries are also from Jäschke’s dictionary (1881). Since the dictionary entries are organized according to the Classical Tibetan forms, i.e. from an onomasiological perspective, the meaning in the modern languages may be

1. See e.g. the site http://www.himalaylanlanguages.org/cdtd. It provides the following information: Work on the Comparative Dictionary of Tibetan Dialects commenced in 1992. As the culmination of sixteen years of collaborative effort, the Comparative Dictionary of Tibetan Dialects was completed in 2008, and pre-print copies together with A Short Guide to the CDTD were circulated amongst colleagues from that year.
quite different and moreover a given form may lack reflexes in some of the modern languages. This makes it difficult to compare a given meaning in the various modern languages, but it is still possible thanks to an index of the meanings in English.

An additional publication deserves a special mention, [SUM, BHA DON, GRUB TSHE, RING's BOD-KYI YULSKAD RNAM, BSHAD (The Tibetan dialects) (2011)]. This work is the most complete presentation of the Tibetic languages written in Tibetan language to this day. It uses a very accurate linguistic terminology in modern literary Tibetan. The only caveat is that it also includes in the classification a few rGyalrongic, Qiangic and Bodish languages.

9.2. Our classification

The inner classification of the Tibetic family is a very complex task because of the extraordinary dialectal diversity and because this linguistic area extends over an immense territory. Let us remind again that if the Tibetic linguistic area would constitute a sovereign state, it would be the tenth largest country in the world (after India, seventh place Argentina, eighth place and Kazakhstan, ninth place). The extreme linguistic diversity is probably partly due to the geographic environment and the fact that the Tibetic area corresponds to the highest plateau in the world, interspersed with mountain ranges and large rivers which constitute natural obstacles to human transportation (see Chapter 2 and Appendix 1). These obstacles are yet relative if we consider the Tibetan history. Tibetans have always traveled extensively throughout the High Plateau for economic, religious or military reasons. These constant population movements, the scarce population figures, the high level of illiteracy together with the lack of a real standardisation have resulted in a very complicated linguistic and sociolinguistic situation. The various dialects often became quite interwoven. For example, dialects of Kham are also spoken in Ngari, Gertse, Bhutan and even Myanmar (Burma) (see below, 9.3). Conversely, some dialects of Amdo are spoken in Kham province, among the pastoralist population (see below, 9.5).

The general characteristics that we have just mentioned suggest that in at least some areas, we are dealing with a rather fragile ecolinguistic system. A significant number of Tibetic languages are spoken by less than 10,000 speakers and some by less
than 1,000 speakers, as we have seen in Chapter 2. These languages, some of which are not well documented are clearly endangered. Additionally, the existence of isolated non-Tibetic languages spread throughout the Tibetic area and particularly in the Kham area, in South-western Sichuan and Yunnan, but also in Bhutan, Nepal and India makes the linguistic situation even more complex.

Another more specifically linguistic reason which makes the classification task very difficult is the existence of a geolinguistic continuum or, more precisely, a set of geolinguistic continua. In theory, the geographic area of a given dialect must correspond to the convergence of phonological, lexical and grammatical isoglosses. However, there is rarely a total convergence of these criteria. For an impressive and visual illustration of this complex situation, one can look at the chart of Kenhat and Shamskat isoglossis in Zeisler (2004). In the case of a dialectal continuum, there are rarely absolute or natural borders. The boundaries established by the linguists depend on selected parameters.

In our classification, we have listed forty-five groups of dialects in the TAR and TAPs of China and thirty-one groups of dialects in the five other countries, with a total of seventy-six groups of dialects (see also the maps in Appendix 3) divided in 308 dialects.

If we consider the number of administrative counties or districts in the six countries (slightly more than 200) of the Tibetic area, we see that the dialectal diversity is usually quite high since there are usually more than one dialect per county. The level of dialectal differentiation may be at the lower level of townships. Moreover, it should also be noted that administrative units do not necessarily follow dialectal boundaries. Thus in many cases, one may encounter several distinct languages or dialects within a single county or district. For example, in Nyemo County (TAR), one can hear varieties of Tsang, Ü, and Hor. Ü and Tsang belong to the same group, whereas Hor belongs to a different “group of dialects” (or “language”). Another example concerns the village of Drip located less than ten km from Lhasa on the other side of the Kyichu river. The dialect spoken in this village had some specific features, however due to the rapid urbanization of the capital, it is likely than these dialectal specificities are disappearing.
More rarely two (or more) counties may correspond to a single dialectal unit. For example, Chushur or Tagtse Counties in Lhasa Municipality do not currently present significant differences with Lhasa dialect.

The best way to estimate the linguistic variation and to make a relevant classification is to rely on previous publications, documentations and on fieldwork (concerning our fieldwork, see the introduction). However, obviously, a number of countries and districts have not been properly documented and the township level is even less documented. Fortunately, since 1980, the number of fieldwork and publications working on the Tibetic dialectology has risen over the years and it is now possible to have a global picture of the linguistic diversity. This progress has been possible thanks to the involvement of the international community of scholars from many countries or regions of the world: Tibet, China, Bhutan, India, Nepal, Pakistan, Taiwan, Japan, Switzerland, Germany, Austria, Netherlands, France, Italy, USA, Australia, Russia, Czech Republic, Poland, Thailand, Finland, Norway, Israel, Spain, etc.

The classification that we propose here is a tentative one. A more precise mapping of the dialects would require having linguistic data down to the township level.

Based on our fieldwork and on the existing materials, we propose to classify the family into eight major sections: ² South-eastern section (SE); Eastern section (E); North-eastern section (NE); Central section (C); Southern section (S); South-western section (SW); Western section (W); North-western section (NW).

This classification shares some major characteristics with the previous classification of Nishi (1986) and the CDTD. These groupings are based on geolinguistic parameters and the terms used for the various groups are essentially motivated by the geography. These three classifications regroup together dialects in a fairly similar way. However, our classification differs in several ways.

a) We have divided the dialects into eight “sections” while Nishi and the CDTD proposed to distinguish respectively seven or six major groupings.

b) We do not use the term “Tibetan dialect” but instead the term “Tibetic languages”

² Called “net” in Dalby’s terminology.
(see 1.2). The reason, as mentioned earlier is that there is no intelligibility between the various sections, and in some cases, even inside a given section.
c) The classification proposed here is essentially based on a genetic approach, but it also includes the notion of mutual intelligibility as well as geographical parameters, migration and language contact factors.
d) Our classification terminology is essentially geographic whereas the two other classifications mix the geographic parameter and the degree of linguistic innovation (innovative or archaic). The reason to maintain a strict geographic terminology is that the degree of innovation varies a lot depending on the linguistic domains and moreover it varies within the same group. For example, one could claim that the so-called dialects of the “Western Innovative” group, such as the “Kenhat dialects” of upper Ladakh are indeed quite archaic compared to the Spiti, Khunu and Garzha dialects.
e) The CDTD introduces a distinction between “Northern Kham” and “Eastern Kham.” As we shall see, the linguistic divisions within Kham are in fact much more complex than a simple binary division. The Kham region (our Southeast section) corresponds to a very complex geolinguistic continuum which we subdivide into twelve subgroups.
f) There are some overlaps in the grouping of the three classifications. For example CDTD lumps together the Ngari Tholing dialect of Western Tibet with the “Western Innovative Tibetan” located outside Tibet, whereas we group it with the other dialects of western Tibet within the Central section. Since the Ngari dialects are part of a geolinguistic continuum, thus it is difficult to justify the CDTD choice. Moreover the boundaries of the so-called Tholing dialect are not specified. Another overlap between the groupings concerns the grouping of the northern dialects of Nepal together with “Central Tibetan.” This is not convincing, since the distance between Sherpa, Jirel and the dialects of Central Tibet seems greater than the one between these dialects and those of “Western Innovative Tibetan,” which is classified separately.
g) One of the main differences in our approach is to introduce a separate group
“Eastern section.” Languages of this area such as Čone, Thewo, Drugchu, Baima, etc. are clearly neither Amdo (our NE section) nor Kham (our SE section).

h) Apart from the strictly linguistic parameters, our classification also takes into account the geopolitical boundaries. In fact, these boundaries do have a strong impact on the linguistic situation because in this region of the world political boundaries clearly limit the exchange between the two communities on each side of the border. Another reason to place Northern Tibetic dialects of Nepal into a SW section, i.e., a distinct section specific to Nepal is that all the Tibetic languages have borrowed loanwords from Nepali, the national language, whereas modern loanwords in the Central section usually come from Chinese.

As we just explained the geopolitical borders are taken into account whenever they are linguistically relevant. Thus four sections are located in Tibet, i.e. currently within the People’s Republic of China3 (the Northeastern, Eastern, Southeastern and Central sections), while the remaining four sections (the Northwestern, Western, Southwestern and Southern sections) are located outside Tibet in Nepal, Bhutan, India and Pakistan. There are only two minor exceptions of the correspondence between the linguistic classification and the geopolitical border between China and the “Indian subcontinent.” These two exceptions are motivated by linguistic arguments related to their affiliation, but they are also clearly supported by geographic reasons. The two dialects of Kyirong and Dromo are respectively regrouped with the Southwestern section (Nepal) and with the Southern section (Tibetic languages of Sikkim and Bhutan), but the two dialects are spoken in lower valleys that penetrate deeply into the southern Himalayas.

While the Southwestern section is delimited by the geopolitical borders of Nepal, this is not the case with the other sections outside Tibet. The Southern section extends over the Indo-Bhutanese border since it includes the Tibetic languages of Bhutan and Sikkim, which used to be an independent kingdom but is nowadays a state of India.

Before we describe the characteristics of each the linguistic groups, let us briefly give some information about the content of each section:

3. There are only three exceptions, such as Kyirong, Dromo and the Kham dialect of Myanmar which are affiliated respectively to the South-western, Southern and South-eastern sections.
1) South-eastern section (SE)
Northern route Kham, Hor Nagchu, Yülshül, Southern route Kham, Minyak Rabgang, and other smaller dialect groups of southern Kham;
2) Eastern section (E)
Drugchu, Khöpokhok (Zitsakhok), Thewo-tö, Thewo-mä, Čone, Baima, Sharkhok, Zhongu, Pälkyi (/Pashi/), Throchu;
3) North-eastern section (NE)
Amdo dialects as well as dialects surrounding the rGyalrong and Kham areas;
4) Central (C)
Ú, Tsang, Phänpo, Lhokha, Tö Ngari, Kongpo;
5) Southern section (S)
Dzongkha, Lhoke (or Dränjong, Sikkim), Choča-ngrača, Dromo, Lakha, Dur Brokkat, Mera Sakteng, Brokpa-ke;
6) South-western section (SW)
Sherpa, Jirel, Lo-kā and other dialects spoken along the Sino-Nepalese border;
7) Western section (W)
Spiti, Garzha (Lahul), Khunu, Jadang;
8) North-western section (NW)
Ladaks, Zanhar (Zangskar), Balti, Purik.

The rest of this chapter will be devoted to the presentation of the various linguistic “sections” of the Tibetic area. We will provide a general introduction for each section, which includes information about its sociolinguistic situation, its cultures and religions. We will also indicate some of the major institutions such as universities, monasteries and traditional medical centers of a given area. There are two reasons to provide information about some major monasteries. First, they are considered not only as important religious centers but also as cultural institutions and play a significant role in preserving the history and the memory of a given area. Additionally, indications about the school of Vajrāyana or Bön or other religions often provide useful information about the dialectal diversity.
Then for each section, we will briefly examine a) the migration patterns (whenever information is available), b) the various linguistic groups and subgroups of each section, c) the geographic extension of each section, d) an estimation of the number of speakers, e) ethnic and sociolinguistic information, f) some characteristic features of the phonology, g) some characteristic features of the grammar. Given the linguistic diversity of the Tibetic area, it is of course beyond the scope of this chapter to provide a detailed presentation for each language and even less for each dialect. Rather, the idea is to convey general information about each of the eight linguistic sections and refer to the existing literature whenever available. Each section will also contain maps with indication about the localization of the main dialects. Concerning the phonology, we will discuss only a few major characteristics concerning suprasegmental and segmental features as well as reflexes of Classical Tibetan. Concerning the grammar, we will only mention some general information about core grammatical categories of the Tibetic languages such as grammatical cases, nominalizers, verbal inflections, linking verbs (copulative and existential verbs) and auxiliary verbs as well as negation markers.

9.3. The South-eastern section

The South-eastern section (henceforth SE section) roughly corresponds to the linguistic area of the “Kham (or Kham-Hor) group of dialects” (according the traditional classification, see above), spoken on a territory extending over Sichuan, Tibet Autonomous Region, Qinghai and Yunnan of China as well as an enclave of Myanmar. However, the linguistic diversity within this “group” is so large that we can not regard it as one language (see 9.3.2). There is no mutual intelligibility between some of the Kham groups. For us, the terms Kham and Hor refer to geolinguistic areas and not to language groupings.

Khampas are traditionally either cultivators or pastoralists (particularly in Northern Kham) or agropastoralist called samdrok ཞབས་མ་འབྲོག་ or yulmadrok ཡུལ་མ་འབྲོག་ (see Chapter 2) The majority of Horpas are pastoralists, but some agropastoralists are also found.

Mandarin Chinese is currently the official language in the Kham area of Sichuan, Qinghai and Yunnan but in the Hor and Kham areas of TAR, Tibetan also has an
official status. However, in both cases, Chinese is de facto the dominant language of the school curriculum.

The languages of the SE section can fairly easily be transcribed in Tibetan script but they are usually not written down, and when people write in Tibetan, they normally write in literary Tibetan. This written language is used in the Buddhist and Bönpo monasteries, in the institutes of Traditional medicine and to a certain extent in some cultural organizations and media, particularly on the internet. Depending on their home province (Sichuan, Yunnan, Qinghai or TAR), Kham and Hor students may attend one of the universities which have Tibetan language departments: the Southwest University for Nationalities (Lhonu Mirik Lobdra Chenmo in Chengdu, the Sichuan College for Nationalities (Sithrön Mirik Lobling in Kangding (Sichuan), the Qinghai University for Nationalities (Tshongön Mirik Lobdra Chenmo in Xining (Qinghai) and the Tibet University (Böjong Lobdra Chenmo) located in Lhasa.

In Kham and Hor (Nagchu) areas, Vajrayāna Buddhism is the main religion and the four sects (Nyingmapa, Kagyüpa, Gelugpa and Sakyapa) are represented. Bönpo communities are also found in the region. Small Christian communities are attested in southern Kham.

Major monastic institutions of the SE section include: Gandän Jampaling in Chamdo, Gandän Sumtsenling in Gyalthang, Gandän Kandze Gön (Geluk), Dzogchen Monastery (Nyingma) in Derge County; Riwoche Tsuglagkhang (Kagyü); Tenchen Gön and Tse Tsedruk (Bön) both in the Tengchen Khyungpo area. One also ought to mention the major printing house, Derge Parkhang.

Among the main studies on the Kham dialects one should mention: SKAL BZANG 'GYUR MED & SKAL BZANG DBYANGS CAN (2002) and Hasler (1999) on Derge, Bartee (2007) on gTormarong (Dongwang), Suzuki (2011b) on dGudzong, and
Kraft & Hu (1998) on Minyak Rabgang and Derge. There are other grammatical sketches regarding Zhollam (Suzuki 2010), Sakar (Suzuki 2012b), Choswateng (Suzuki 2014a), Nangehen (Causemann 1989), Drag-yab (Schwieger 1989), and Lhagang (Suzuki & Sonam Wangmo 2016a). There is also a Comparative Dictionary of Southern Khams Language Varieties (Bartee & Hugoniot, 2020: webonary.org.SIL International).

9.3.1. Migration patterns, legends and historical records

A Kham variety related to Yülshül or Nanchen dialects is spoken in Gertse and Gegyä counties (Ngari Prefecture). Settlements of Khampas in Ngari (Western Tibet) have been reported by foreign travelers over a century ago. In the History of Ngari Korsum, GESLONG BSTAN’DZIN DBANG.GRAG (1996: 269) makes a brief mention about the historical migration from Kham:

“From oral accounts as well as archival sources, seven families known as the ‘seven families of Teryik’ that were from Kham made a pilgrimage to Tö and settled in a place which is now called Gertse. This toponym is the same as a pastoralist settlement in the Yülshül area in Kham. The people of Gertse have very similar traditions and dialectal pronunciation as the people of Yülshül so we consider that they originally came from there.” (our translation)

In Ngari area, nomadic tribes, mainly Hor, have migrated from the Nagchu area. These migrations are confirmed by the linguistic data and the various toponyms found in the Kailash area: Hor, Möntser (which indicates Mön or Bhutanese origin) and Gertse (a Northern Kham toponym). One of the striking characteristics of Ngari is the way Hor and Tö pastoralists live together and yet have kept their own dialects, which belong respectively to the Kham-Hor group (our South-eastern section) and the Ü-Tsang group (Central section). Kham pastoralists are also found in Eastern Bhutan, in Trashi Yangtse district.

Some Kham speakers have migrated from Dzayül to Myanmar less than two hundred years ago. They have settled in four villages of the Dazundam village tract in

4. In Ngari, Hor usually refers to Uighur people. In Ngari there are also two Mongol tribes: SOG.STOD and SOG.SMAD which are both situated in Gar. They still celebrate the Mongol New Year and follow the same calendar.
the Kachin State, Myanmar, on the China-Myanmar border, near the Tibet Autonomous Region. They still have relatives in the Dzayül villages on the other side of the border.

9.3.2. Linguistic groups of the SE section

The dialects of the SE section present a greater diversity than Amdo or Central Tibet. Some dialects such as Rongdrak, Semkyi Nyida (locally pronounced /Shanggi Nyila/) and Khyungpo (pastoralist dialect) are so distinct that they do not allow for easy communication with other surrounding linguistic groups.

In the north, the Hor dialects (Nagchu, etc.) and the Kham dialects of the “Northern Route” (Chamdo, Derge, Kandze) and, to a lesser extent, Yülshül allow a fairly good mutual intelligibility. In the east, there are two smaller groups of dialects: Minyak Rabgang and Rongdrak. The former allows good intelligibility with dialects of Northern Route, whereas the latter is considerably different from any Tibetic varieties spoken in Kham.

The dialects spoken on the “Southern route,” which include Bathang ཐི་བཐང་, Lithang བི་ཐང་ and Markham སངས་ཁམས་ also allow some mutual intelligibility.

Further in the south, one finds four groups: Derong-Jol སྡེ་རོང་འཇོལ་, Semkyi Nyida (/Shanggi Nyila/) སེམས་ཀྱི་ཉི་ཟླ་, Chagthreng གཞག་ཤྱེང་ and Pomborgang གཉིས་བོད་ཀྱང་ (formerly Muli-Dabpa ཚུ་མི་ལི་དང་འདབ་པ་). These have a very low intelligibility between each other and, of course, with the dialects spoken further in the north (concerning lexical variation regarding Tibetic languages in Yunnan, see Suzuki 2018, 2022).

Thus, it is clearly not suitable to speak of “one Kham language,” even less of “one Kham dialect,” since we have at least 4 major dialect groups or “languages” (i.e. Hor, Northern Kham, Southern route Kham and Minyak Rabgang) and many smaller groups. Nevertheless, we can still make use of “Kham Tibetan” in order to refer to the Tibetic languages spoken in the Kham region.

5. Muli is often written སྨི་ལི་ Mili in Tibetan, however the etymology and the history of the toponym is not clear and we will continue to use the spelling Muli. Suzuki (2018) has proposed to replace the name of the group Muli-nDabpa by “Pomborgang group.”
For the classification of the SE section, we propose the following fourteen groups:

- **Nagchu** (traditionally called Hor dialects)
  - Nagchu, Amdo, Nyānrong, Sok, north Biru (Driu), Lhari, Pāngon, Tshonyi, Shāntsa, Nyima-Hor, Tshochen-Hor, Gertse-Hor, Damzhung and Nyemo.

- **Drachen/Bachen**
  - West Bachen (Drachen). This variety is also considered traditionally as a Hor dialect.

- **Kyegu**
  - Kyegundo, Nangchen, Chumarlep-Kham, Thridu-Kham (partly), Dzātō, Driō, This group also includes Gertse Kham and Gegyā Kham spoken in Ngari, which originally come from Kyegu area.

- **Pāmbar**
  - Pāmbar and South Biru (Driu).

- **Khyungpo**
  - Tengchen and Bachen (east).

- **Northern route (Chamdo, Derge, Kandze)**
  - Chamdo, Jonda, Gonjo, Drayap, Riwoche, Lhorong, Kandze, Nyagrong (Kham), Pālyul, Derge, Sershiul, Drango (Kham) and Tau (Kham).

- **Rongdrak**
  - Rongdrak, Pronang (Rongdrak dzong), Jintang in Dartsendo (East) and Khoryül in Dartsendo (NE).

- **Minyak Rabgang**
  - Rangakha in Dartsendo (West), Drongsam (East Nyagchukha), Morim, and Basme (spoken in Tau).
Southern route (Markham, Bathang, Lithang)

Bathang (east), Lithang-Kham, Gola (Dzogang), Markham, Pasho, Pomu, Potö (in Powo district), Nyagchukha-Kham (West), Dzayül (central), Metok (partly) and probably Sangdam, spoken in Myanmar.

Dzayül

Dzayül Gola (Dzayul East), Dzayül Tö, Dzayül Rongmä, Dzayül Drungchu and Tshawarong.

Derong-nJol

Derong, nJol-Tö (Dechen) and Pomtserag.

Chagthreng

Chagthreng, Rwata (in NW of Chagthreng and in a small area in the north of Derong) and Tormaron (Dongvang).

Pomborgang

Cingdrol (in the north Dabpa), Mundzin (south of Dabpa), Mairi (north of Muli), Mola (south of Lithang) and Nyayülzhab (South Nyagchukha, Darmdo Minyag (Southernmost) and Gyäzur Minyag (Northernmost)).

Semkyi Nyida

Nyishar (West of Gyälthang and a small area of Dechen and Balung), Gyälthang, Yangthang (south of Gyälthang), Thachu (central area of Balung), Zhollam, Melung (also in Balung), Lijiang, Yongsheng and Yongning and Muli (south).

The great linguistic complexity of Kham might be due partly to the existence of pockets of other Tibeto-Burman languages namely Situ, Geshitsa, sTaus, Nyagrong-Minyag, Darmdo Minyag, nGochang, nDrapa, Choyu, Lhagang Choyu, Lamo, Larong sMar, Drag-yab sMar, Ersu, Lüzu, Nosu, Prinmi, Shuering, Namuzi, Na (or Moso), Laze, Naxi, Malimasa, Lisu, Bai, Anung and even Miao (see Chapter 10).
9.3.3. Geographic extent of the SE section

The SE section extends over the historical Kham region (ཁམས་) as well as the Hor region (ཧོར་) located mainly in Nagchu area and in the eastern Janghang (ཇང་ཧོང་). This area of Do-Kham (མདོ་ཁམས་) is drained by major rivers and it is for this reason sometimes referred to as Four Rivers Six Ranges (ཆུ་བཞི་སྒང་དྲུག):

The six plateaux called སྒང་དྲུག་(gang druk) are: ཕྱིན་ཐང་ཚལ་མོ་སྒང་ (Zälmo Gang), ཕྱིན་ཐང་ཚ་བ་སྒང་ (Tshawa Gang), བོད་ཁམས་སྒང་ (Markham Gang), བོད་ཁམས་སྒང་ (Pombor Gang), མོང་ལྷེའི་སྒང་ (Mardza Gang), and རྡོ་རྗེ་སྒང་ (Minyak Rabgang). The four rivers are: in the west, the Gyälmo Ngülchu river (རྒྱལ་མོ་རྔུལ་ཆུ་) (upper course of the Salween or Nu Jiang), the Dachu [Dzachu/Lachu] river (ཝྲོ་ཆུ་) (upper course of the Mekong), and the Drichu river (འབྲི་ཆུ་) (upper course of the Yangtze) as well as in the east the Nyagchu river (ཉག་ཆུ་) (Yalong Jiang) a large tributary of the Yangtze. Additionally, one should mention two other large rivers, found respectively at the western and eastern limits of Kham: the Dzayül river (རྲི་བུལ་ཆུ་) (known as གསང་ཆུ་Sangchu in its upper course), which flows in the south-eastern TAR (China) before crossing the border with Arunachal Pradesh (India), where it becomes the Lohit river, and in the east, the Gyälmo Ngülchu river (རྒྱལ་མོ་རྔུལ་ཆུ) (Dadu river), which flows east of Dartsendo (Kangding).

The SE section corresponds to Chamdo District (TAR), the south-eastern part of Nyingthi Municipality (TAR), Kandze TAP (Sichuan), Muli TAC (Sichuan), Yülshül TAP (Qinghai) and Dechen TAP (Yunnan). The Hor linguistic area is located mainly in Nagchu Prefecture (TAR) but extends to some counties of Lhasa and Ngari Prefectures.

This linguistic area extends over a huge territory in the south-eastern region of the Tibetan speaking area. It covers about 600,000 km².

---

6. The difference of these names is due to the phonetic reflex of the word ZLA.CHU.
7. According to some commentaries, the fourth river is sometimes the KMA.CHU (Huang He) which flows through Amdo. However, since the Four Rivers Six Ranges refer to Kham, the inclusion of the Huang He is not likely to be valid.
8. One should not confuse the Salween and the Dadu rivers which bear the same name in Tibetan: RGYAL.MO.RNGUL.CHU.
The map IX.2 focuses on the Kham region, reflecting the linguistic variation attested within eastern Kham.
MAP IX.2. – Linguistic classification of the Kham area

Legend: ■: Northern Route; ○: Rongdrak; ●: Minyak Rabgang;
■: Southern Route; ●: Pomborgang; ○: Chagthreng;
■: Derong-njol; ○: Semkyi Nyida; ●: Dzayül
Detailed location of the Kham and Hor dialects

Here is a list of the counties where Kham dialects are spoken:

- in Chamdo Municipality (TAR): Chamdo, Tengchen, Jonda, Gonjo, Drag-yap, Dzogang, Riwoche, Markham, Lhorong, Pasho, and Pambar;
- in Nyingthi Municipality (TAR): Pomā and Dzayül;
- in Ngari Prefecture (TAR): Gertse and Gegyä;
- in Kandze Prefecture (TAP, Sichuan): Kandze, Dartsendo, Drango, Nyagrong, Palyü, Derge, Sershül, Rongdruk, Gyäsur, Nyagchukha, Tau, Bathang, Lithang, Derge, Sershül, Dabpa, and Derong;
- in Muli (TAC, Sichuan): Muli;
- in Yülshül Prefecture (TAP, Qinghai): Kyegundo, Nangchen, Chumarlep, Thrindu (partly), Dzatö, and Dritö;
- in Dechen Prefecture (TAP, Yunnan): Semkyi Nyida, Dechen, and Balung;
- a few settlements of Kham speakers are also found in Lijiang Municipality, Xuehua (Maoniuping) Village of Yulong District, in Yongsheng County (Daan Township), in Ninglang (Yongning Township) and Gongshan County (Bingzhongluo and Bangra Townships) located in Nujiang Prefecture;
- in Myanmar (Burma): Dazundam Village Tract;

Any linguistic description on this variety is unavailable; hence, it is still impossible to claim to which dialect group of Kham this variety belong.

Here is a detailed list of the counties where Hor dialects are spoken:

- in Nagchu Municipality (TAR): Nagchu, Amdo, Nyänrong, Drachen (Bachen), Sok, Drirü (Biru), Lhari, and in the following counties (together with a Tö pastoralists dialect): Pangön.
9.3.4. **Number of speakers**

The total number of speakers of the Tibetic languages located in the SE section is hard to ascertain due to the lack of a reliable census and to the linguistic diversity of the area. According to Qu (1996), the total of Kham and Hor Nagchu speakers is about 1,500,000.

This figure might be slightly underestimated. However, one should bear in mind, that given the linguistic diversity in the area, the population can not be considered speakers of one single language. Thus, to have linguistic relevance, a census should be based on a precise inner classification of the SE section.

9.3.5. **Ethnic and sociolinguistic groups**

The South-eastern section corresponds to the grouping of dialects spoken mainly by two slightly distinct Tibetan ethnic groups: དཔལ་ 'Khampa' and སྨོ་ 'Horpa' herders from the Jangthang area, also called གཡུལ་ 'Apho Hor' (this term is sometimes slightly derogatory or ironical). Qu Aitang (1996) has referred to Hor dialect as Western Kham. We do not follow this suggestion because Horpas do not consider themselves as Khampas.

Additionally, in Rongdrak area, people speaking Kham dialects usually define themselves as Gyäróngwa (རྒྱལ་རོང་བ་) and in the southern most region of Kham, in Dechen and Gyäthang, people until recently defined themselves according to the local regional names such as Gyäthangwa (རྒྱལ་ཐང་བ་).

In Kham, the cultivators’ and the pastoralists’ dialects are distinct. There are also many communities of agropastoralists called ས་མ་འབྲོག་ samadrok or ང་མ་འབྲོག་ yulmadrok (for alternative terms, see 3.1). Nagchu Horpas are essentially pastoralists རྡོག་ droga or in some rare cases agropastoralists.
Note that the Tibetans of Muli, who speak a Kham Tibetan dialect, are sometimes called “Kami,” but this is an exonym.

9.3.6. Phonological characteristics of the SE section

It is not possible to list common phonological features to all the dialects of the SE section. The phonological characteristics are usually valid only at the level of the groups or even sometimes the dialects.

Suprasegmental features

The Tibetic languages in the SE section have a pitch tone system without exception. Multiple types of tonal distinction are attested: a majority of dialects have a four-way contrast (high, rising, falling and rising-falling) and a minority of dialects have either a two way-contrast (high and low) or five way-contrast (high, low, rising, falling and rising-falling).

Segmental features

Synchronic approach

The sound systems of the SE section are characterized by the following frequent features:

- Existence of voiced non resonant sounds (b, d, d, g, dz, j, z, zh, ɣ, f).
- Voiceless nasals (m', n', ng', ny').
- Prenasalization is pervasive (b', d', g', etc.). In many Kham dialects, combinations such as t', p', k', etc., are also found.
- Most Kham dialects have an aspirated fricative series: s', x', sh'.
- Preaspiration is found in many Kham dialects (t', h', k', d', b', g', s, etc.).
- A limited set of final consonants (except in some north-western dialects).
- A large set of vowels and the existence of /ə/. The opposition between /ɑ/ and /a/. The distinction between long and short vowels is quite frequent.

---

9. Not found in Hor dialects. In the phonetic alphabet, the voiceless nasals are noted as [m̥, n̥, ŋ̊].
Diachronic approach and reflexes of Classical Tibetan

- In the Kham groups, the reflexes of preradical sounds are realized as preaspiration including prenasals. A minority of dialects such as Khyungpo have traces of segmental realization for the preradicals.
- Voiced non resonant sounds (b, d, g, dz, j) are derived from the consonants with preradicals (except M and ʼ).
- Voiceless nasals (m’, n’, ng’, ny’) are the reflexes of nasals with the preradical S. This is the case in words like སྨན་ ‘medicine’, སྣ་ ‘nose’, སྙིང་ ‘heart’, སྔོ་ ‘blue’.
- Prenasalization correspond to the reflex of radicals with a predical M or ʼ.
- As for aspirated fricative series, CT S, SH, PHY without any preradicals correspond to s’, x’, sh’ (in the northern route Kham), but the sound correspondence is different depending on each dialectal group. In some Kham dialects, SR without preradical B also corresponds to an aspirated /s’/.
- Preaspiration in many Kham dialects is triggered by all the CT preradicals (except M and ʼ).
- CT final consonants B, D, G often changed into a glottal stop; M, N, NG triggered the nasalization of vowels. The consonants R, L and S often caused the lengthening of the vowel.
- The final consonant triggered the change of tongue position resulting in a vowel change. In addition i without final is usually realized as /ə/.
- In Hor and northern Kham dialects, the combination SL and ZL respectively yield: /ts/ and /dz/. This is the case in words like སླ་མོ་ ‘easy’, སློབ་ ‘to teach’, སླེབས་ ‘arrive’.

9.3.7. Grammatical characteristics of the SE section

As we will see from the grammatical sketch below, the dialect groups of the SE section exhibit significant differences in their grammar. For example, concerning the verbal inflectional morphology, the northern Kham dialects have well preserved two or three inflections inherited from CT whereas the southern dialects have lost these
forms and the lexical verb has become invariable. This feature is shared with the languages of the S section such as Dzongkha and Lhoke. The languages of the SE section also largely differ in their verb auxiliaries from the point of view of both form and function.

9.3.7.1. **Case markers**

As we will see, cases in the various Tibetic languages differ in form, function as well as the number. However, one can say that all the various case systems minimally include three case markers: genitive/ergative, dative and ablative, as well as the unmarked absolutive case. In most of modern Tibetic languages ergative and genitive case markers are either identical or very similar.

Some dialects of the SE section distinguish up to seven or eight case markers. Frequent cases include ergative, absolutive, dative, ablative, genitive and locative,\(^{10}\) comparative and associative.

We will first present the “core syntactic cases” and then introduce the peripheral cases (or local cases). The case markers that indicate core syntactic roles are the absolutive, the ergative and the dative.

The ergative is used with both controllable and non-controllable verbs. It is found in all the dialects of the SE section. It is used in the various tenses (present, past, future) and aspects (completed and uncompleted). However, the ergative is only compulsory in some contexts otherwise it is optional and used to show an emphasis on the grammatical agent. In some southern Kham dialects, such as Melung (Sems-kyi-nyila), ergative has acquired a more pragmatic function.

In the SE section, one finds the following markers for the ergative: /-ki/, /-ka/ and other forms derived from ྡིས་GIS. However, some dialects have an ergative as: /-je/ which may be derived from the allomorph: མི་YIS. Some dialects have special ergative forms for the personal pronouns (Derge, Kandze, Chamdo, Rongdrak, Khyungpo).

\(^{10}\) Depending on the languages, the semantic meaning may be a general locative, an inessive or an illative.
In the SE section, ergative and genitive markers may be identical as in Derge or still distinct as in many dialects of Melung, but usually they have a very similar form.

The absolutive is always marked by zero marker (Ø) in all SE languages. In some dialects such as Derge and Rongdrak, the Sa (the subject of an intransitive controllable verb) is optionally marked with an ergative, while the Sp (the subject of an intransitive non-controllable verb) must be in the absolutive.

The dative marker is derived in most dialects from CT དལ་ LA. The vowel varies a lot depending on the dialect: /-le, -la, -lā, -lo/. In a marginal way, one finds other forms whose origin is unclear: /-tsa/ (most dialects spoken in Yunnan). It might be derived from the word རྩ་ RTSA ‘root’ which is very often used as a postposition and even a case in some dialects of other sections.

The local or peripheral case markers are ablative, genitive, comparative, locative (inessive, illative, etc.).

The three main forms for the ablative are བོད་ NAS /-ne:/ as well as /-ts'a/ and /-de/. The origin of the /-ts'a/ and /-de/ is unclear. The former is quite common in Yunnan (except for Melung area), while the latter is found in Rongdrak and Melung. In some dialects, the ablative is used for the comparative.

The comparative is /-ya/ /-yi/ in some Kham dialects (Derge), a form probably derived from CT ཡིས་ YIS. Another form /-we/, also derived from CT བས་ BAS (also having a comparative function), is found in Gyalthang and Chamdo. The dative is also used for the comparative function, for example in Derge (see Häslar 1999: 119). In Derong a special form /-kərə/ is found.

For the locative case, the dative དལ་ LA is often used. Some dialects have a special form with a more specific locative meaning such as inessive or illative. For example བོད་ NAS /-na/ is found in Derge. Another form /na/ is widespread in Minyak Rabgang, Rongdrak, Derong, etc. It is probably derived from the classical locative case བོད་ NA or from བོད་ NANG ‘interior’. In some dialects the two forms can alternate.
The instrumental in many SE dialects is formally identical to the ergative case གིས་GIS. However, a few dialects have a special form: /-ka/ (Zhollam) or /-he/ (Gyalthang).

The genitive form is usually derived from གི GL. As mentioned above, genitive markers are often distinct from the ergative, although they may be very similar. Yunnan dialects have an alternative form གི་/da/ (of unclear origin). In the Melung subgroup, there is also an exception form for the genitive གི་/k’ong/ < CT KHONGS ’to belong’.

The associative /tö/ derived from the CT form གི DANG (see Hässler 1999). 11

9.3.7.2. Nominalizers

Nominalizers play a major role in the Tibetic languages. They have a specific argument function related to the verb which is nominalized. They may indicate the agent of the verbal action, the place, the agent, the patient, the instrument, etc. A characteristic feature of Tibetic nominalizers is that they also function as relative clause marker and some nominalizers may also be used with auxiliaries to form verb endings.

Various nominalizers are found in the SE section. We list below eight relatively frequent nominalizers of this section. Some dialects may have only three markers. The nominalizers have a specific argument function related to the verb which is nominalized.

- In Kham and Hor, there is a very widespread nominalizer, མི MI, which is pronounced in various ways /ma, mo/. Other forms such as མི་NÌ, /na, -no/ and མྱི MI /-nya/ are also attested. All these are probably derived from མི MI or its archaic variant མྱི MI which originally means ‘person, human being’. One strong argument in favor of this hypothesis is that in the given Kham dialects, ‘human being’ is also pronounced /na/ (Gyalthang) or /nya/ (Chagthreng). This nominalizer has different functions depending on the dialect. It may indicate the general nominalization (infinitive-like), the patient or the agent of the verbal action (‘the doer’, ‘the thing to be done’ or ‘the thing done’).

11. It seems that the associative case unlike some other dialect groups and CT is not commanded by lexical verbs in the SE section.
The nominalizer \( \text{མཁན} / \text{MKHAN} \) is found in some Kham (Khyungpo, Derge, Minyak Rabgang) and Hor dialects. It usually indicates the agent of the verbal action.

The nominalizer \( /\text{pa}/ \) identical to the universal nominalizer of CT is found in several Kham dialects (Notably Derge).

In Hor (Nagechu) and some northern Kham dialects (e.g. Jonda, Derge) one encounters the nominalizer \( /\text{le}/ \) which indicates the patient of the verbal action similar to \( /\text{pa}/ \) in CT.

Another frequent nominalizer is \( /\text{s'a}/ \). It is pervasive in Kham and Hor dialects. This nominalizer is derived from CT \( /\text{sa}/ \) ‘place’ and has various meanings such as the place of the verbal action and the instrument of the verbal action (e.g. Rongdrak).

The form \( /\text{zä}/ \) is found in some Kham dialects such as Derge to indicate the instrument of the verbal action. It is probably derived from CT \( /\text{rdzas}/ \) ‘thing, object’.

The nominalizer \( /\text{ja}/ \) or \( /\text{gya}/ \) is found in several Kham (Derge, Nagchu, Rongdrak, etc.). It is derived from CT \( /\text{rgyu}/ \) and indicates the patient of the verbal action. For this function, the nominalizer \( /\text{da}/ \) (of unclear origin) is used in Melung.

The nominalizer \( /\text{stang}/ \) is used in several northern dialects (Derge). It indicates “the way of the verbal action.”

### 9.3.7.3. Verbal inflections

In most southern Kham dialects, the lexical verb is invariable and has no tense or modal inflections (see 8.3.8). In the northern Kham and Hor dialects, verbs have often two or three forms. The verbs with two forms have an opposition between “present” and “past-imperative” or between “present-past” and “imperative.” The verbs with three forms distinguish “past,” “present” and “imperative.” The Khyungpo Thromtshang dialect is exceptional and has a few verbs with four inflectional forms (“completed,” “uncompleted,” “future” and “imperative”), however these forms are generally innovative and not inherited from CT.
In some dialects a few verbs have suppletive forms: ‘to go’ ལེགས་SONG and ལེགས་‘GRO; ‘to come’: ལེགས་‘ONG and ལེགས་‘SHOG; ‘to say’: ལེགས་‘ZER and ལེགས་‘BZLAS.

The general trend shows a reduction of verbal inflections and the younger generations have a tendency to use a single verb form. This trend can also be observed in the dialects of Ü-Tsang.

9.3.7.4. Linking verbs and auxiliary verbs

Copulative verbs

The verbs ལེགས་YIN and ལེགས་‘RED are used both as equative copulative verbs. The copula YIN functions as an egophoric meaning whereas RED indicates a factual meaning. These two markers are widespread in the SE section, however, in a few dialectal groups, instead of ལེགས་YIN and ལེགས་‘RED, other pairs are used: ལེགས་‘YIN vs ལེགས་‘YIN-SNANG (Rongdrak), ལེགས་‘YIN vs ལེགས་‘SNANG (Melung). In Gyälthang, one can find a form ལེགས་GRAG with a non visual sensory evidential meaning.

These auxiliary verbs have the following negations: ལེགས་YIN or ལེགས་‘OD > ལེགས་MED or its archaic form ལེགས་‘MYED.

Existential verbs

Existential auxiliaries exhibit a greater variation both in form and meaning. རོག་‘YOD or its variant རོག་‘OD are found in all the dialects, however they convey different grammatical meanings.

A) In many dialects, རོག་‘YOD; or its variant རོག་‘OD, is used to refer to personal information (egophoric) and generally occurs with the first person subject.

B) In southern Kham (e.g. Gyälthang, Derong, Dechen, etc.) རོག་‘YOD indicates the possession and may be used with possessors referring to any of the three persons.

These auxiliary verbs have the following negations: རོག་‘YOD / རོག་‘OD > རོག་MED or its archaic form རོག་‘MYED.
In many dialects the following auxiliaries indicate sensory access to information, often visual but not exclusively: བདའ་ (e.g. Hor, Yülshül and many herders' dialects), འགི་ (e.g. Derge, Kandze, Lithang), SNANG (southern Kham), DUG (Gyälthang, Derong, Mulō). For this function, Minyak Rabgang uses a compound form for the existential verb: ཡོད་དུ་. In Southern Kham, some of the above auxiliaries have acquired different grammatical meanings. For example, in Gyälthang, SNANG is used to indicate exclusively visual perceptions whereas DUG indicates the existence of animate beings especially humans. In some dialects (Bathang, Derge and Gyälthang, etc.), GRAG is used for non-visual perceptions.

**Compound linking verbs**

Compound verbs are very frequent in the SE section. They involve the combination of various copulas or auxiliaries. Sometimes, the first verb is followed by a relator (connective or nominalizer). ཡོད་, རེད་, GDA’, འགི་, SNANG, DUG.

The following combinations are frequently found in northern Kham and Hor: ཡོད་ རེད་, ཡོད་ འགི་, ཡོད་ སྣང་, ཡིན་ རེད་, ཡིན་ འགི་, ཡིན་ སྣང་, etc.

**Auxiliaries**

Frequent auxiliary verbs of the SE section consist of linking verbs (sometimes preceded by a relator): ཡོད་, རེད་, GDA’, འགི་, SNANG, DUG.

Other frequent auxiliaries include: THAL (all dialects), BYUNG (most dialects), GRAG (many Kham dialects), ZUG (Derge, Kandze), DO (Derge), KHAL (njol, Minyak Rabgang).

From a semantico-cognitive point of view, the auxiliaries (together with relators) convey complex temporal, aspectual, evidential and epistemic meanings.

Most of the systems have special forms to mark sensory access to information, as well as factual, egophoric, inferential, hearsay and epistemic meanings.
9.3.7.5. Negation

The negation has two forms in all the dialects of the SE section: ང་ MA and ངི་ MI (Kham, Hor) or its archaic form གྱི་ MYI (in Yunnan) which occur before the verb or the auxiliary. In general, for the imperative, one can only use the form ང་ MA. In synchrony, negative morphemes ང་ MA and ངི་ MI together with auxiliaries constitute one single verb affix, which cannot be analyzed. This suffix conveys TAM and evidential-epistemic values as well as negation.

In addition to these, especially in Southern Kham, another negation prefix is emerging: གར་ GAR, which originally denotes ‘where’ as an interrogative word. Gyalthang Kham has already grammaticalized it as a negation prefix, and it functions as a negation, mostly used for the negation of egophoric inferential (see Suzuki & Lozong Lhamo 2020).

9.4. The Eastern section

The Eastern section (henceforth E section) is made up of several linguistic groups of dialects spoken in a relatively small area in northern Sichuan and southern Gansu. The great linguistic diversity of this section may be due to the contact with other Tibeto-Burman languages such as the Qiangic languages. The mutual intelligibility between the various groups of dialects is really limited. There is no traditional term to designate this linguistic area, which is considered to be a part of the cultural Amdo region.

In the E section, whether in northern Sichuan or southern Gansu, Mandarin Chinese is currently the official language and is the language used in the school curriculum. It has a strong impact on some languages of the E section, particularly on Khöpokhok (Jiuzhaigou) and Baima. In addition, because of the distribution of the languages in the E section, communications with Amdo pastoralists are usually done by using Amdo Tibetan, hence many Tibetans who are natives of a language of the E section can also speak Amdo to some extent, or even fluently, without any strong accent.

The languages of the E section are usually not written down, and when people write in Tibetan, they normally write in Literary Tibetan. The Tibetan written language is used in the Buddhist and Bönpo monasteries, in the institutes of Traditional medicine and to a certain extent in some cultural organizations and media (internet, etc.).
universities located not far from the area of the E section have Tibetan language departments: the Northwest University for Nationalities (Nubjang Mirik Lobdra Chenmo located in Lanzhou, Gansu), the Gansu Normal University for Nationalities (Kansuu Mirik Geö Lobdra Chenmo Tsö (Hezuo, Gansu) and the Southwest University for Nationalities (Lhonup Mirik Lobdra Chenmo in Chengdu, Sichuan).

In the E section, Vajrayāna Buddhism is the main religion, with a predominance of the Gelugpa sect. There is also a strong Bönpo presence in this region, particularly in Thewo and Zungchu counties. In Dzorge as well as Zungchu one also finds Sakya monasteries. The Buddhist and Bön communities of this section are in contact with Hui communities of southern Gansu and Northern Sichuan.

Major monasteries of the E section include: Tagtshang Lhamo, Čone Gönchen Shädrubling (Geluk) and Sharkhok Gamäl (Bön) in Zungchu County.


### 9.4.1. Migration patterns, legends and historical records

The speakers of each language in the E section have their own history of migration. It is difficult to prove whether this history is based on facts or folklore, but the common

12. This monastery is located in Dzorge at the border with Thewo, Čone and Luchu. It also corresponds to the border between the Eastern section and the Northeastern section where Amdo is spoken. Thus the monastery of Tagtshang Lhamo and nearby monasteries concentrate monks who speak very different dialects.
feature to all the language groups is that the speakers claim they originally come from different parts of TAR, and in the period of the Tibetan Empire (from the seventh to the ninth century) their ancestors were forced to migrate to the eastern region as a military force against Tang dynasty. According to each tradition, Čone ancestors are from Phänpo, Thewo ancestors from Dagpo (DWAGS.P0), Drugchu ancestors from Dagpo, Khöpokhok ancestors from Kongpo, and Baima ancestors from Dagpo. In the case of Sharkhok, Pashi, Thromjekhok and Zhongu, the ancestral origin is unknown and there is no specific folklore about their origin.

9.4.2. Linguistic groups of the E section

The E section is characterized by great diversity, and as mentioned earlier, a very limited mutual intelligibility. Furthermore, this low intelligibility is usually asymmetrical.

In previous works conducted in China, the languages and dialects of this section are often classified as a member of Kham or agricultural Amdo (e.g. Qu & Jin 1981; Nishida & Sun 1990; Zhang 1996; Wang 2012). Sun (2003a) described one Tibetic language of the E section, namely Zhongu. Suzuki (2008) proposes to establish a new category called "Shar" for several Tibetic languages of this section. Sun (2013; 2018) has presented a brief description of two undescribed varieties spoken in Throchu County: Khalong and Taku. The genetic affiliation of Baima, a Tibetic language of this section, is still disputed by several scholars. Until 1949, the Baima ethnic group was still considered as a Qiang minority (Chirkova 2008). It seems that Baima has preserved a Qiangic substrate. When considering the languages of this section, we should also note the existence of linguistic substrata, that is previous potential ethnic minorities’ languages. In this regard, some of the present languages in E Section cannot simply be considered as languages “derived from Old Tibetan.”

For the dialect classification of the E section, we propose the following eleven groups:

- Čone ནོ་ནེའི་སྐད་
- Thewo-tö ཐེ་བོའི་སྟོད་སྐད་
Nearly all these dialect groups or "languages" form a quasi-geographic continuum, but two or more languages of this section are never found in the same village. Within the E section, Baima and Drugchu are clearly distinct languages, not allowing intelligibility with surrounding languages. Other languages, although more closely related, have sharp dialectal differences. This is the case for example of Čone (Čone Nyinpa vs other dialects) or Pashi (Pashi vs Babzo). Čone (except for Nyinpa) does not allow a good intelligibility with Thewo. The mutual intelligibility of Thewo-tö and Khöpokhok is also limited. Pashi speakers do understand Thewo-tö as well as Sharkhok to some extent, but the reverse is not true (a case of asymmetrical intelligibility). Similarly, Khöpokhok speakers understand Sharkhok well while the reverse is not true. Sharkhok speakers understand Thromjekhok to some extent, while speakers of the latter have a much better understanding of the former. The mutual intelligibility of Sharkhok and Zhongu is limited. As mentioned earlier, Baima and Zhongu have a non Tibetic substratum, presumably Qiangic, thus they are typologically quite different from the other languages of the E section. Some dialects of the E section are nearly extinct. That is for example the case of Dramtsher /Batse/ (Chin: Lintan), Meri (Chin: Minxian) and Thergyü (Chin: Dangchang, locally pronounced /Tangchang/).

13. Except for the conference papers of Sun (2013; 2018), no information is available. In addition, these varieties are to a greater extent different from each other. Here we just propose a temporary classification based on the geography.
9.4.3. Geographic extent of the E section

The E section is located at the easternmost region of the Tibetan cultural area at the border between Sichuan and Gansu. It corresponds to the area of the upper Zungchu river འབྲུག་ཆུ་ (Druchu) called Min Jiang (岷江) in Chinese. Other significant rivers of this SE area include the Luchu river དཀར་ཆུ་ (Karchu) and one of its tributaries, the Throchu river དྲོ་ཆུ་ (Throchu), as well as the Drugchu river འབྲུག་ཆུ་ (Drugchu) in Chinese. Other rivers include the Luchu river དཀར་ཆུ་ (Karchu), the Drugchu river འབྲུག་ཆུ་ (Drugchu) also called ‘Thewo Chunak’ གཞེན་ཆུན་ (Chin: 黑水河 Heishui He) which is running through the eponym county and is the main tributary of the Zungchu (Min Jiang).

The E area roughly covers the territories of Cone, Thewo and Drugchu Counties (Gannan TAP) in Gansu as well as a part of Dzorge, Zitsadegu (Jiuzhaigou) and Zungchu Counties (rNgawa TQAP) in Sichuan. Other than these areas, a part of Pingwu and Wen Counties as well as a part of Wudu District are also included in the E area.
PART 2 – CHAP 9. Inner classification of the Tibetic languages

MAP IX.3. – Linguistic area of E section

Legend: ■: Čone; ■: Thewo-tö; ●: Thewo-mä; ○: Drugchu; ▲: Baima; ▵: Khöpokhok; ▲: Pälkyi; □: Sharkhok; △: Thromjekhok; ●: Zhongu; ○: Throchu

Sources: Esri, ArcGIS, USGS, NGA, NASA, CGIAR, N. Robinson, NCEAS, NLS, OS, NMA, Geodatastyrelsen, Rijkswaterstaat, GSA, Geoland, FEMA,

1:2,845,624

0 20 40 60 80 mi

0 15 30 45 60 75 90 km
Detailed location of the dialect groups

- **Čone** spoken in Čone County, in Liulin, Muer, Nalang, Taojian, Tsangpawa, Kache, Zhagulu, Malu, Daogao townships and a part of Niba and Wanmao townships.

- **Thewo-tö** spoken mainly in the westernmost part of Thewo County in Dianga Town and Yiwa Township as well as the northernmost part of Dzorge (north) in Jiangzhai, Zhangwa, Chonger, Donglie and Reer townships. A part of the dialects spoken in Niba Township, Čone County, also belongs to this group.

- **Thewo-mä** spoken mainly in the central and eastern parts of Thewo County in Kaba, Ni’ao, Wangzang, Huayuan, Luoda, Axi, Lazikou, and Sangba townships as well as the western part of Drugchu County in Quwa, Bazang, Hanban, Dayu, and Fengli townships.

- **Drugchu** spoken mainly in a part of Drugchu County in Jiangpan, Nanyu, Baleng, Guoye, Wuping, Chagang, Gongba, Danian, Tieba, and Boyu townships as well as Wudu District (Longnan city, outside of the Gannan TAP), in Pingya township, and Dangchang County in Xinchengzi and Guan’e villages.

- **Pälkyi**/Pashi/ spoken mainly in Pälkyi District of Dzorge County in Baxi, Axirong, Qiuji (Chos-rjes), Baozuo townships as well as in the northern part of Zitsadegu County in Dalu, Yuwa, Heihe townships. In addition to these, it is also spoken in Dala and Duaer townships of Thewo County.

- **Khöpokhok** spoken mainly in the central part of Zitsadegu County (formerly Namphel) in Zhangza Town and Baihe township.

- **Sharkhok** spoken mainly in Zungchu County in Shanba and Shuijing townships as well as Chuanzhusi town, Huanglong, Shili, Dazhai, Qingrun, Anhong and Daxing townships.
Inner classification of the Tibetic languages

- **Thromjekhok** རྟོལ་ཐེས་ཧོ་, spoken mainly in Thromjekhok (Munigou), in Anhong and Muni townships of Zungchu County བོད་དམག་རྫོང་, in Hongtu and along the Zhongu river ཟུང་ཆུ་ཁོང་, in Hongza and Xiaoxing townships and in a few villages of Throchu ཁྲོ་ཆུ་རྫོང་ [Heishui] according to Sun (2003a).

- **Zhongu** ཞོང་ངུའི་སྐད།, spoken mainly in Zhongukhok (Rewugou) ཟུང་ཆུ་རྫོང་ of Zungchu County ཟུང་ཆུ་, in Hongtu along the Zhongu river ཟུང་ངུ་ཆུ་, in Hongzha and Xiaoxing townships and in a few villages of Throchu ཁྲོ་ཆུ་[Heishui] according to Sun (2003a).

- **Throchu** ཁྲོ་ཆུའི་སྐད།, spoken mainly in several villages of Throchu ཁྲོ་ཆུ་[Heishui] in Shashiduo, Luhua, and Qinglang townships according to Sun (2013; 2018). This includes Khalong and Dagu which have been described by Sun (ibid.). Note that these dialects are very different from each other.

- **Baima** བོད་དམག་སྐད། 14† spoken mainly in the eastern part of Zitsadegu County ཞེས་རྒྱ་སྡེ་དགུ་ [Jiuzhaigou] (Guoyuan, Wujiao, Majia, Anle) as well as a part of Pingwu (in Baima, Muzuo and Mupi townships and possibly Huya Township), Wen (Tielou, Liping, Zhongcai, Shangtanbo) and Drugchu counties (in a few villages of Boyu).

Each of the above dialect groups or “languages” have an extremely low intelligibility with the others and in some cases, do not allow for any communication. Particularly Drugchu and Baima are never understood by speakers of other languages. Some speakers of these languages learn Amdo, while Amdo-speaking people generally do not understand nor learn these languages.

There are some varieties spoken between Thewo-mā and Pālkyi /Pashi/, which are not intelligible with the speakers of both languages.

### 9.4.4. Number of speakers

The total number of speakers of the Tibetic languages located in the E section is hard to know due to the lack of reliable census and to the great linguistic diversity. Qu (1996) mentions 65,000 speakers for Çone and Drugchu.†5 According to Bradley

---

14. There is no definite spelling. Some authors propose the following spelling བོད་དམག་སྐད། BOD.DMAG.SKAD.

15. Zhouqu xianzhi (1996) mentions 36,000 speakers for Drugchu, Baima and Thewo speakers in Drugchu County.
(2007), the number of Baima speakers amounts to 10,000. The number of Zhongu speakers of Zhongu is probably less than 5,000 (Sun 2003a). No figures are available for Sharkhok and Thromjekhok.

The total number of speakers for the various languages of the E section does not exceed two hundred thousand speakers. This figure might even be overestimated since some languages of the E section, such as Khöpokhok and Baima, are now endangered and no longer transmitted to the younger generation: Chinese has become the dominant language.

9.4.5. Ethnic and sociolinguistic groups

The speakers of the Tibetic languages in the E section generally identify themselves as Tibetan and specifically as Amdowa except for the Baima-speaking people, who often regard themselves as non-Tibetan (particularly the Baima people living outside rNgawa prefecture) despite the fact that they are officially classified as Tibetan nationality (Tib: BOD.RIGS; Chin: Zangzu) by the Chinese government. The speakers of Čone, Thewo and Drugchu often mention their historical origin (from Central Tibet) but their ethnic identity is likely to be melded with Amdo.

The languages of the E section are spoken by རྣ་པ་ zhingpa 'cultivators', sometimes locally called སྲག་པ།/lakanyi:/ or རུང་པ་ lu:wa/. There are no pastoralist communities, but agropastoralists རོང་མ་འབྲོག་ rongmadrok are encountered.

9.4.6. Phonological characteristics of the E section

It is not possible to list common phonological features to all the dialects of the E section. The phonological characteristics are usually valid only at the level of the groups or even sometimes the dialects.

Suprasegmental features

The Tibetic languages in the E section present various types of suprasegmentals, mainly three:

- Zhongu and Throchu lack any systematical suprasegmental distinctive features. But Sun (2003a) describes for Zhongu a stress contrast in some examples, which does have a phonological function.
PART 2 – CHAP 9. INNER CLASSIFICATION OF THE TIBETIC LANGUAGES

- Pitch tone type: Čone and Baima.
- Register type (or non-suprasegmental contrast but with a phonetically clear reflex): Thewo-tö, Thewo-mä, Drugchu, Pashi, Khöpokhok, Sharkhok and Thromjekhok.

As for the pitch tone, the number of pitch pattern in Čone is two (high and low) whereas that in Baima is at least four (high, low, rising, falling). As for the register tone, Drugchu and Pashi have breathy-based suprasegmentals (low-register may be more marked) whereas the others have creaky-based ones (high-register may be more marked).

**Segmental features**

**Synchronic approach**

The sound systems of the E section are characterized by the following frequent features:

- Existence of voiced non resonant sounds (b, d, ɖ, g, dz, j, z, zh, ɣ, fi).
- Prenasalization is pervasive (“d, “b, “g, etc.). Prenasal with aspirated consonants such as “t, “p, “k, etc., are also found.
- Most dialects have an aspirated fricative series: s’, x’, sh’.
- Preaspiration is found in many dialects of the SE section (t’, h’p, h’k, h’d, h’b, h’g, h’s, etc.). The voicing of the preaspiration sound is sometimes different from the main initial (in Pashi, Khöpokhok and Drugchu)
- The rhyme form is quite simple. Thewo, Drugchu, Zhongu and Baima just have an open-syllable. In these dialects, the nasal feature in the rhyme is not attested.
- A large set of vowels and the existence of /ə/. The opposition between /a/ and /ɑ/. The distinction between long and short vowels is quite frequent.

**Diachronic approach and reflexes of Classical Tibetan**

- In the languages of the E section, the reflexes of preradical sounds are realized either as segmental features, preaspiration and prenasals (cf. Sharkhok) or more
commonly as preaspiration and prenasals (Drugchu, Thewo-tö).

- Voiced non-resonant sounds (b, d, ɖ, g, dz, j) are derived from the consonants with preradicals (except M and') in Sharkhok, Thromjekhok, Thewo-tö, Čone and Baima. In Pashi, b, d, ɖ, g, dz, j) are derived from the consonants without preradicals.

- Prenasalization corresponds to the reflex of radicals with a predical M or '.

- As for aspirated fricative series, CT S, SH, PHY without any preradicals correspond to s’, x’ and sh’ but the sound correspondence in the articulatory position is different depending on each dialectal group.

- Preaspiration is normally triggered by all the CT preradicals (except M and’).

- CT final consonants B, D, often changed into a glottal stop; G changes into the epiglottal V or a glottal stop; M, N, NG changed into nasalized vowels; R, L, S often caused the lengthening of the vowel (except for Thewo-mä, Drugchu, Zhongu and Baima).

- The final consonant triggered a change in the tongue position resulting in a vowel change. In addition, i without final is usually realized as /ə/.

- In Sharkhok and Thromjekhok or Thewo-mä, the combination SL and ZL respectively yields: /ts/ and /dz/ just as in many Kham dialects. This is the case in words like སླ་མོ SLAMO 'easy', སློབ SLOB 'to teach', རླ་བ ZLABA 'moon'.

9.4.7. Grammatical characteristics of the E section

Languages of the E section such as Baima or Zhongu have a number of grammatical features not found in the other Tibetic languages (For example Baima has an accusative case). These specificities may be partly due to the existence of a Qiangic substrate. The languages of the E section are quite diverse and the differences also extend to grammar. Concerning verbal inflections, the languages of this section have inherited many forms found in CT, but have also innovated a number of forms not attested in CT.
9.4.7.1. Case markers

Some dialects of the E section distinguish up to five or six cases (e.g. Sharkhok). Frequent cases include ergative, absolutive, dative, ablative, genitive and comparative.

The ergative is used with both controllable and non-controllable verbs. All languages of the SE section have ergative constructions (Čone, Thewo, Sharkhok, Drugchu, Khöpokhok, Pashi, etc.) except Baima which has an accusative marker (Huang 2007: 159) and lacks a real ergative marker. Chirkova (2005) however mentions for Baima the existence of an agentive case, but specifies that it is used mainly to disambiguate potential agents.

The ergative in Sharkhok can be used in the past and present but not in the future. In some dialects, the pronouns have a special form. It is also worth noting that the ergative marking is more obligatory for pronouns but remains optional for nouns. Thus, one could say that the ergativity is weakened in the E section compared to other sections.

In the various languages, the ergative marker is often marked by /ɣə/ derived from གིས་GIS.

In the SE section, ergative, genitive markers and even dative are distinctive but have a very similar form.

The absolutive is always marked by a zero marker (∅). It is used to mark the patient except for Baima which has dominant accusative constructions.

The accusative case is found only in Baima. The form of the case is /tæ/ of unclear origin.

The dative marks the beneficiary and sometimes the patient. In Sharkhok, the form is /ɣe/ and in Čone /ɦe/. The form /-tsə/ or /-sə/ is used in Zhongu. These origins are unclear. The dative is also often used to mark the locative; In Čone, pronouns have a special form for the dative case.

The local or peripheral case markers are ablative, genitive, comparative, locative (inessive, illative).

The ablative is often གནས NAS /-ncə/ or /-nə/. In Baima the ablative form is /ɣə/ of unclear origin.
The comparative has several non-related forms in the various languages: /-ʃũmbo/ (Sharkhok), /-s’ɔ/ (Khöpokhok), /-ʃuŋz/ (Baima). These forms could be related to the comparative form /sang/ used in Ladakhi.

For the locative function, the dative is often used. Some dialects have the form /-nɔ/ or /-na/ (e.g. Zhongu, Drugchu) derived from CT locative case: š’ NA.

The instrumental in many E dialects is formally identical to the ergative case GIS /-ɡə̲, -kə̲/. Baima, which lacks an ergative construction, has preserved the instrumental form: /-kə̲/.

The genitive form in Sharkhok and Khöpokhok is /-ɡə̲/ derived from GI. In many dialects, pronouns have a special form for the genitive case. According to Huang (2007) Baima entirely lacks a genitive case, except for pronouns where the genitive is formally identical to the accusative.

9.4.7.2. Nominalizers

Various nominalizers are found in the E section (for a general description of the nominalizers, see section 8.3.13). We list below three frequent nominalizers of this section. Data about nominalizers in these languages are not sufficient.

- In some dialects, one finds the nominalizer /-mo, -nyi, -m(n)ya, -m(ə)w/ (Čone, Khöpokhok, Drugchu, etc.). These forms are probably also derived respectively from MI and MYI which originally means ‘person, human being’. The nominalizer and ‘human being’ have a similar form in many dialects. It may indicate the general nominalisation (infinitive-like), the patient or the agent (of the verbal action) or the subject of an intransitive verb (the ‘doer’, the ‘the thing to be done’ or ‘the thing done’). It refers to persons and is not used for inanimate beings.

- Another frequent nominalizer is SA /-s’a/. It is pervasive in many dialects of the E section. This nominalizer is derived from CT SA, which means ‘place’ has various meaning ‘the place of the verbal action’ and ‘the instrument of the verbal action’.

- The nominalizer /fiə/ probably related to PA is found in Drugchu.
9.4.7.3. Verbal inflections

In the dialects of the E section (Thewo, Čone, Khöpokhok, etc.), the lexical verb often has two forms. Sharkhok and Pashi have even three forms for some verbs (“past,” “present” and “imperative”). These forms are either inherited from CT or correspond to specific innovations.

For example, some dialects have specific aspirated forms for the imperative, which are not inherited from CT. In some dialects, the verbs ‘to go’ and ‘to do’ often have suppletive forms to indicate the various tenses and the imperative: ‘to go’ མོང་SONG, སྐལ་THAL and ཤོོ་GRO (e.g. Sharkhok); ‘to do’: མོོ་BYED and ཤོོ་BGYID (Čone).

9.4.7.4. Linking verbs and auxiliary verbs

Linking verbs correspond to equative and existential verbs.

Copulative verbs

Copulative verbs རེད་RED and རེད་MIN are found nearly everywhere in the E section, however, in a few dialectal groups other pairs are used as: རེད་MIN and སྐུ་GI /gi/ (Thewo-mä, Drugchu). They convey respectively egophoric and factual meanings. These auxiliary verbs have the following negations: རེད་MIN > རེད་MIN, རེད་RED > རེད་MA-RED (most dialects), སྐུ་GI > སྐུ་GI /magi/ MAGI (Thewo, Drugchu).

Existential verbs

Existential verbs such as སྣང་SNANG are quite common to the Tibetic languages in the E section. Usually SNANG conveys an egophoric meaning while SNANG has a sensory meaning. A compound form སྣང་GI is found in Čone and Sharkhok.

སྣང་DUG is quite marginal but found in a couple of dialects such as Zhongu.

These auxiliary verbs have the following negations:

སྣང་SNANG > སྣང་MI-SNANG (nearly all the dialects), རེད་YOD > རེད་MED (most dialects), རེད་MYED (Thewo-mä).

Compound linking verbs

Compound verbs are very frequent in the E section. They involve the combination of various copulative verbs, existential verbs or auxiliaries. For example, we find རེད་པར་
Frequent auxiliary verbs of the E section consist of copulative verbs and existential verbs (sometimes preceded by a relator): རེད་ RED, སྣང་ SNANG, དྱོད་ YOD.

Other frequent auxiliaries include: ཕུལ་ THAL (all dialects), སྦྱུང་ BYUNG (Baima), དགོས་ DGOS (Sharkhok, Pashi). In some languages such as Sharkhok, the verb often occurs at the end of a sentence without any auxiliary.

From a semantico-cognitive point of view, the auxiliaries (together with relators) convey complex temporal, aspectual, evidential and epistemic meanings.

Evidential systems of the E section usually make at least a distinction between egophoric, sensory, factual and hearsay markers.

9.4.7.5. Negation

The negation has two forms in all the dialects of the E section: མ་ MA and མི་ MI. Generally speaking, the negation in the various tenses and aspects often comes before the auxiliary and in some cases in front of the lexical verb. See above the specific negation of the copulas and auxiliaries འགོད་ YIN and འགོད་ YOD.

9.5. The North-eastern section

The north-eastern section (henceforth NE section) is made up of one large linguistic set of dialects traditionally called Amdo. Thus, we can say that the NE section is made up of one “single Amdo language” with a significant dialect variation. However, there is not any real standardization of Amdo even if some dialects such as Rebgong, Labrang or Thrika often play a prominent role in the media.

Some intellectuals want to establish an Amdo standard language. See for example DPAI. LDAN BKRA SHIS recent publication Amdo Tibetan language, an introduction to normative oral Amdo (2016).

The dialectal differences are particularly salient in the field of phonology but to a lesser extent they also involve some aspects of the grammar and the lexicon. In Amdo, the phonological discrepancies between the dialects are not only geographically based
but they also depend upon sociolinguistic parameters, namely the way of life and activity of speakers: nomadic and pastoral versus sedentary and agricultural.

Mandarin Chinese is the official language and is used in the school curriculum of the NE section, but Literary Tibetan is still taught in some schools of the Amdo area. The literary production of Amdo is currently the most productive of all the Tibetic regions.

Amdo dialects can easily be transcribed in Tibetan script but they are usually not written down, and when people write in Tibetan, they normally write in Literary Tibetan. However, some novels are clearly influenced by vernacular Amdo. As in the other Tibetic areas of China, Literary Tibetan is also used in the Buddhist and Bönpo monasteries, in the institutes of Traditional medicine and to a certain extent in some cultural organizations and media (particularly on the internet). Depending on their home province (Qinghai, Gansu or Sichuan), Amdo students may attend one of the three universities, which have Tibetan language departments: the Northwest University for Nationalities 甘洛藏文语言文学系 Nubjang Mirik Lobdra Chenmo located in Lanzhou (Gansu), Qinghai University for Nationalities 青藏语言文学系 Tshongön Mirik Lobdra Chenmo in Xining (Qinghai) and the Southwest University for Nationalities 甘洛藏文语言文学系 Lhonup Mirik Lobdra Chenmo in Chengdu (Sichuan).

In the NE section, Vajrayāna Buddhism is the main religion. The followers belong predominantly to the Gelugpa sect, but Nyingmapa, Kāyupa, and to a lesser extent Sakyapa sects are also represented. Bönpo communities are also found in the region, particularly in Rebgong, Thrrika and Çāntsha. Jonangpa communities are found in Golok and Ngawa areas.

Among the major monastic institutions of the NE section, we find: Kumbum, Labrang Trashi Khyil, Rebgong Rongwo Gön, Dragkar Tredzong in Tsigorthang (Geluk), Achung Namdzong (Nyingma), Darthang (Nyingma) and Jonang in Gabde County.
Amdo speaking Muslim minorities are attested in a few counties namely, in Xunhua and Hualong. More generally, there is a strong “Hui” or Chinese Muslim community in Qinghai and Gansu. There are also Amdo-speaking Mongols living in Sogwo County.


9.5.1. Migration patterns, legends and historical records

Some parts of the Amdo region were incorporated into the Tibetan Empire in the second half of the seventh century. Before this date, the Amdo region was inhabited by other ethnic groups such as ’Azha (probably Turkic speaking people), Qiang and Tangut people. During the second half of the seventh century, a significant part of Amdo was under the administration of the mGar family, who was mandated by the Tibetan emperor.

In the more recent periods, Amdo speakers have migrated both within Amdo and outside the traditional Province. During the Muslim warlord Ma Bufang’s rule, people form the Amdo area of Kha-sgang migrated from Palung /Hwalung/ to Chabcha, Thrika, Mangra and Ba Districts in order to flee the forced Islamization conducted by the Qinghai warlord (see SUM.BHA.DON.GRUB.TSHE.RING 2011: 21). Amdo speakers have also migrated to various Kham areas such as Dartsendo, Lithang, Drango and

16. It is noteworthy citing a toponym Axia, a township located at the south of Thewo County. According to the folklore, the inhabitants of this township are descendants of ’Azha. However, their language exhibits the characteristics as a Tibetan language, belonging to Thewo-mi; cf. 9.4.3. More investigation is needed.
Nyagrong (see the classification below) from the present south-Kokonor and Rebgong regions.

9.5.2. Linguistic groups of the NE section

Amdo language is in contact with a few Non-Tibetic languages belonging to three genetic stocks: Mongolic, Sinitic (Chinese), and Turkic. Some dialects of Amdo are also in contact in the south with Tibeto-Burman languages such as rGyalrongic and Qiangic languages.

According to Janhunen (2005: 114), the Mongolic languages spoken in Amdo are

“Shira Yughur, (Huzhu) Monghul, Mongghuor, (Minhe) Mangghuer, Qinghai (or Buddhist) Bonan, Gansu (or Moslem) Bonan, Kangja, and Santa (or Dongxiang). [...] Several Mongol and Oirat groups notably the so-called Henan Mongols, have been linguistically assimilated and today use Amdo Tibetan as their native language.”

Janhunen adds:

“The Turkic family is represented in the Amdo by two distinct languages: Salar and Sarygh Yugur (or Western Yellow Uighur). In addition, Kazakh is spoken marginally in the region (in Western Qinghai close to the border of Xinjiang).”

The two languages Salar and Sarygh Yugur (also spelled Yughur) do not present mutual intelligibility. The first language, Salar is spoken essentially in Xunhua County (Qinghai) and the second, Sarygh Yugur is spoken in Sunan County (Gansu), located roughly 400 km away from Xunhua. Northeast of the Kokonor lake, there is also an Autonomous Hui County in Menyuan Huizu Zizhixian, (Traditionally, the Tibetic area within Menyuan belongs to Pari /Hwari/).

This section is made of six main groups of dialects: Tsho Ngönpo (Kokonor), Labrang-Rebgong, Tsongkha, Rwanak, Ngawa and Washül and several smaller groups. All the dialects of this section generally allow a rather high mutual intelligibility, except
for some striking exceptions, such as Khalong, gSerpa, Gorkä, Dungnak and rTarmnyik. The first two dialects clearly have a Qiangic or rGyalrongic substratum. 17

The various dialects are grouped together into a language which is named Amdo, locally called Amdö-kä, འམདོའི་སྐད་/amdo / kät/, /amdo käl/ or /amdo kā/ often shortened as Am-kä, འམ་མདོའི་སྐད་/amkā/ in the northern pastoralists’ area, /amkā/ in Ngawa and /amkā/ in the cultivators’ region. Amdo language is also often simply called བཤད་སྐད་/workā/ or /workā(t)/ i.e. literally ‘Tibetan language’ according to the local pronunciation of བཤད་སྐད་/BOD.SKAD. When they refer to the language of Central Tibetan, the term བཤད་སྐད་/BOD.SKAD is pronounced as /po kāt/. The same is true for the designation of the literary language. བཤད་ཡིག /BOD.YIG /wo(t)yik/ usually refers to the Tibetan script while the Classical Literary language is called /poyik/.

We have already mentioned in 3.3.1 that there is a widespread cliché stating that all the pastoralists’ dialects are quite similar, while the cultivators are very distinct. 18 This point of view is not correct first because Kham, Amdo and Tö Ngari pastoralists’ dialects are very distinct but also because within Amdo, we find quite distinct pastoralist dialects, some of which are innovative while others are more conservative.

Recently, CHAM.TSHANG PADMA LHUN.GRUB (2008) proposed a classification into four groups: a) archaic pastoralists’ dialects འབྲོག་སྐད་རྙིང་མ་/BROG.SKAD RNYING.MA, b) innovative pastoralists’ dialects འབྲོག་སྐད་འཕེལ་མ་/BROG.SKAD PHEL.MA, c) cultivators’ dialects རོང་སྐད་/RONG.SKAD, d) agropastoralists’ dialects རོང་མ་འབྲོག་གི་སྐད་/RONG.MA.BROG.GISKAD.

17. According to this author, the specificity of some lexical items may be explained by the existence of a Showu subtrate, which is a rGyalrongic language. Khalong speakers also speak Showu rGyalrong. Showu rGyalrong is also called Zbu.

18. This idea is for example reflected in the following Wikipedia definition of the Tibetan dialects:

"[...] The Tibetan language is used in Tibet. Generally, one makes a distinction between pastoralists’ and cultivators’ dialects. Concerning the pastoralists’ dialects, there is no distinction between high, low and middle areas [Jangthang, Amdo and Kham areas]. The pronunciation and intonation are harsh and rough and the initial sounds have not disappeared. [...]"
We generally agree with his analysis, but it is not sufficiently detailed and we prefer to use a classification based on geolinguistic and historical categories rather than purely sociolinguistic categories.

The problem with CHAM, TSHANG PADMA L'HUN, GRUB's classification is that it is not always easy to distinguish between "archaic" and "innovative" dialects as a whole. In a single dialect, some features may be archaic while other features may be innovative. But there are other sociolinguistic reasons that prevented us from using this kind of analysis. The notion of "pastoralist" may refer to people who practice the activity of cattle-breeding but also to a sociolinguistic identity. Nowadays, some pastoralists or drogpa locally called /dʰoχwa/ may still practice transhumance, while other drogpa have settled down but still breed cattle. A third category of drogpa do no longer possess yaks, sheep, goats, and horses and may work as merchants, lamas, civil servants, etc., but still consider themselves as drogpa.

Two dialects Dungnak and rTarmnyik were recently discovered in Gansu by Shao Mingyuan in 2012. These two dialects need further research. Dongnak speakers claim that they originally came from Chamdo. So it may turn out that these two dialects are Kham enclaves in Amdo (just as there are Amdo enclaves in Kham; see the following description).

For the dialect classification of the NE section, we propose the following fourteen groups:

- **Tsho Ngönpo (or Kokonor) group** མཚོ་སྔོན་པོའི་ཡུལ་སྐད་ཚོགས་པ་
  This group corresponds to the dialects spoken by pastoralists who live around the lake Tsho Ngönpo. They speak pastoralist innovative dialects’ མཚོ་སྔོན་པོའི་ཡུལ་སྐད་ཚོགས་པ་. In some

---

19. rTarmnyik was mentioned by Qu Aitang (1996) but it has only recently been studied by Shao Mingyuan (pers. comm.).

20. They use the existential verbs SNANG / MI SNANG and the reflex of the initial labial $b$ pronounce /p/ and not /w/ as it is usually pronounced in Amdo. The first feature, however, does not correspond to the present Chamdo dialect belonging to Northern Route (Zälmogang) Kham. (see 9.3, in the SE section)
Chinese literature, this group is called 环湖方言  *Huanhu fangyan* ‘circum-lake dialect.’

- **Tsongkha group** རྩོང་ཁ་ཡུལ་སྐད་ཚོགས་པ།
  
  This group corresponds to cultivators’ dialects རང་ཁ.

- **Labrang-Rebgong group** ལླ་བྲང་རྩོང་ཁོང་ཡུལ་སྐད་ཚོགས་པ།
  
  This group corresponds to agropastoralist dialects རང་མ་རྩོང་ཁོང་གི་སྐད་ཚོགས་པ།.

- **Rwanak (Banak) pastoralist group** སྦྲ་ནག་ཡུལ་སྐད་ཚོགས་པ།
  
  This group locally called Rwanak is also known in the literature as Banak, lit. 'Black tents.' It is located in Amnye Machen ཨ་མྱེས་རྨ་ཆེན་ས་ཁུལ་ and corresponds to pastoralist archaic dialects འབྲོག་སྐད་རྙིང་མ་.

- **Ngawa group** རྨ་བ་ཡུལ་སྐད་ཚོགས་པ།
  
  This group corresponds to archaic dialects mainly spoken by sedentary pastoralists འབྲོག་སྐད་རྙིང་མ་.

As well as the following smaller groups:

- **Arik group** ཅི་རིག་ཡུལ་སྐད་ཚོགས་པ།
  
  This group corresponds to pastoralists’ archaic dialects འབྲོག་སྐད་རྙིང་མ་ spoken in Dola Ringmo area མདོ་ལ་རིང་མོའི་ས་ཁུལ་.

- **Hwari (Pari) group** དཔའ་རིས་ཡུལ་སྐད་ཚོགས་པ།
  
  This group corresponds to pastoralists’ archaic dialects འབྲོག་སྐད་རྙིང་མ་.

*Southern groups of dialects with settlements in Kham*

- **Mewa pastoralists’ group** (with settlements in Kham) རྨེ་བའི་ཡུལ་སྐད་ཚོགས་པ།
  
  This group also corresponds to pastoralist archaic dialects འབྲོག་སྐད་རྙིང་མ་. It includes dialects spoken by pastoralists in some areas of the Kham region.

---

21. The tribes’ name does not always allow identification for the younger generation.
• **Washül pastoralists’ group** (with migrations into Kham) དབའ་ཤུལ་ཡུལ་སྐད་ཚོགས་པ
This group also corresponds to pastoralist archaic dialects ཀྲོག་སྐད་རྙིང་མ་. It includes dialects spoken by pastoralists in some areas of the Kham region and the Kham-Amdo border zone.

**Divergent dialects**

• **Gorkä group** ཆོས་ལྡོན་བོད་སྐད་ཁོ་བར
This group corresponds to agropastoralist and pastoralist dialects of the Golok area རྒྱལ་རོང་དང་ཐག་ཉེ་བའི་ཨ་མདོ་འིཡུལ་སྐད་

• **Gyårongo-spheric Amdo** རྒྱལ་ལྕོང་དང་ཐག་ཉེ་བའི་ཨ་མདོ་འིཡུལ་སྐད་
This group corresponds to archaic dialects spoken by pastoralists ཀྲོག་སྐད་རྙིང་མ་.

As well as the following small group:

• **Dungnak and rTamnyik** dialects neighboring the Turkic-speaking Western Yughurs of Gansu: བྲལ་ལྡན་བོད་སྐད་ཁོ་བར.

One of the striking features of the North-Eastern section is the generally high intelligibility across different groups. Our classification differs from previous classifications that were mainly based on the distinctions between various sociolinguistic groups: cultivators’ group, agropastoralists’ group and pastoralists’ group. This is true for recent classifications in Tibet (see **SUMBHA DON GRUB TSHERING** 2011). Although these sociolinguistic distinctions are quite significant, they prove to be insufficient for the linguistic classification of Amdo. As it is the case in other sections, Amdo dialects have some significant differences in the fields of phonology, grammar, and lexicon. At the two ends of the geolinguistic continuum of Amdo, the differences may be quite significant, but they would not be an obstacle to basic conversation.

We have noticed significant differences between the pastoralists’ communities within Amdo, thus, we have introduced traditional communities’ names, often referred to as དོན་ tshona ‘tribes’ and འབྲོག་སྐད་ shogpa e.g., འཛིན་ Arik, རྡལ་ Rwanak, དབའ་ཤུལ་ Washul /ya fi’u/ (or /ya fi’u/). For example, in Rebgong, there are twelve shogpa: མཐུ་ ལུམ་པོ་ལེབ་
9.5.3. Geographic extent of the NE section

The area of the NE section is nearly equivalent to the historical Amdo province except for the part corresponding to the E section mentioned above (where Amdo is often a second language). It corresponds to all of TAPs in Qinghai (except most of Yülshül) and to Gannan TAP and Pari (locally /Hwari/) TAC in Gansu. In Sichuan, it extends to the northwestern part of Ngawa TQAP and to a part of Kandze TAP.

The Amdo linguistic area is located around the Tsho Ngönpo Lake མཚོ་སྔོན་པོ་ (Kokonor or Qinghai lake) as well as the Ma-chu རྒྱ་ཆུ་, which corresponds to the upper course of the Huanghe (Yellow river), and its tributaries, the Sang-chu སང་ཆུ་ (夏河 Xia He) and the Lu-chu ོལ་ཆུ་ (洮河 Tao He). Another significant river of Amdo, although smaller, include the Tsongchu བསང་ཆུ་ (夏河 Xia He) that runs through Ziling རྒྱ་ཆུར་ཉིས་ ' (西 "Xining). In the northwest of Amdo, an important river is the Chunak river འོ་ཆུ་ (黑河 Hei He) that runs through Arik (祁连 Qilian). Other rivers include the Darlak-chu སྦོང་ཆུ་, Tse-chu ནག་ཆུ་, Serch-en-chu ཉེར་ཆེན་ཆུ་, Chungön ཚོང་ཆུ་ (blue water), Ba-chu སྦོང་ཆུ་ and Gu-chu སྦོང་ཆུ་ (also called Ka river ཚོང་ཆུ་).

South of the Machu river, the Amdo speaking area extends to the Golok region, in the upper course of the Nyagchu river ཉག་ཆུ་ as well as to Ngawa region.

Some southern Amdo dialects are spoken in Gyälrong རྒྱལ་རོང་ and Kham ཁམས་ areas in the following counties: Chuchen བྱུུ་ཆུ་ (rNgawa TQAP), Rongdrak རོང་རུ།, Tau ར་, Dartsendo སྦོང་ཆུ་, Dranggo དྲང་གོ, Kandze, Paðyul, Nyagrong ཉག་རོང་, and Lithang སྐུ་ཐང་ (Kandze TAP). In these areas, dialects of Amdo are spoken in the high pastures by pastoralist communities. The Amdo speakers of this region (except Chuchen) do not consider themselves as གུམ་པ་ 'Amdowa' and usually call themselves either གུམ་པ་ Khampas or རྩོ་པ་ Drogpa i.e., 'pastoralists' locally pronounced /mḍøgwa/. Amdo speaking communities are also found in the west of Nyagchu ཉག་ཆུ་ in Honglong and Kela townships (Nyachukha County ཉག་ཆུ་ཁ་).
Detailed location of the dialect groups

The major groups include:

**Tsho Ngönpo group**

This group corresponds to pastoralist innovative dialects that are spoken in the north-west of Amdo around the Kokonor lake in Kangtsha མཚོ་སྔོན་ཡུལ་སྐད་ཚོགས་པ་, Themchen རྐང་ཚ་, Dazhi སྤྲེལ་པོ་, Chabcha (north) སྒོ་མེ་, Mangra སྣང་, and extends to the region north of Amnye Machen in Gápasamdo སན་པུ་སྐད་ཚོགས་པ་/Kawasumdo/ སྐད་ཚོགས་པ་(south), and Tsigorthang གོ་རོ་ཆེན་པོ་(alt. རྗེ་རེ་མ་, BRAG.DKAR.SPREL). Some pastoralist communities in Thrika ཐྱིག་, Cántsha སྤི་ཚ་ and Yadzi ཰དི་(Kangtsha dewa) also use this variety. Additionally the dialects spoken by pastoralist migrants in Tuulän and Bayánkhar are probably also affiliated to this group.

The main *tsowas* and their distributions are: བོང་སྟག་ BONG.STAG (Themchen, Tsigorthang), རྐང་ཚ་ RKANG.TSHA (Kangtsha, Thrika, Xunhua), གཡོན་རུ་ G‧YON.RU (Chabcha, Cántsha), ཰ུས་རྒྱན་ RUS.KANGAN (Mangra, Dulan), མདའ་བཞི་ MD.A.BZHI (Dazhi), རུ་སྔན་ RU.SNGAN (Mangra, Dulan), འབྲག་དཀར་སྤྲེལ་ BRAG.DKAR.SPREL (Mangra), སྒོ་མེ་ བདུད་ཤུལ་ SGOME BDS.UL (Thrika, Chabcha), འབྲག་དཀར་སྤྲེལ་ རྩི་གོར་ཐང་ BRAG.DKAR.SPREL (Thrika, Tsigorthang) and りན་ལུགས་ RIN.LUGS (Mangra).

The area on the west of the lake, including the Tarim basin and the piedmont no longer has any significant Tibetan settlements. This region is presently inhabited mainly by Chinese, Hui, and Mongol populations. This is the case of Terlenkha city གཏེར་ལེན་ཁ་གྲོང་ཁྱེར་ Delingha, Nagorno or Golmud city གཉིམ་ཐུང་གོང་གི་ དུར་ལམ་ GERM.NGAM.DUR.LAM (Chin: 德令哈 Delingha), Nagorno or Golmud city གཉིམ་ཐུང་གོང་གི་ གཉིམ་ཐུང་གོང་གི་ GERM.NGAM.GERM.NGAM (Chin: 德令哈 Germmu), Mangne སྣང་ (Chin: 芒崖 Mangya), Lunhu རུང་ཧུ རུང་ཧུ LUN.HU (Chin: 冷湖 Lenghu), Datsha dam སྣང་གུ སྣང་གུ DACHAI.DAN (Chin: 大柴旦 Dachaidan) and Wulän རུའུ་ལན་ WUL.IAN (Chin: 都兰 Dulan). The only exception is the county སྨུན་ལུགས་ SNUN.LUGS (Chin: 都兰 Dulan), which includes several Tibetan villages, that are mainly cultivators and agropastoralists. 23 The only exception is the county སྨུན་ལུགས་ SNUN.LUGS (Chin: 都兰 Dulan), which includes several Tibetan villages, that are mainly cultivators and agropastoralists.

23. Wulän is a Mongolian Autonomous County, which is mainly inhabited by Mongolian speaking people and Chinese. Recently about 20 Tibetan families have settled in this area.

24. Some scholars believe the real name was གྲོུན་ལུགས་ DUR.LAM (the path of the graves) near the great walls.
• **Tsongkha group**

This group corresponds to cultivators’ dialects spoken mainly in Tsongkha area near Xining and in the cultivated areas near the Yellow river and the Tsongchu River. The core dialect area is the former, such as Čäntsha གཅན་ཚ་(snap: SNANG.RA, mag: MAGI, (strict skyar:engs, 严格格独 bra:gi SNAKHA, 严格格独 SKYAGSKYA), Palung དཔའ་ལུང་ DPA.LUNG /Hwalong/ (Wayänkhar), ཉི pst: BIMDO /Windo/ in Yadzi སྣང་ར་(Xunhua); part of Pari ང་རིང་/Hwari/ (BAIYAN RONG.KAID), Drotshang གྲོ་ཚང་ (Ch: 乐都 Ledu), Tsongkhakhar མཁར་ (Ch: 平安 Pingan) in བར་ཚང་village, Kamalok རྗེ་མ་ལོག་(Ch: 民和 Minhe) in གཤོང་ཐང་village, as well as in Thrika སྣང་ར་(Ch: 互助Tu Autonomous County) and Semnyi སེམས་ཉིད་(Ch: 互助Tu Autonomous County) and Semnyi སེམས་ཉིད་(Ch: 互助Tu Autonomous County). In some areas of Gäpasumdo /kawasumdo/ བད་པ་སུམ་དམོ (also called ར་ Ba), Mangra གྲོན་པ་(Ch: 贵南 Guinan), Tsigorthang གྲོང་ཐང་, and east Chabcha གྲང་ (county seat and its surroundings), immigrants from Hwalung རྩེ་མ་འབྲོག་ (Wayänkhar) and Čäntsha (less than 100 years ago) speak this group’s dialect (see Roche 2015 and Tsering Samdrup & Suzuki 2017); therefore, they should be closely related to the Hwalung dialect. Some speakers from Hwalung have also migrated to Tuulän.

• **Labrang-Rebgong group**

This group corresponds to agropastoralist dialects ཨོོང་མ་འབྲོག་'Rongmadrok-gi ka’. These dialects are spoken mainly in Labrang རྙིང་པོ་, Tsö སོ་ and Rebgong རེབ་ by communities, which practice both cattle breeding and culture. We can, however, find small differences between speech in Rebgong and Labrang. The area extends to some communities located in Čäntsha གཙོད (snap: STENG.SÖ, mag: NANG.KHOG and snap: DO.RGYA), Dobi /Dowi/ སྦིང་མ་and Thrika སྣང་ར་(snap SK.BA and
STONG.CHE), the last of which is considered a slightly different subgroup from other dialects in this group called  JAM.PA’SKAD, considered by locals as a different variety. The last subgroup is  AMCHOG YUL.SKAD (Sangchu, Chin: Xiahe), and  KHAGYA TSHO DRUG, and  Samtsha have very specific features which are maybe related to the Ngawa group.

- **Rwanak group**

This group corresponds to pastoralist archaic dialects spoken in a large area south of Amnye Machen roughly in the north of the Golok prefecture, in Malho prefecture and in the northeastern edge of the Yülshül prefecture, in the following counties: Tsekhok (Tsekhog), Soqwo (Gäpasumdo), Machen (Machen), Matö (Gäpasumdo), Chumarlep (Machen), Thringdu (north) (Rhingdu). Some pastoralist or agropastoralist communities in Mangra, Luchu, Arik, Rebgong, and the western part of Labrang speak a dialect of this group.

The main tshowas and their distributions are: HOR (Tsekhog), BON.BRYGA (Tsekhog, Mangra), 2ARIG (Machen), MGAR.RTSE (Gäpasumdo) and SHA.SBRANG (Gäpasumdo).

- **Ngawa group**

This group also corresponds to sedentary pastoralists’ archaic dialects  principally spoken in Ngakhok. These pastoralists have settled down and have become sedentary cattle breeders. These dialects are spoken in Ngawa (Machu), and a part of Çigdril. This group also includes some dialects located in Amdo in Dzorge and Murge (in Zungchu County), north Dzamthang pastoralists as well as southern pastoralists’ part of Golok Prefecture such as Gabde. Additionally, a few speakers of this group are found in Marthang and Darlak.

---

25. They claim they are of LDONG (Minyak) origins.
26. In Soqwo, the official Mongolian language was still used about 150 years ago (Jangbu pers. comm.). Nowadays there are less than ten percent that can speak Mongolian. See also Balogh (2017). There are principally three local varieties in Soqwo.
• **Arik group**

This group is essentially made of the Arik dialect spoken in Dola County about 100 km north of the lake on the banks of Chunak river. Some branches of the Arik tribe live in the present Sogwo and Machen Counties. However, most of the speakers have adopted the local dialect which belongs to the Rwanak group. The Tibetan dialect spoken in Sunan is probably also affiliated to Arik, but it is not sufficiently documented.

• **Hwari (Pari) group**

Hwari is a pastoralists’ dialect spoken in Hwari (West) in Gansu as well as in Semnyi, Serkhok and probably, by pastoralists in some parts of Gönlung (Dgon.lung, also referred as Hordrong) in Qinghai.

**Southern groups of dialects with settlements in Kham**

• **Washül group**

This Washül group corresponds to the traditional tribes of དབའ་ཤུལ་ (Dba'.shul) or རྒྱ་ཤུལ་ (Rgya.shul) and designates dialects spoken by pastoralist communities in Serta (or Serthar) སེར་རྟ་, Darlak སྲེ་ལག and partly Padma/Panma/ གནའ་མ་ and Sershül སེར་ཤུལ་.

Some of these communities’ dialects are spoken in the traditional Kham region. That is the case of the following dialects:

- Dranggo pastoralist dialect is spoken on the height of Dranggo བྲག་འགོ (Chin: Luhuo) called Likhok བི་ཁོག་.
- Some pastoralists areas of Kandze སྙིན་རྫོང་.
- Nyagrong pastoralist dialect is spoken in Larima ལ་རི་མ་ Township of Nyagrong རྲི་མ་.
- Lithang pastoralist dialect is spoken in central and north-west Lithang གཡོན་རུ་ around the county seat (Gyon.ru) and སྲེ་གཞུང་མ་ Sde.Gzhung.ma tshowa) as well as in Gemu Township of its southern border area to Dabpa County སྐྱབས་ཤུལ་ skyabs.shul tshowa which descends from གཡོན་རུ་.

27. The traditional names are བོད་པའི་བོད་པ་ HOR.BA RNING.PA, བོད་པའི་དཔོན་པོད་པ་ DPAONSKOR, རྫོང་པོད་པ་ GYON.RU, སྲེ་གཞུང་པོས་ SDE.GZHUNG.MATSHOWA, and རྫོང་པོད་པ་ MCHOD.Rten. Many pastoralists in this region have practised a nomadic lifestyle without having any determined domicile.
SDE.GZHUNG.MA), and in the western border of Nyagchukha in Honglong and Kela townships (by the 70.THOG tribe). BDE.GZHUNG.MA’s variety is to some extent different from the others.

Note that Tibetans in Lithang call Amdo-speaking pastoralists WA.SHUL and their language WA.SKAD regardless of their origin. They refer to all pastoralists from Tsho Ngonpo area and are related to Mongols. Several Lithang pastoralists have narratives that their ancestors came from the Tsho Ngonpo area and are related to Mongols. The GYON.RU (locally pronounced as GYON.GRU) is one of the pastoralist groups in Lithang and even now one can find GYON.RU to the south of Tsho Ngonpo. As far as linguistic features are concerned, every variety spoken in Lithang shares many features with the Tsho Ngonpo group, but there are also features close to the dialects spoken in Serta, the original place of the WA.SHUL tshowa. Wherever their ancestors came from, the influence of WA.SHUL.

The pastoralist dialect in Shingnya hamlet within Lhagang Town (Dartsendo) is directly related to Lithang pastoralist dialect according to their migration history (see Suzuki & Sonam Wangmo 2019a).

Some Washül people live in Marthang (Hongyuan) and Dzorge as well. However, their language does not reflect the Washül group’s features, but it is rather similar to rNgawa group.

**Mewa group**

This group corresponds to pastoralist dialects spoken on the Mardzagang plateau which is located in the traditional region of Kham.

It includes:

- Mardzagang pastoralist dialect spoken in Longdeng, Seka and Xiede townships and Bamei town in Ta’u County (RME.BATSHOWA).
- Lhagang pastoralist dialect is spoken in Tagong (LHA.SGANG) township

---

28. The word drogs in Lithang denotes Tibetan pastoralists speaking varieties of Kham. They principally live in the DRA.KHOG area, in the westernmost area of Lithang, at the bottom of the sacred mountain GNAS.CHEN.DGE.BSNYAN, connecting to Bathang County.

29. In the local languages, Mongolians are always called sog (Po’); however, in some proper names, HOR is also understood as ‘Mongolian’, as in HOR.RNying.PA (Hen’ Township), HOR.LUNG (Honglong Township) and HOR.CHU (Huoqu River).
in Dartsendo མཁྲིག (Kangding) County (RME.BA, NAG.LONG.MA, ICANG.PA, NANG.SKOR, RAL.LI, BAR.NANG GSUM.MDO tshosus).

Marthang (Hongyuan) ཆ་ལུང

The above varieties are closely related.

The original center of མཁྲིག RME.BA tribe was located in a part of the present Kandze County. However, most families migrated firstly to Martsagang, and then to Marthang (Hongyuan). This migration dates back to the move of Mewa Monastery from Kandze to Marthang in the nineteenth century. At present, the majority of the original མཁྲིག RME.BA tribe’s territory is inhabited by speakers of the Washiul group (see above).

Divergent dialects

• Gorkä group གོར་ཁ མགོ་ལོག་རོང་མ་འབྲོག་ཡུལ་སྐད་ཚོགས་པ་

Gorkä refers to an indigenous group of dialects spoken by cultivators and agro-pastoralists in the Golok region. It locally called གོར་ཁ MGO.SKAD/gorke/30 and is a tonal dialect according the description made by Wang (2012). Gorkä is spoken in the lower valleys of the following counties: Panma ལྭ་མ་, Gabde གདའ་བདེ, Darlak ཆར་ལག, Matö རྨ་སྟོད, Machen རྨ་ཆེན as well as part of Čigdril ཆིག་སྲིལ, Serta སེར་རྟ་ (Serthar སེར་ཐར) and Kandze ཆུ་ལོང་ཆེན་

• Gyālrongo-spheric Amdo དཀོན་ཟེར་འབྲུག་པའི་ཡུལ་སྐད།

A few dialects are spoken in the rGyalrong area and are influenced by the neighboring rGyalrongic languages.

➢ The Khalong dialect མཁྲིག, described by Sun (2007), is spoken by a community of sedentary pastoralists in Wuyi township up to Gangmuda township, in Dzamthang County རྣམ་ཐང་ (Sichuan). More research is needed to evaluate precisely its affiliation.

➢ The gSerpa dialect གསེར་པ་ is spoken in Yangge (Tib: ཡག་འགོ YAG.GO), Jiaxue (Tib: རྒྱ་ཤོ RGY.ASH.O), Xuri (Tib: SHO.RIB) townships and Wengda town

30. Jangbu (pers. comm.) proposes to write the language as MGO.SKAS (just as GLUSKAS’ melody pronounced as /gorke/ and not /gorkä/ or /gorkal/). However, it is not probable because the local pronunciation is /gorke/.
Inner classification of the Tibetic languages

(Tib: རྒོ་མདའ་ RBO.MDA) by an agricultural community at the south eastern corner of Serta County (Sichuan) near the GSER.CHU (Chin: Sequ) river. The neighboring rGyalrongic Zbu may have had an impact on gSerpa.

- Barkham Amdo is a dialect spoken by sedentary pastoralists in Barkham area and Phösul (South Dzamthang).

- Gyalrong Amdo was a lingua franca spoken by pastoralists and various speakers of various rGyalrongic languages, such as Situ, Tshobdun, Japhug or Geshitsa. It is still spoken essentially by merchants in these areas, distributed over several counties such as Barkham, Chuchen, bTsanlha, and Rongdrak.

- Yukhok dialect is spoken in Tau County (Yuke District) by pastoralists communities.

- Mroha dialect (derived from the word BROG.PA, and called Mosika 莫斯卡 in Chinese) is spoken by pastoralist communities at border of Chuchen County (Akeli Township) and Rongdrak (Dando Township).

**Dungnak and rTarmnyik** dialects neighboring the Turkic speaking Western Yughurs of Gansu:

- Dungnak is spoken in the Qifeng (Chin: 祁丰) Tibetan township (Qifeng Zangzu xiang) of Sunan Yughur Autonomous County, Gansu. It is located at the northwestern limit of the Amdo. See Shao (2018) for details.

- rTarmyik which literally means ‘horse’s hoof’ is spoken also in the Sunan Yuguzu Autonomous County in the Tibetan township of Mati (Chin: 马蹄, Mati Zangzu xiang), which is the Chinese translation of rTarmyik. The village of Mati is located near the eponym Tibetan monastery (also called RTABJE ‘horse trace’).
MAP IX. – Linguistic area of NE section

Legend:
- rTarmyik-Dungnak
- Hwari
- Arik
- Tsho Ngonpo
- Tsongkha
- Labrang-Rebgong
- Rwanak
- Gorkä
- Ngawa
- Mewa
- Washül
- Gyalrong surroundings
9.5.4. Number of speakers

The total number of speakers of the Tibetic languages located in the NE section is 809,360 (Qu 1996). Kalsang Norbu et al. (2000) gives the figure of 1,500,000 speakers. It is hard to give a precise figure because of the lack of a recent and reliable census and also because Hui (Chinese Muslims) in some counties such as Hwalong and Xunhua and even Rebgong area are native speakers of Amdo. Additionally, many Gyārlóng as well as speakers of Eastern languages such as Čone, Thewo, etc. also know Amdo as a second language.

The Mongols of Sogwo have been Tibetanized and speak mainly Amdo Tibetan. Some very specific dialects such as gSerpa are spoken by small numbers of people. Sun (2006) gives the figure of 6,500 residents for gSerpa.

9.5.5. Ethnic and sociolinguistic groups

The speakers of Amdo consider themselves as Amdowa, which is perceived as a strong identity marker among the Tibetans. Within Amdo, the Golok tshowes have a strong representation of their identity. However, as mentioned above, a minority of Amdo speakers living in Kandze area or in Chuchen do not regard themselves as Amdowa, but as Khampas or simply as drogpa ‘pastoralists’.

9.5.6. Phonological characteristics of the NE section

It is not possible to list common phonological features to the all the dialects of the NE section. The phonological characteristics are usually valid only at the level of the groups or even sometimes the dialects.

Suprasegmental features

It is well-known that Amdo, the Tibetic language of the NE section, does not have distinctive suprasegmental features. However, the position of stress may affect some phonetic realisations. Additionally, Gorkā Amdo has been reported to be a tonal dialect.

31. The figure of 1,800,000 is even mentioned by Ethnologue.
**Segmental features**

*Synchonic approach*

The sound systems of the NE section are characterized by the following frequent features:

- Multiple combinations of initials with a preinitial: prenasalization and preaspiration (see Chapter 7) are widespread features. Additionally, labial, velar and uvular preinitials are also found.

- Existence of voiced non resonant sounds (b, d, ɖ, g, dz, j, z, zh, y, fi).

- Some dialects have voiceless resonant series (m’, n’, ng’, ny’, l’).

- Labial prenasalization is found in all the pastoralists’ dialects (“d, “g, “ng, etc.) and homorganic prenasalization is pervasive (“b,”d, “g). Combinations such as “t’, “p’, “k’, or their labial equivalents (“r’, “k’), etc., are also found.

- Most dialects have an aspirated fricative series: s’, x’, sh’.

- Most dialects have uvular consonants: q (final), ϱ, χ.

- Amdo dialects have a rich set of final consonants, particularly pastoralist dialects.

- A limited set of vowels, which always include /ə/.

*Diachronic approach and reflexes of Classical Tibetan*

- In the languages of the NE section, the reflexes of preradical sounds are realized as segmental features, preaspiration and prenasals.

- Voiced non-resonant sounds (b, d, ɖ, g, dz, j) are derived from the consonants with preradicals (except M and ʃ). In some dialects, b, d, ɖ, g, dz, j may also be derived from the consonants without preradicals.

- Prenasalization corresponds to the reflex of radicals with a predical M or ʃ.

- As for aspirated fricative series, CT’S, SH, PHY without any preradicals correspond to s’, x’ and ʃh’.

32. In the Gyälrong surrounding dialects the reflexes yield mainly segmental features (rather than preaspiration).
• Preinitial (labial, retroflex, uvular, glottal, etc.) is triggered by all the CT preradicals (except ˚M and ˚).

• All CT final consonants B, G, M, N, NG, R are well preserved. In many pastoralist dialects the final ˚d changed into ˚l. The final ˚l is rarely preserved and the final ˚s never appears.

• The combination ˚SI and ˚ZL respectively yield: /ts/ and /dz/. This is the case in words like སློམ ‘easy’, སློབ ‘to teach’, བྲ་ སྲ ད བ་ ‘moon’.

9.5.7. Grammatical characteristics of the NE section

The Amdo dialects do not exhibit substantial differences in their grammatical systems. Languages and dialects which are spoken in the Gyālrong and Kham areas may have more grammatical specificities, but further research is needed to support this point. Among the major characteristics of the Amdo dialects, one can mention the preservation of inflectional verb morphology inherited from CT. In most case, two or three forms (present-future, past and imperative) are well preserved. In this aspect, Amdo language is the most conservative of the Tibetic languages. Another characteristic feature of Amdo dialects is ergativity. Whereas in many languages, the ergative is either limited to some tenses and aspects (mostly completed past) or largely optional and based on pragmatic factors, in Amdo, the ergative case is quite syntactic and usually compulsory for all tenses and aspects. Finally, demonstratives are postponed to the head nouns as in Central Tibet unlike many languages of the southern Himalayas.

9.5.7.1. Case markers

Some dialects of the NE section distinguish up to seven case markers. Frequent cases include ergative, absolutive, dative, ablative, genitive, locative, and associative.

The ergative marks agent of a transitive verb. It is used with both controllable and non-controllable verbs. The ergative case is found in all the dialects of the NE section.

It is compulsory in the various tenses (present, past, future) and aspects (completed-uncompleted)33 unlike in the languages of most other sections.

33. Or “perfective” versus “imperfective” as they are often called. About the terminological choice, see Chapter 8.
In the NE section, the marker for the ergative is /-ka/, which is derived from CT གིས GiS. In gSerpa, the ergative form is /yi/ which is derived from CT དིན YIS. In Amdo, ergative and genitive markers are usually homophones in most environments. However, the pronouns have usually distinct forms.

The absolutive marks the intransitive subject and the patient. It is always marked by zero (∅) in all NE languages.

The dative marks the beneficiary and sometimes the patient. In Amdo dialects, the dative marker is usually /-a/. In some contexts, the dative is not clearly pronounced. The dative is derived from CT ང་ LA. In gSerpa, the dative is /-la/.

The local or peripheral case markers are ablative, genitive, comparative, locative (inessive, illative).

The ablative is /-ni/, which is derived from CT ང་ NAS. Sometimes in Amdo, the ergative /ka/ is used to indicate the origin and functions as an ablative.

Most dialects lack a comparative case marker. In many dialects, the comparison is marked by various constructions in the Amdo dialects: one frequent construction is X(DAT)+BLTAS-NAY(ABS) V (སློབ་ནུ) lit. ‘if we look at X, Y is more’. The gSerpa comparative case /ve/ is probably derived from CT བས་ BAS, whereas it is /ni/ < CT ང་ NAS in Dungnak.

For the locative and allative functions, the dative /-a/ is often used. For the locative (without movement), /-na/, which is derived from CT ང་ na is also used.

The instrumental in Amdo dialects is formally identical to the ergative case: /-ka/ (གིས GIS) however in gSerpa, there is a special form: /-ke/.

As mentioned above, the genitive and ergative markers are homophones in most environments. The genitive is derived from CT དིན GI /go/. In gSerpa, the genitive is /yi/ as the ergative and is derived from the CT form དིན YL.

The associative corresponds to /-ra/ and its allomorph /-pa/ which has several functions (including connective functions). It is probably derived from CT ང་ R or ང་ LA.

4.5.7.2. Nominalizers

Various nominalizers are found in the NE section (concerning nominalizers, see section 8.3.13). We list below seven relatively frequent nominalizers of this section.
In Amdo /-na/ is very widespread nominalizers. It is probably derived from CT MI or MI which originally means 'person, human being'. It indicates generally the A (agent) of a verbal action.

Another very similar form /-no/ also derived from MI or MI followed by a definite marker BO /-po, -wo/ (also attested in Ladaks and Balti). It indicates the P (patient) and the subject of an intransitive verb (with completed past).

The nominalizer /-ja, -gya/ is frequent in Amdo. It is derived from CT RGYU and often followed by the definite marker /po, bo/ < BO and pronounced /-jo, -gyo/. It indicates the P (patient) and the subject of an intransitive verb (with the uncompleted aspect and the future).

Another frequent nominalizer is SA /-s'a/. This nominalizer is derived from CT SA, which means 'place' indicates 'the place of the verbal action'. In Amdo dialects, the nominalizer is often followed by the definite marker /po, wo/ < BO and pronounced /s'o/.

The nominalizer /-spa/ /-če/ is derived from CT spyad 'to use'. SKAL.BZANG 'GYUR.MED & SKAL.BZANG DBYANGS.CAN (2002) give the form BYED /-če/, which is also used in Literary Tibetan for the instrument. It is often followed by the definite marker /po, wo/ < BO and pronounced /-če-po/ or /-če-ko/.

The form /-so/ /-go/ derived from the noun SROL 'tradition, custom, habit' is found in many Amdo dialects.

The nominalizer /-t'o/ PHRO 'remainder' is a nominalizer used in some Amdo dialects to indicate what remains to be done.

9.5.7.3. Verbal inflections

One important characteristic of the Amdo dialects is the preservation of inflectional forms for many verbs. Lexical verbs often have two or three forms. In a few peculiar

---

34. However the reflex of BYED in Amdo is /she/ and not /če/ and the reflex of the spoken form is /ye/, so this etymology is improbable.
dialects such as gSerpa, one finds up to four inflections forms (“past”, “present”, “future” and “imperative”). The imperative form is an innovation as shown by Sun (2007). Most dialects have specific aspirated forms for the imperative, which are not inherited from CT.

In some dialects, a few frequent verbs have suppletive forms to indicate the various tenses and the imperative. For example, ‘to go’: སྣོང་ SONG and རྒོ་ GRO, ‘to come’: རྨྲ YONG and གོང་ SHOG; ‘to say’: ཐེར ZER and བྱེས་ BZLAS, ‘to give’: རིན STER and བྱིན BYIN.

9.5.7.4. Linking verbs and auxiliary verbs
Linking verbs correspond to equative (copulative) and existential verbs.

Copulative verbs

The verbs ཤིན་ YIN and རེད་ RED are used both as copulative verbs. YIN usually conveys an egophoric meaning while RED has a factual meaning.

These auxiliary verbs have the following negations in most dialects: ཤིན་ YIN > རིན MIN; རེད་ RED > རིན MA-RED.

Existential verbs

Existential auxiliaries exhibit some variations both in form and meaning. In nearly all the dialects ཤོད་ YOD is used. gSerpa has a reflex of the archaic form རྲོད་ ’OD.

ཤོད་ YOD or its variant རྲོད་ ’OD are used to refer to personal information (egophoric) and generally occurs with the first person subject.

The compound existential verb རྲོད་ བཞིན YOD.GI indicates a sensory marking while རྲོད་ བཞིན YOD.NLRED indicates a factual meaning.

These auxiliary verbs have the following negations:

- ཤོད་ YOD > རིན MED. The archaic form རིན MYED is found in gSerpa.
- རྲོད་ བཞིན YOD.GI > རིན MED.GI
- རྲོད་ བཞིན YOD.NLRED > རིནMED.GI YOD.NLMA-RED
Compound linking verbs

Compound verbs are very frequent. They involve the combination of various copulas or auxiliaries. Aside from ◊ཡོད་ནི་རེད་ YOD.NL.RED mentioned above, we find many combinations of linking verbs such as ◊ཡིན་ནི་རེད་ YIN.NL.RED, ◊ཡིན་རྒྱུ་རེད་ YIN.RGYU.RED, etc.

Auxiliaries

Frequent auxiliary verbs of the NE section consist of the linking verbs (sometimes preceded by a relator): ◊ཡིན་ YIN, ◊རེད་ RED. Other frequent auxiliaries include: ◊ཐལ་ THAL (all dialects), ◊ཟུག་ ZUG, ◊སོང་ SONG, ◊བཏང་ BTANG, ◊གདའ་ GDA’.

From a semantico-cognitive point of view, the auxiliaries (together with relators) convey complex temporal, aspectual, evidential and epistemic meanings.

Most of the systems have special forms to mark sensory access to information, as well as factual, egophoric and inferential meanings. (See Tournadre & Shao, forthcoming.)

9.5.7.5. Negation

The negation has two forms in all the Amdo dialects of the NE section: ◊མ་ MA and ◊མི་ MI. The archaic form ◊མྱི་ MI is attested in gSerpa. For the imperative, one can only use the form ◊མ་ MA. Other than these markers, the form ◊ཆིས་ CHIS, maybe derived from ◊ཆིས་ CHIS ‘by what’, as well as the marker ◊ན་ NA are also marginally attested (see also 9.3.7.5 and 8.4.11). Generally speaking, in Amdo dialects, the negation in the various tenses and aspects often comes before the lexical verb, even in the presence of some auxiliaries such as ◊ཐལ་ THAL, but it may also occur after the lexical verb, before the auxiliary.

9.6. The Central section

The Central section (henceforth C section) is made up of one large linguistic set of dialects spoken in Central and Western Tibet. This groups divides into several groups in a similar way as the NE section (Amdo). Thus, we can say that the C section corresponds to a single language which is often referred to as ‘U-Tsang’, with considerable dialectal variation.
‘Common Tibetan’ བོད་སྲིད་ ‘Chikä’, which is based on བོད་ཁབ་ ‘Lhasa dialect’ /lhäs a kä/, the prestigious dialect of the capital Lhasa, a variety of བོད་ཁབ་ ‘Ü-kä’, is now spreading through the media. People in Nagchu and Chando areas are frequently exposed to Common Tibetan for administrative reasons. Although Kham and Hor dialects are spoken in these two prefectures, their administrative affiliation to the TAR facilitates the diffusion of Common Tibetan.

It is also the language used in the Tibetan diaspora. However, the standardization is not totally achieved since “Common Tibetan” is not taught in schools and is usually not written down.

Both “Tibetan” and “Chinese” are official languages of the Tibet Autonomous Region. The Chinese law does not specify which type of “Tibetan” is the official language: “Literary Tibetan” or Common Tibetan based on the “Lhasa dialect,” or both? Given the very significant diglossy, it is an important issue. However, despite the legislation, the main language used in the school curriculum is de facto Mandarin Chinese. This language has thus some impact on the oral dialects of Central Tibet. The ‘Ü-Tsang’ dialects are usually not written down, and educated people prefer to use Literary Tibetan for written purposes. The Literary language is also used in the Buddhist and Bönpo monasteries, in the institutes of Traditional medicine (བོད་ཁབ་ ’Mäntsikhang) and to a certain extent in some cultural organizations and in the media, particularly on the internet. The Tibet University བོད་ལྗོངས་སློབ་གྲྭ་ཆེན་མོ་ Böjong Lobdra Chenmo located in Lhasa, has a Tibetan language department and various departments using Tibetan in the curriculum. It also provides a Tibetan language course for foreigners, in which the Tibetan oral language is often called by the Tibetans ‘Tibet University speech’, and perceived as a sociolectal variety.

The cultural and economic capital of Tibet is also traditionally a great center for pilgrimages. The Potala palace རྡེ་ཕ་བྲང་/པོ་ཏ་ལ་ and the Norbu Lingkha སེར་བུ་གླིང་ཁ་ were traditionally the residences of the Dalai Lamas and the Tibetan government.

35. From a political point of view, it is only nowadays the capital of the Tibet Autonomous Region.
In the C section, Vajrayāna Buddhism is the main religion with followers of the four sects i.e., Gelugpa, Sakya, Kagyü and Nyingmapa. Bönpo communities are also attested throughout the region. There are also Muslim and a small Christian minority, mainly living in Lhasa.

Among the major Buddhist institutions of the C section, we find: the Jokhang, Ramoche as well as the ‘three seats’ of Drāpung, Sera and Gandān all located in Lhasa or in the vicinity. In other Central areas, outside the capital, we find the following major Buddhist and Bön monasteries: Radreng (alt. Reting), Drigung thil, Tshurphu (Kagyü), Trashi Lhünpo, Pālkhor Chöde (Geluk), Nyethang Drölma Lhakhang (Geluk); Sakya and Gongkar Chöde (Sakya); Mindröl Ling (ecumenical), Lama Ling (Nyingma), Yungdrung Ling (Bön) and Mänri (Bön).

The main works on the languages of C section are mainly devoted to the Lhasa dialect or “Common Tibetan.” We will only mention the major textbooks or grammars here and some pioneer works: Roerich and Phuntshok (1957), Chang Kun and Betty Shefts (1964), Goldtsein and Nornang (1978), Losang Thonden (1986), Hu Tan et al. (1989), THUB.RBSTAN DBANG.PO et al. (1996), Kitamura (1977), Hoshi M. (1988), Wang Zhijing (1994), Tourmadre (1996a), Tourmadre & Sangda Dorje (1998, 2003), Blondeau et al. (2002, 2014), Chonjore Tseten (2003), Denwood (1999), Mélac et al. (2014). Tsang dialects have been much less described. The essential contribution by Haller (2000) is devoted to the Zhikatse dialect. The dialects of Ngari have received more attention. The main references are: Qu and Tan (1983) for a survey of the Tō dialects of Ngari Prefecture (but it does not include the Tō dialect spoken in Zhikatse Municipality), Herrmann (1989) on Tō cultivators’ dialect of Dingri, Kretschmar (1986) on Western Drokpa, a Tō pastoralists’ dialect spoken in Drongpa County. The various dialects of Lhokha, Phānpo, Kongpo, Lhobrak and Dagpo have received little
attention and there isn’t any available overview of these dialects. Tournadre and Jiatso (2001) mention some elements of the verb morphology of the Phänpo dialects.

9.6.1. Migration patterns, legends and historical records

From a historical point of view, the Chonggyä and Yarlung valleys are considered the cradle of the Tibetan culture during the rise of the Tibetan Empire. The implantation of Tibetans in the Kongpo area is also very ancient. According to historical records, the migration westwards towards the Ngari area dates back to the first half of the seventh century, when the Tibetan armies annexed the Zhangzhung Kingdom.

Since Lhasa was the capital of Tibet during the Tibetan Empire and subsequently during the Dalai Lamas’ reign, and it was also a main center for pilgrimages, it has received numerous influences from all the regions of Tibet.

9.6.2. Linguistic groups of the C section

The dialectal variation within the C section is mainly geographic, but it also depends upon sociolinguistic parameters, and one can distinguish in some cases pastoralists’ and cultivators’ dialects, especially in the Tö Ngari region, in Western Tibet.

Communities of གོང་པ་ (drogap) ‘pastoralists’ and བཞིང་པ་ (zhingpa) ‘cultivators’ or གསང་པོ་ (samadrok) ‘agropastoralists’ are also found in Ü, Tsang, Lhokha, Kongpo areas, however, the dialectal variation is not clearly based on these sociolinguistic parameters unlike in the case of Tö Ngari or Amdo.

All the dialects spoken in this section generally allow a rather good mutual intelligibility, but there is a gradual variation between the eastern and western limits of this section. This section can be better described in terms of a geolinguistic continuum. Although Kongpo dialect in the east can fairly easily communicate with adjacent dialects such as Ü or Lhokha and these latter dialects are also mutually intelligible with Tsang and Tö, one can safely state that dialects at the two extremes of the area, such as Tö and Kongpo, will not allow for easy communication.

The central section is made up of eight groups spoken in of dialects spoken in Central and Western Tibet.
For the dialect classification of the C section, we propose the following eight groups:

- Ü འབྲུས་སྐད་
- Tsang གཙང་སྐད་
- Phänpo འཕན་པོའི་སྐད་
- Tö pastoralists’ dialects འབྲོག་པའི་སྟོད་སྐད་ ‘Drogpä Tö-kä’
- Eastern Tö cultivators’ dialects དབྱངས་པའི་སྟོད་སྐད་ ‘Sharchok Rongpä Tö-kä’
- Western Tö cultivators’ dialects དཔོན་པའི་སྟོད་སྐད་ ‘Nubchok Rongpä Tö-kä’
- Kongpo ཀོང་པོའི་སྐད་
- Lhokha ལྷོ་ཁའི་སྐད་

Two historical regions of southern Tibet, འབྲོག་སྐད་ Lhobrak and འབྲུག་གསུམ་མཚོ་ Dagpo spoken in the Lhokha prefecture may have linguistic specificities that would require the creation of two additional groups. However, we do not have reliable data on the dialects spoken in these areas.

Several dialects of the C section are in contact with a few Non-Tibetic languages: Ba-ke བྲག་སྐད་ (a.k.a Basum, see 10.3), spoken in Kongpo around the Dragsum lake ལྷོ་གསུམ་མཚོ་, Dakpa (a.k.a ‘Tshona Mönpa’) and Tshangla (in Lhokha and Kongpo), Bokar and Idu (Tani languages spoken in Kongpo).

9.6.3. Geographic extent of the C section

From an administrative point of view, the dialects of the C section are all spoken in the southern and western parts of the Tibet Autonomous Region which has been created in 1964 after the integration of Tibet into the People’s Republic of China in 1950. Note that in the northern and eastern prefectures of Nagchu and Chamdo, which also belong to the TAR, Hor and Kham dialects are spoken.

The area of the C section corresponds to the historical regions of Ü འབྲུས་, Tsang གཙང་, Ngari རྒྱ་, Lhokha ལྷོ་ཁ་ and Kongpo ཀོང་པོ་.
The cradle of the Tibetan civilization is often associated with the 'Yarlung valley' also spelled 'YAR.KLUNGS'. This valley is located at the confluence of the Yarlhasham River (eponym of the mountain and deity Yarlhashampo) and the Chonggye river, which is located near Tsethang in Lhokha. In the Chonggye valley are located the tombs of the Tibetan emperors.

The main river of the Tibet Autonomous Region is the Yarlung Tsangpo river also known as the Brahmaputra (on the Indian side of the border), which originates from the Tsho Mapham lake located in Ngari Province (in Western Tibet), and flows towards the Kongpo area (in Eastern Tibet) through Tsang, Lhokha and Central Tibet. In its upper course, the Yarlung Tsangpo river is called Tamchok 'flowing from the horse’s mouth'. The Yarlung Tsangpo is about 2,900 km long (see Appendix 1).

Other significant rivers of the C section include the main tributaries of the Yarlung Tsangpo: the Nyangchu (flowing through Gyantsé and Zhikatse), the Kyichu (flowing through Lhasa), the Kongpo Nyangchu (flowing through Kongpo Gyamda and Nyingthri) and the Pharlung Tsangpo (in Pomā area). Thus, the Yarlung Tsangpo and its tributaries represent the core of the C section area. Additionally, two other large rivers, which flow in Western Tibet should be mentioned: The Sengge Khabap (Indus river) and the Langchen Khabap (Sutlej river).

36. Both are pronounced in the same way as /yarlung/. Note that KLUNG(s) means 'river' while LUNG refers to 'valley' (see Appendix 2). The name 'Yarlung valley' generates some confusion since it does not designate the valley of the Yarlung Tsangpo, the main river of Tibet (called Brahmaputra in India), but the lower course of the Yarlhasham river (also sometimes called Yarlung river!), which joints the Yarlung Tsangpo near Tsethang.

37. Also spelled /yarg/ (in the ancient document).

38. One should not confuse MYANG-CHU and NYANG-CHU which are both pronounced 'Nyangchu' (see Appendix 1).
Detailed location of the Ü-Tsang dialects

- Ü བུས་སྐད་ is spoken in the following counties of Lhasa Municipality: Lhasa ལྷ་ས་, Chushur ཅུ་ཤུར་, Tönlung Dechen སྟོད་ལུང་བདེ་ཆེན་ and Tagtse སྟག་རྩེ་.

- Phänpo འཕན་པོ་སྐད་ is spoken in the following counties of Lhasa Municipality: Lhündrup ལྷུན་གྲུབ, Madrogungkar སྲོད་གང་རྒྱུས་, and part of Tagtse སྟག་རྩེ་.

- Tsang གཙང་སྐད་ is spoken in Zhikatse Municipality in Zhikatse གཞིས་ཀ་རྩེ་ and in some counties of Lhasa Municipality, Namling རྣམ་གླིང་, Rinpung རིན་སྤུངས་, Gyantse གཡུང་རྟེ་, Panam གོན་པ་, Khangmar གང་མར་, Sakya སྐྱ་, Tingkye དིངས་རྒྱུས་, Gampa གམ་པ་, Zhāthongmön བཞད་མཐོང་སྨོན་, Chushur ཅུ་ཤུར་(partly) and most of Nyemo County སྙེ་མོ་ and north of Dromo རྱུས་ for the southern part of Dromo, see 9.7)

- Lhokha ལྷོ་ཁ་སྐད་ is spoken in Lhokha Municipality in Nedong (Tsethang) རྩེད་ཐང་, Chonggya ང་གྱེས་, Gongkar ཐོང་དཀར་, Nankartse སྣ་དཀར་རྩེ་, Dranang ཆུ་ནང་, Chusum སྣུ་མ་, Zangri བཟངས་རི་, Lhüntse གླུ་རྩེ་, Tsomä བསྟོད་ཚད་, and part of Tsona སྡོམ་ནོར་.

As mentioned earlier, the dialects spoken in the Dagpo region of Gyatsa County རྒྱ་ཚ་ and Nang County སྣང་ are not documented. The same is true for the dialect spoken in the region of Lhobrak རྒྱལ་. So their affiliation needs further research.

- Kongpo གོང་པོ་ is spoken in Nyingthri Municipality, in Kongpo Gyamda, Nyingthi, Mänlung and Powo County.

- Tö cultivators' dialects རོང་པའི་སྟོད་སྐད་ ‘Rongpä Tö-kä’ may be subdivided in two groups: the eastern and the western groups. The Western Tö cultivators' dialects are spoken in Ngari Prefecture in the following districts: Sengge Tsangpo (west of Gar ཙང་, Ruthok (west), Tsanda ཡུལ་དཀར་, and Purang (Pureng) གླུ་རྡེ་. Eastern Tö cultivators' dialects are spoken in Zhikatse Prefecture in Ngamring གཞིས་ཀྱི་(the southern part), Nyanang ཉེ་ཐང་ (also called Nyalam), Dingri རུང་, Lhatse སྣ་སྟེ་ and Kyirong (north) གཞི་རོང་.

- Tö pastoralists' dialects ཀྲོད་པའི་སྟོད་སྐད་ ‘Drogpä Tö-kä’ are spoken in Ngari

39. In the northern part of Nyemo County (Markyang township), a Hor variety is found.
prefecture, in the northern part of Zhikatse Prefecture and in the western part of Nagchu prefecture. In Ngari Prefecture: Ruthok རུ་ཐོག་, Sengge Tsangpo (the eastern part of Gar སངས་, Götshang Tö གོ་ཐོ, Götshang Mä གོ་ཐོ་མ་ and Gyamuk གོ་མུ), Geyä གླ་ཡའ (north), Gertse གེར་ཙེ་ 40 and Tshochen གློ་ཆེན་. In Zhi katse Prefecture: Drongpa ཀློང་པ་, Saga སྐག་, Ngamrim གང་རིམ་ (the north Latö area), Zhäthongmön བཞེན་ཐོང་མོན་ (north) and Namling སྐྱིད་ (north). In Nagchu prefecture: Nyima ཐི་མ་ (Nagtshang area), Shäntsa སྣྱ་ཝ་, Pängön གང་ོན་ and Tshonyi གོ་ནོད་. In the latter prefecture, Tö dialects are spoken together with Hor dialects.

9.6.4. Number of speakers

There are about 1,000,000 of Ü-Tsang dialects in a territory of half a million km², i.e. approximately half of the Tibet Autonomous Region (in the other half on the TAR the Kham-Hor dialects of the SE section are spoken). Ethnologue also gives a figure of more than 1,000,000 speakers. In order to give an idea of the total repartition within the Ü-Tsang group of Ü, Tsang and Tö dialects, we give Qu Aitang’s estimation (1996): Ü—569,222, Tsang—457,660, Tö—38,319. According to Qu (ibid.), the total of ‘Ü-Tsang’ speakers is 1,065,201. However, it is probably an underestimation and we do not have reliable recent figures. There is no data for the number of Kongpo and Lhokha speakers.

Due to the lack of official statistics and the constant evolution of the situation, it is difficult to give any precise figure. In some areas of Kongpo such as Bayi, some young Tibetans are no longer able to speak in Tibetan and use mainly Chinese. Conversely, the number of speakers using སྤྱི་སྐད་ ‘Common Tibetan’ (based on Lhasa dialect) as a second language is rapidly growing within the Tibet Autonomous Region and to a lesser extent among the elite of Kham and Amdo. ‘Common Tibetan’ is also the main language of the Tibetan diaspora which amounts to 130,000 people, throughout the world.

40. In Gertse “downtown” a Kham variety is spoken. See SE section.
Inner classification of the Tibetic languages

Legend:
- : Tö pastoralists
- : Tö cultivators
- : Tsang
- : Ü
- : Lhokha
- : Phânpo
- : Kongpo

MAP IX.5. – Linguistic area of C section
9.6.5. Ethnic and Sociolinguistic groups

The speakers in the C section are predominantly རིག་པ། zhingpa ‘cultivators’ and agropastoralists (notably in Tsang, Phänpo and Kongpo areas) which are locally called དམ་པོ་ samdrok or འབྲེལ་པོ་ bömdrok (in Nyemo County). The pastoralists or དྲོག་པ droga are essentially found in Tö Ngari, but some small groups of cattle-breeders are also encountered elsewhere in Kongpo, Ü, Tsang and Lhokha. Within the Central area, ལྷ་ས་ Lhasawa /läsä:/ and གཙང་པ་ Tsangpa are generally perceived as strong identities.

9.6.6. Phonological characteristics of the C section

The phonological diversity of the C section is rather limited. However, it is not possible to assert common phonological features for all the dialects of the C section. The phonological characteristics are usually valid only at the level of the groups.

Suprasegmental features

It is well-known that the dialects of the C section have distinctive suprasegmental features and generally have four distinctive pitch and contour tones.

Segmental features

Synchronic approach

The sound systems of the C section are characterized by the following frequent features:

- Most dialects have simple initials and a few dialects (notably Tö) have prenasalization before voiced obstruents.
- The majority of dialects lack voiced non-resonant sounds (b, d, q, g, dz, j, z, zh)
- Some dialects have voiceless resonant series (r’, m’). The /m’/ occurs with the negation in Lhasa (some speakers) and Tsang in front of aspirated MA [m’a] PHYIN. But these sounds have a marginal status.
- Final /p/, /m/, /ng/ are usually preserved.
- A limited set of vowels which include central rounded vowels /ö/ and /ũ/ as well as midhigh unrounded vowel /ä/. 
• The majority of dialects have nasalized vowels. The vowel length is distinctive.

Diachronic approach and reflexes of Classical Tibetan

• In the languages of the C section, the reflexes of preradical sounds are essentially prenasals (see Tö, Ü) or they yield no trace (Tsang).

• Low pitch aspirated sounds are derived from voiced non-resonant sounds (B, D, G, DZ, J).

• Prenasalization corresponds to the reflex of radicals with a preradical M or ' before voiced consonants.

• In Ü and Tsang, the preinitial L of the second syllable triggers a nasal sound as in /བོད་ལྗོངས་/BOD.LJONGS//’p’o’jong/ and /དགའ་ལྡན/ /DGA’.LDAN//’ka’dän/.

• The following final consonants B, M, NG, R are well preserved. The final L is preserved only in Tö dialects and the final S and D never appears.

• The combination ZL yields: /’d/ or /’t/, /ཟླ་བ/ ZLA.BA//’ndlawa/, /’tawa/ ‘moon’.

9.6.7. Grammatical characteristics of the C section

The languages and dialects of the C section exhibit some grammatical differences particularly in their verbal and nominal morphologies. Discrepancies in the grammatical functions are also attested. Among the major differences, we may mention that the dialects of Ngari have a non-visual form of sensory /’tä/ < CT རྒྱ་GRAG, which is also present in Spiti, Ladaks and Dolpo but not in the central dialects of the C section. The auxiliary of Ü, Ts, Phänpo and Tö are morphologically very diverse, particularly the verb ‘to be’ (see e.g. Tournadre and Konchok Jiatso 2001). Concerning verbal inflections, the languages of this section have preserved a limited number of forms found in CT.

Another significant difference is found in the nominal morphology. The Ngari dialects have a comparative form *སངʼ SANG (also found in Ladakh, Spiti and a few other areas) which corresponds to བས་ LAS /’lâ/ in the Central dialects. These remarks confirm the existence of a geolinguistic continuum between the Central, Western and North-Western sections.
Case markers

Most dialects of the C section distinguish usually six case markers. Frequent cases include ergative, absolutive, dative, comparative, ablative and genitive (for the Kongpo dialect, further research is needed). Three additional marginal cases, the locative, adessive and associative are found in some dialects (for example, the locative /tu/ in Shigatse, see Haller 2000).

The ergative marks the agent of a transitive verb. It is used with both controllable and non-controllable verbs. The ergative case is found in all the dialects of the C section. It is not always compulsory and in some dialects is mainly restricted to the completed past (or “perfective”).

In the C section, the marker of the ergative is /-ki/ which is derived from CT གིས GIS, except Tsanda which has /-shi/. Exceptionally, in Tsanda (Tö Ngari) and Spiti, the form /-su/ related to CT ལྟ་ LA is also found.

The absolutive marks the intransitive subject and the patient. It is always marked by zero (∅) in all NE languages.

The dative marks the Beneficiary and sometimes the patient. In the dialects of the C section, the dative marker is usually /-la/ or /-t/. The dative is derived from CT ལས་ LAS. LA.

The local or peripheral case markers are ablative, genitive, comparative, dative, locative adessive and associative.

The genitive is either identical to the ergative or differs only by a distinct tone. The genitive is derived from CT གི། GI.

The various forms of ablative are /-ni/, /-na/, /-nä/, which are derived from CT ལས་ NAS.

For the comparative case, the dative /-lä/ is often used in most cases. It is derived from CT ལས་ LAS. In some Tö dialects, other forms are found /sum/ (Tsanda, Gar),
The forms ꞌsangꞌ /-sang/ is also found in the languages of the Western and Northwestern sections such as Spiti and Ladaks.\textsuperscript{41}

The instrumental in the dialects is formally identical to the ergative case: /-ki:/ (སྐོར་ Gṣ).

The associative corresponds to /-tang/ or /-ta/ which are derived from CT ꞌDANG.’

9.6.7.2. Nominalizers

Various nominalizers are found in the C section (see 8.3.13). We list below six frequent nominalizers. They are normally shared by all the dialects, but they may differ in their pronunciations and specific functions.

- In this section /-ngän, -kän, -k’an/ is a very widespread nominalizer. These marks are derived from CT ꞌmḳan ‘expert person’. It indicates generally the A (agent) of a verbal action.

- The nominalizer /-pa/ or /-pi/ derived from CT ꞌPA indicates the P (patient) and the subject of an intransitive verb (with completed past).

- The nominalizer /-ya/ or its variants /-yä/ or /-zha/ is frequent in the C section. It is probably derived from CT ꞌCHAS ‘thing, tool’. It is also used in the Western, North-western and Southern sections. Thus it is used for the infinitive in Ladaks: ཇོ་ཀ་ZACES /sa-čes/ and in Lhoke: ཇོ་ཀ་ZASHAD /sa-shäʔ/ (the Tibetan spellings SHAD and CES reflect only the pronunciation and not the etymology). The nominalizer /ya/ (or its variants) indicates the P (patient) and the subject of an intransitive verb (with the uncompleted aspect and the future) and has a similar meaning as RGYU.

- The nominalizer /-gyu/ is attested in some dialects. It is derived from CT ꞌRGYU ‘object’ which is also used as a nominalizer in CT. It indicates the P (patient) and the S (single argument) of a monovalent verb (with the uncompleted aspect and the future).

\textsuperscript{41} They might be derived from CT TSANA or TSANG which indicates the cause or the origin. Another hypothesis is that these markers are derived from a more archaic form present in proto-TB as suggested by Nishida & Sun (1990). Cognate forms are found in Baima, in Tangut and Ngari as well as Qiang (see e.g. Huang 2007: 123).
Another frequent nominalizer is མ་/sa/. This nominalizer is derived from CT མ་, which indicates "the place of the verbal action" and various other grammatical meanings.

The nominalizer /-tang/ derived from the noun སྟངས STANGS 'way, manner, appearance, look' is found in nearly all dialects of the C section.

9.6.7.3. Verbal inflection

The number of inflectional forms is quite limited. Most lexical verbs are either invariable or have two, rarely three forms ("past", "present-future" and "imperative"). See the section 8.3.6 of Chapter 8 for examples. In some dialects, a few frequent verbs have suppletive forms: 'to go': བྱིན PHYIN (past) and གྲོ GRO (present-future), རྒྱུགས RGYUGS (imperative) 'to come' < CT 'to run'; ཡོང YONG (present-future-past) and ལྷོག SHOG (imperative) < CT GHEGS 'to go, come'.

9.6.7.4. Linking verbs and auxiliary verbs

Copulative verbs

The verbs ཨིན yin is used as a copulative verb in all the dialects with an egophoric meaning (usually associated with the first person). The negative forms are /man, män/ ་མན་.

For the factual meaning, various forms are found བྱིན /'re:/ in Ü, Phânpo and Kongpo as well as some dialects of Ngari (Gertse, Gegye, Tshochen), where a varieties of Kham and Hor are spoken, བྱིན བྱིན /'pa:/, /pie/ and its variants /'po:/, /'pu:/, /'po:/ (Tsang and Lhokha) and བྱིན /'da:/ or /'ta:/ (in most Tö dialects) (see Qu & Tan 1983), /'da:/ (Tö, Lhatse) or /'te/ (Tö, Dingri). For the origin of these markers, see section 8.3.3.

The negative forms of these copulas are:

• 'mare: བྱིན བྱིན in Ü, Phânpo and Kongpo as well as some dialects of Ngari (Gegye, Tshochen).

• /män-pi:/, /män-po/, /män-po/, /män-po/ (Tsang and Lhokha) which may correspond to བྱིན བྱིན MAN-SBAD < ? CT བྱིན MA-BYED.
Existential verbs

In the dialects, various forms */^yo:/, */^hö:/, */^wö:/ are found for the existential egophoric verb. They correspond to the reflexes of CT 'YOD or the variant 'OD; These forms are used to refer to personal information (egophoric) and generally occur with the first person subject.

The sensory forms are */^du?/ or */^nu/. They correspond to the reflex of CT 'DUG. The auxiliary */^nang/ derived from CT SNANG is found in the Phänpo varieties. GRAG is used in Ngari for non-visual sensory meaning (see 8.4.3).

These auxiliary verbs have the following negations:

- */^me?/ (Tö), */^me:/ (Ü) */^mi:/ (Ts) < CT 'MED;
- */^mi-nang/ < CT 'MLSNANG;
- */^mi^du?/ < CT 'ÌNANG;
- */^mi^ta?/ < CT 'ÌNANG.

Compound linking verbs

Compound linking verbs involve the combination of various auxiliaries.

The forms indicating an existential factual meaning are:

- */^yo:re:/ (Ü), */^hö:pare:/, */^hö:re:/ (Phänpo), */^wö:re:/ (pastoralists' Tö) are the reflexes of CT 'YOD.PARED or 'OD.PARED or 'OD.DE.RED.
- */^yo:pe/, */^yo:pa/, */^yo:po/ (Tsang and Lhoka); 'YOD.BA.BSAD < 'YOD.BA.BYED.
- */^wö:da?/, */^hö:tae/, */^hö:te/, */^höka^da?/ (Tö) respectively derived from 'OD.PARDAG, 'OD.DE'.DAG or 'OD.MKHAN.DAG.
Auxiliaries

Frequent auxiliary verbs of the C section are made of the linking verbs (sometimes preceded by a relator):

- ིིན་YIN
- རེད་RED (Ü), འཇིག་SBAD /"pä:/, /"pa:/, /"po:/, /"pu:/ /"pa:/ (in Tsang)
- ང་འདག་/"da?/, /"da:/ or /"te/ (in Tö). The latter forms are also probably cognate with /nak/ attested in Loke (Mustang, see South-western section), and /'na/ or /-'da/ (in the negation /'man-da/ and inferential /'yin-da/) in Brokpa, Bhutan (see Southern section).
- ིོད་YOD or ཡོད་/"yö:/, /"ʰö:/, /"woʔ/
- ཐུག་/"DAG (Ü, Ts) or སྣང་SNANG (Phänpo)
- བྱུང་BYUNG
- སོང་SONG (or སྣང་SNANG)
- གྲག་GRAG /'täʔ/
- བཞག་BZHAG

Other frequent auxiliaries include: བོགས་DGOS, ཡོང་YONG, བསྡད་BSDAD.

From a semantico-cognitive point of view, the auxiliaries (together with relators) convey complex temporal, aspectual, evidential and epistemic meanings.

The systems have special forms to mark sensory access to information, as well as factual, egophoric, inferential and hearsay meanings.

9.6.7.5. Negation

The negation has two forms in all the dialects of the C section: མ་MA used and མི་MI. For the imperative, one can only use the form མ་MA. Generally speaking, in most dialects (with the notable exception of Tö Purang), the negation used in the various tenses and aspects comes before the auxiliary verb (i.e. after the lexical verb) and not before the lexical verb. As in most other Tibetic languages, the negative auxiliaries ིིན་MIN and རེད་MED often mark the negation.
In Lhasa, the pronunciation of the negation prefix depends on an initial consonant of the verb stem. If it is voiceless aspirated, the initial of the negation prefix often becomes voiceless /m'/.

9.7. The Southern section

The S section is made up of several linguistic groups of dialects mainly spoken in Bhutan and Sikkim (India) and the Dromo County of TAR (China). There is limited linguistic diversity within this section. However, there is no mutual intelligibility between the groups of dialects.

Dzongkha དོང་ཁ་, lit. ‘the castle language’, is the national and official language of Bhutan. Although English does not have an official status, it is widely used in administration, education, media and businesses.

Both English and Dzongkha are taught in the schools, but English is the main medium. Nepali, a strong regional language, has also had an impact on the linguistic situation in Bhutan. Written Dzongkha is increasingly used in the media, in the schools and on the internet. Literary Tibetan, which is called Chökä ཆོས་སྐད་ i.e. ‘Dharma language’ is used in the Buddhist monasteries.

In the Indian state of Sikkim, Lhoke ལོ་སྐད་ (LHO,SKAD) Lit. ‘the southern Language’ is an official language of Sikkim. There are alternative names such as ཉི་བཙོན་བོན་ ལོ་སྐད་ (BRAS,LJONGS,SKAD) Dränjong-kä often spelled ‘Denjongke’ (lit. ‘language of the fruit/rice region’) or Bhutia. However English and Nepali, the lingua franca, are used as the main medium in the school curriculum. Since the end of the 1960s a written form of Lhoke, called བློ་ཡིག་ ‘Lhoyik’, has emerged. It is written in the Tibetan script. Lhoyik has developed a fairly high level of standardization and is taught in the schools as one of the eleven “official languages”: Nepali, Lhoke (or ‘Bhutia’), Lepcha, Limbu, Newari, Rai, Gurung, Magar, Sherpa, Tamang and Sunwar.

A few novels and plays are written in Lhoke. However, Lhoyik has still a limited diffusion and unlike Dzongkha is not present on the internet.

Thus, in both Bhutan and Sikkim, English and Nepali have a significant impact on the ecolinguistic system. Within the S section, the Tibetic languages are also in
contact with various Tibeto-Burman languages depending on the area: Lepcha, Gurung, Magar, Sunwar, Rai, Tamang, Limbu, Bumthang, Kurtö, Tshangla or Sharchopkha, etc.

Among the scientific and cultural institutions of Bhutan, one should mention the National Library of Bhutan Druk Gyälyong Pedzö located in the capital and the Khesar Gyalpo University of Medical Science Gesar Gyalpo Sorig Tsuglagkhang Lobde; the department of traditional medicine (ministry of Health) Nangmän Zhabto Läkhung; the Dzongkha development commission Dzongkha gongphel lhántshok.

Unlike in Tibet, most Bhutanese dzongs 'castles' are well preserved. Major dzongs include: Trashi Chödzong (in Thimphu), Punakha, Trongsar, Lhuntse, Paro (alt. Rinpung) and Trashigang.

In the S section, Vajrayāna Buddhism is the main religion. In Bhutan, Buddhists are mainly followers of the Drugpa Kagyü and to a lesser extent the Nyingma sect. There is also a minority of Hindus.

Major monasteries of Bhutan include: Chagri Mingyur Dorjedän (near Thimphu), Dechen Phodrang (near Thimphu), Paro Tagtshang, Gangteng (in Wangdü Phodrang) (Nyingma), Kuje Lhakhang (in Bumthang district), Nalanda Buddhist Institute (NBI), Phajoding, Sumthrang Samdrup Chödzong (Bumthang) and Tamzhing Lhundrup chöling.

In Sikkim, Buddhists are followers of the Nyingma and Kagyü sects. There is also a small Christian community. They are also in contact with neighboring Shamanist and Hindu communities. The major monasteries of Sikkim include: Pema Yangtse (Nyingma), Trashi Ding and Ralang Pälchen Chöling.

Among the cultural institutions of Sikkim, one should also mention the Namgyal Institute of Tibetology.

Lhoke, despite the fact that it is an official language of Sikkim, is not well documented yet. There has been a pioneer work by Graham Sandberg (1894). However, some publications (textbooks or articles) have recently appeared: Denzongpo et al. (2009; 2011), and Yeshe Rigzin Bhutia (2008). Yliniemi (2019) has written a comprehensive grammar of Lhoke for his doctoral dissertation, which is a major contribution to the description of this language.

9.7.1. Migration patterns, legends and historical records

According to Balikci (2008: 6):

“Tibetan settlers came to Sikkim from the neighboring valleys of Chumbi and Ha and regions beyond these southern valleys such as Kham Minyak from the thirteenth century onwards and established the kingdom in 1642. Their descendants call themselves Lhopo (Lhopa, lit. ‘people from the south’) but are generally known as Bhutia, Sikkimese or even Denjongpas [Dränjongas].”

Sikkim is called ‘Dränjong’ འབྲས་ལོངས་ both in Tibetan and in Lhoke, which means “the land of rice.” However ‘Dränjong’ is probably a shortening for ‘Drämjong’ and is thus better translated as “the land of fruits.” It is also often referred to as a བསྒྲ་ཡུལ་‘hidden valley’ in the Buddhist tradition. Sikkim was a Kingdom until 1975 when it became a state of India. Literary Tibetan was the official language
of Sikkim since the creation of the Namgyal dynasty and remained as such until 1977, when Lhoke was introduced in the schools.

The descendants of the Tibetans, called བློ་པོ་ Lhopos in the local language and བློ་པ་ Lhopas in Tibetan, held a prominent position in Sikkimese society for many centuries. However, the indigenous people of Sikkim are the Lepchas, who are called མོན་པ་ monpa in Lhoke. Since the eighteenth century they have developed their own script distantly related to the Tibetan script. Lepcha is a TB language, but does not belong to the Tibetic family and is very remote from Lhoke. If most toponyms and village names are in Lhoke, some are also in Lepcha language such as Thimjin and Kabi. Lepchas represent 15% percent of Sikkimese. They practice both Shamanism and Buddhism.

Another important ethnic group is the Limbu, who are called གཙོང་པ་ gTsongpa in Lhoke and Tibetan and are also referred to as Subba. They also have an original script, which was invented during the eighteenth century (about the Limbu script, see van Driem 2001 and Plaisier 2008). Limbu is also a TB language, but it is very different from Lhoke.

During the past fourty years, the demographic situation of Sikkim has undergone major changes. There has been an influx of immigrants from Nepal and to a lesser extent from neighboring states of India such as West Bengal and Bihar. The Nepalis now form three-quarters of the Sikkimese population. This massive immigration was due mainly to political and economic reasons: the incorporation of Sikkim into India and the development of modern agriculture.

9.7.2. Linguistic groups of the S section

The S section is made up of several linguistic groups, which allow a rather limited mutual intelligibility. However, the region can be described in terms of a geolinguistic continuum, particularly between the Chumbi valley (Dromo), the Ha valley in Bhutan and the central Sikkim area of Gangtok.

These southern Himalayan dialects are traditionally spoken by cultivators or agropastoralists. In some cases, such as Sāphuk, Mera, Sakteng and Dur in Bhutan, as well
as in Lachen in northern Sikkim, one also encounters small pastoralists’ communities. Some communities such as Dur are now living a sedentary way of life.

For the dialect classification of the S section, we propose the following seven groups:

- Dzongkha རྫོང་ཁ་
- Lhoke ལྷོ་སྐད་
- Cho ça-nga ça རོ་ཅ་ང་ཅ་ཁ་ also called Tsamang-Tsakhaling རུང་ཕྲང་རྒྱལ་ཁང་
- Brokpa (Mera Sakteng pastoralists’ dialect) རུང་ཕྲང་མེ་རེ་སྟེང་གི་འབྲོག་པའི་སྐད་
- Dur pastoralists’ dialect དུར་འབྲོག་སྐད་
- Lakha ལ་ཁ་ or Säphuk pastoralists’ dialect ལྲས་ཕུག་འབྲོག་སྐད་
- Dromo གྲོ་མོའི་སྐད་

The Dzongkha language has several dialects, but an inner classification still needs to be done.

According to Driem (1998), they include the Há dialect spoken in Há ཝཱ་ district and the ‘standard dialect’. As noted by van Driem (1998: 5), the dialect of Há is more similar to the Dränjong dialects, spoken in Sikkim. Additionally, three varieties of Dzongkha are spoken by yakherd communities in northwest Bhutan: Laya ལ་ཡ་, Lingzhi ལིང་བཞི and Lungnagnak /Lunana/ རུང་ནག་ནག་. These varieties have so far not been studied.


The pastoralists of Merak (locally pronounced Mera) and Sagteng སག་སྟེང་ call their dialect རུང་ཕྲང་པ་ ’Drogpa-kā’, which simply means ‘Pastoralists’ dialect’. Quoting Roerich (1961), van Driem reports that “the language of the Brokpa
[of Merak and Sagteng] is an archaic dialect preserving many ancient phonetic and lexical features of Old Tibetan” (1998: 15). The pastoralists of Dur call their language ‘Drog-kä’ locally pronounced /Brokkat/, which also means ‘pastoralists’ dialect’. The community of cattle breeders located in Saphuk calls their dialect ‘Lakha’ ལ་ཁ་, lit. the ‘language of mountain passes’. In Dzongkha, this language is known as Tshangkha གཞིང་ཁ་. There is little available information and a general survey is needed about these three pastoralist dialects.

Lhoke has two very distinctive dialects (Yliniemi 2019): the Northern dialect of Lachen-Lachung བེ་སྦྱེས་/ཕར་རི་, which includes communities of cattle breeders, and the southern varieties spoken elsewhere in South, West and East districts.

The Dromo dialect གྲོ་མོ་སྐད་ is divided into two main subdialects: གྲོ་མོ་སྟོད་ Upper Dromo and གྲོ་མོ་སྨད་ Lower Dromo. Upper Dromo is more similar to the Dzongkha dialect of Há, and lower Dromo, which includes the Chumbi valley, is more similar to Drengjong. As a local proverb says: “The people of Dromo and Há are one” གྲོ་ཧ་རྩ་གཅིག. Two varieties are spoken in Phagriཕག་རི་and Sharsingma ཤར་གསིང་མ་. These two places were trading centers and Sharsingma was under British control in the beginning of the twentieth century. It was handed over to the Chinese in 1954.

Several dialects of the S section are in contact with a few Non-Tibetic languages either Tibeto-Burman such as Lepcha, Limbu, Bumthang, Kheng, Tshangla, Dzala, Kurtö, Chali or Indic languages such as Nepali and Assamese. There is a rather limited intelligibility between Dzongkha and Lhoke, the two major languages of the S section.

**9.7.3. Geographic extent of the S section**

The area of the S section is located on the southern slopes of the Himalayas. It extends over Bhutan, which is called འབྲུག་ཡུལ་ Drug-yül i.e. the ‘Land of the Dragon’ and over the Indian state of Sikkim called འབྲས་ལྗོངས་ Dränjong as well as Dromo County གྲོ་མོ་རྫོང་ in the Tibet Autonomous Region (China) especially the Chumbi valley ཆུ་འབི་ལུང་. Additionnally one should mention the two towns of Dorjeling རྡོ་རྗེ་གླིང་, known abroad as Darjeeling (famous for its tea), and Kalönbug སྦུག་ in the Tibetan Autonomous Region (China) especially the Chumbi valley ཆུ་འབི་ལུང་.
Kalimpong, both located in the Indian State of West Bengal. Historically, many Dzongkha and Lhoke speakers have settled in these two towns.

In Western Bhutan, the largest rivers are: The Wang-chu river འབང་ཆུ (upper course of the Raidak river, a tributary of the Brahmaputra), which flows through Thimphu, the Paro-chu river སྤ་རོ་ཆུ, which runs in the eponym district. The Mo-chu river མོ་ཆུ (upper course of the Sankosh), a tributary of the Brahmaputra, which runs near Punakha. Another significant valley from the economic and historic point of view is the Há valley eponym of the district བཱ་རྫོང་.

In Eastern Bhutan, the main valleys are those of the Amo-chu river ཨ་མོ་ཆུ (upper course of the Torsa in West Bengal), the Bumthang-chu སྤུ་ཐང་ཆུ which runs through the eponym district of Bumthang སྤུ་ཐང་རྫོང་ as well as the Mangde-chu river གང་སྙེ་ཆུ and the Drangme-chu དྲང་མེད་ཆུ, which is the biggest river in eastern Bhutan.

In Sikkim, the main valley is formed by the Teesta River, which runs through the entire state and is a tributary of the Brahmaputra.

Finally, the S section extends to the Chumbi valley located in Dromo County, Tibet Autonomous Region (China). Although, the county is located in Tibet, the valley belongs to the geographic environment of the southern Himalayas and borders with both Sikkim and the Há valley of Bhutan.

Detailed location of the dialect groups

• Dzongkha རྫོང་ཁ

Dzongkha is traditionally spoken Western Bhutan in Thimphu ཐིམ་ཕུ, Punakha སྤུ་ན་ཁ, Paro སྤ་རོ (sometimes spelled with the conservative orthography སྤ་གྲོ), Wangdi Phodrang དབང་འདུས་ཕོ་བྲང་ (Wangdi Phodra/), Garsa བདག་ས་, Darkarnang སར་དཀར་ནང་ (Dagana/) and Chukha རུ་ཁ་ districts. As the national language of Bhutan, Dzongkha is also spoken as a second language in the other districts of the country. Dzongkha includes the specific varieties of Há མ་ and the pastoralists’ varieties of Lingzhi ཞིང་གཞི (located in the eponym Geok of Thimphu district), Laya རྱ་ཡག་, and Lunana དུན་ནག་ (located in the eponym Geoks of Garsa County)
• Choča-ngāča

Choča-ngāča language is spoken in eastern Bhutan in Mongar and Lhüntse districts but some speakers are also found in the neighboring districts of Trashigang and Trashi Yangtse. The two main villages of Tsamang and Tsakaling are located in Mongar district. In Mongar, Choča-ngāča is spoken in Tsakaling, Tsamang, Sheri Muhung, and Saling Geoks. In Lhüntse, it is spoken in Jare, Minje, Menbi, Tsankhar and Metsho Ungar geoks, and in Trashi Yangtse it is spoken in Tongshang and Gangkhapa geoks. In Trashigang, speakers of Choča-ngāča are found in Bartsham, although according to local history Bartsham speakers originally came from Tsamang village. Due to the Bhutanese resettlement policy, speakers of Choča-ngāča can now also be found in the southern districts of Samtse and Sarpang.

• Lhoke

Lhoke also called Bhutia or Drānjong-kā (alt. Denjongke) is spoken in the following districts and villages:

**East district:** Gangtok, Pendrom, Tadong, Tathangchen, Matsong (alt. Machong), Sang-Martham, Khamdong, Simik, Dikling, Pakyong, Pathing, Rumtek, and Asam Lingdze.

**South district:** Raling, Kyozhing (Sosing), Doling, Ben, Yanggang, Sangno, Lingmo, Rabang, Namtsi (Namchi), Tinggitam and Barphung Lingtam.

**West district:** Trashiding (alt. Tashiding), Lingchum (alt. Lingchyum), Singgyang, Yuksam, Drubde, Chumbung, Paling, Gyelshing (alt. Gezing), Nāthang, Tingling, Dorap (alt. Darap), Zilnen (alt. Sinon) and Ganggyap.

---

42. The term ‘Geok’ refers to a block or subdivision of a district.
43. Namtsi is the main town of South district but there are relatively few Lhoke speakers.
44. Ganggyap has a special dialect. A Kham community has settled in this area.
North district: Lachung ལ་ཆུང་, Lachen ལ་ཆེན་, Phodong ཕོ་མདོང་, Phänsang ར་སང་, Kabi ཕེ་བི་, Thingim སྤྱིར་ིས་, Namok ར་མོར་ and Manggän (alt. Mangan) ཟར་གན་.*

• The Pastoralist Tibetic languages of Bhutan

Lakha ལ་ཁ་ is spoken in the area of Säphuk Geok ས་ཕུག་ in Wangdü Phodrang district སང་འདུས་ཕོ་བརང་ (/Wangdi Phodra/). According to van Driem (1998: 16):

“[it] is situated in Wangdi Phodra, a district in the north of the Black mountains, south of the lofty white peaks of Gangs dkar kun bzang [གངས་དཀར་ཀུན་བཟང་] which separate Bhutan from Tibet. […] Lakha speaking villages of Sāphu are B’uso, Langbji, Brabrack, Dzêri, Dârilo, Wangdiqom, Rabu, Kumbu, Bati, Nākha, Sekta and Thanyä.”

Mera Sakteng Brokpa-ke བྲོག་པའི་སྐད་ is spoken in Trashigang district བཀྲ་ཤིས་སྒང་ in and around the two villages of Merak (/Mera/) and Sagteng. D’ur Brokkat དུར་འབྲོག་ is spoken in the Bumthang district བུམ་ཐང་ in the village of D’ur.

• Dromo གྲོ་མོ་དཔོན་

Dromo County is located in the south of Tibet Autonomous Region (China), at the border with both Bhutan, in the east, and Sikkim (India) in the west. It is divided into two dialects: upper Dromo and lower Dromo.

9.7.4. Number of speakers

Dzongkha has about 160,000 native speakers in Bhutan (van Driem 1998). However, since Dzongkha is the official language of Bhutan, a lot of Bhutanese citizens speak Dzongkha as a second language and the number is growing. The total number of speakers (as a first or second language) may be more than 500,000. Additionally, about 10,000 Dzongkha speakers are also found outside Bhutan, in India, Nepal, etc. The second Tibetic language of Bhutan is Tsamang has about 30,000 speakers (according to SIL 2011).

The number of speakers for the other Tibetic languages of Bhutan is much more limited. According to van Driem, Lakha has about 8,000 speakers, for Dur and Mera, there are respectively about 500 and 3,000 speakers.

45. Manggen is mainly a Lepcha speaking area.
Concerning the number of speakers in Sikkim, it is difficult to establish precisely. According to Turin (2011), "the population census of 2001 records Sikkim as being home to only 540,000 residents, of which the indigenous Lepcha and Bhutia make up only a tenth each." Thus the number of Lhoke speakers would amount to around 50,000 speakers. Ethnologue mentions the figure of 70,000 speakers. Turin (2011) adds: "The autochthonous languages of modern Sikkim – Bhutia, Lepcha and Limbu – are at present severely endangered." This statement should be relativized in the case of Lhoke (Dränjong-kä or Bhutia) since the language is taught in schools and has a written standard. Yliniemi (2019) provides a lower estimation for the number of Lhoke speakers: 25–30,000.

Finally, for Dromo, which is the only dialect of the S section spoken on the Tibetan side, the number of speakers is roughly of 6,000 or 7,000 speakers.

Thus the total number of speakers for the S section is less than 300,000 speakers.
9.7.5. Ethnic and sociolinguistic groups

The speakers in the S section are predominantly མིང་པ་ zhingpa ‘cultivators’ and agropastoralists. There are some pastoralists communities of ཁྲོག་པ་ droga, called ཁྲོག་པ། /'djob/ in Dzongkha, in Mera Sakteng, Dur pastoralists’ dialect and Lhaka or Saphuk.

In Bhutan (as mentioned in Chapter 2), there is no generic term to designate the people who speak Tibetic languages as their native tongues, and identity is in general related to the native valley or to the dzong. The term ཉུ་མ་ BRUG /'djob/ in Dzongkha (ཡུ་མ་ BRUG,PA in Tibetan) applies to all the Bhutanese citizens including those who are not native speakers of Dzongkha.

In Sikkim, the Tibetic speaking people call themselves བོ་པོ lhopo which corresponds to Lhopa བོ་པ་ in Classical Tibetan and means ‘southerner’. They also sometimes refer to themselves as Bhutia or bhoṭi (བོ་ཊི) (see 2.5).

9.7.6. Phonological characteristics of the S section

The phonological diversity of the S section is rather limited. However, it is not possible to list common phonological features to the all the dialects of the S section. The phonological characteristics are usually valid only at the level of the groups. Dzongkha དྲོང་ཁ།, Lhoke བྲོད་མོ། and Dromo སྐྱོང་མོ། are in general more innovative than Choča-nça. We lack data about the pastoralists’ dialects which may also be conservative.

Suprasegmental features

The dialects of the S section have distinctive suprasegmental features and generally have two distinctive pitch tones. Additionally, tone contours are found in Dzongkha (see Michailovsky 1986 and Mazaudon and Michailovsky 1989 as well as van Driem 1998).

Segmental features

Synchronic approach

The sound systems of the S section are characterized by the following frequent features:

- The languages of the S section have usually simple initials. Exceptions are found in Dzongkha and Choča-nça. Dzongkha has combinations of a labial and plosive such as /'ch/ and /'j/ as in གྲོ་ བོ་པ། /'djob/. It has also preglottalized
resonant (as in /ˈnam/ ‘sky’). Choča-ngača Tsamang dialect has preserved initial clusters such as /pr, phr, br/ and even /mr/.

- They have voiced non-resonant sounds (b, d, q, dz, j, zh).
- Final /p/, /m/, /ng/ are usually preserved.
- A limited set of vowels which include central rounded vowels /ö/ and /ü/ as well as midhigh unrounded vowel /ä/.
- The vowel length is distinctive.

Diachronic approach and reflexes of Classical Tibetan:

- In the languages of the S section, we find no trace of preradicals in the initial position.
- In Dzongkha, low pitch half-devoiced or breathy sounds [b̥], [g̊], [d̥], etc., are derived from single voiced non-resonant sounds (B, D, G, DZ, J). They are noted in the transcription with a voiced consonant followed by the apostrophe. In Lhoke, they are devoiced.

CT PY, PHY, BY correspond /p̥/̇, /p̊/̇ and /b̥/̇ or /ɡ̊/̇ in Dzongkha and to /p̥/, /p̊/, /by/ or /p̥y/ in Lhoke and Dromo. Ex: གྱུག (PHYUG) ‘rich’; གྱུ། (p̥yuko) (Lho), གྱུ ཁུ (p̊yup) (Dz), ཁུ ‘sand’, ཁུ /p̥yem/ (Lho), ཁུ ‘/p̥yim/ (Dz). Thus, Dzongkha is less conservative than Lhoke for this reflex and presents an intermediate stage between the labial+glide and the affricate as in the Ü-Tsang dialects of the Central section.

CT PR, PHR, BR (with or without preradical) correspond to /k̊/̇, /k̊̇/̇ and /b̥/̇ or /ɡ̊/̇ in Dzongkha and to /k̊/, /k̊̇/, /by/ or /p̥y/ in Lhoke and Dromo. Ex. གྱུག (PHERENG) ‘rosary’; གྱུ ཁུ (p̥yengbu) (Lho), ཁུ /k̊e:m/ (Dz); གྱུ (SBRANG) ‘flying insect’; གྱུ ཁུ /byam/ (Lho), ཁུ /k̊a:m/ (Dz); གྱུ (Brag) ‘rock’; གྱུ ཁུ /p̥ya/ (Lho), /k̊a:/ (Dz).

The reflexes of CT KR, KHR, GR (with or without preradicals) are quite diverse in the S section. For example, in Dzongkha, they correspond to /k̊/, /k̊̇/ and /ɡ/ or /q/; གྱུ (GRL) ‘row’; གྱུ (GYA) /g̊a/, ཁུ (KRID) ‘to lead’; ཁུ (FHID) /k̊i/; or to /ɛ/, /ɛ́/ and /j/ or /j/; ཁུ (GRO) ‘to go’; ཁུ (GYU
PART 2 – CHAP 9. Inner classification of the Tibetic languages

/ˈuju/, གྲང་མོ GRANG.MO ‘cold’; ʃ གྱམ་ GYAM /ʃaːm/; but also sometimes to /ˈaː/; /ˈjʊ/ /ˈd/ or /ˈd/': ʃ བཀྲོང་ BKROWN ‘to kill (H)’; /ˈtong/, འཁྲོ་ KHAG ‘blood’: /ˈtɒː/ (see van Driem 1998). In Lhoke, the reflexes are /ˈky/, /ˈky’/, /ˈgy/ or /ˈgy’/: འཁྲོ་ KHAG ‘blood’: ʃ བཀྲོང་ GYAG /ˈkyaː/; འཁྲོ་ KHID ‘to lead’: ʃ KHID /ˈkik/, འཁྲོ་ KHAG.MO ‘cold’: ʃ འཁྲོ་ KHAG.DRAG /ˈkyaːnɪʔ/; འཁྲོ་ GRO ‘to go’: ʃ འཁྲོ་ ‘gyu’ /ˈgyu/; and /ˈt/, /ˈd/ or /ˈd/’: འཁྲོ་ KHAB ‘to act’: /ˈtap/, འཁྲོ་ GRONG ‘to die (H)’: /ˈtɒːn/. (See Yliniemi 2019.)

• Prenasalization is not found in the S section.

• The following final consonants B, M, NG are well preserved. L, R, S and D do not normally appear.

The elision of the final vowel is a frequent morphophonological phenomenon attested in Dzongkha and to a lesser extent in Lhoke. For example, see above the words; ཕྲ་ བྲང་མ་ SBRANG.MA ‘flying insect’ and བྲ་ བྲང་ MA ‘sand’.

• The combination ZL yields: /d/: ཁྲ་ ZLA.BA /dau/ ‘moon’ (Dzongkha).

9.7.7. Grammatical characteristics of the S section

The dialects of the S section exhibit some grammatical differences particularly in their verbal and nominal morphologies. For example, the grammatical case systems of Dzongkha, Lhoke and Choa-nga differ both in number of cases, forms and functions. In the languages of this section the demonstratives are placed before the head noun.

It is worth noting that the languages of this section have lost the verbal inflections found in CT. Dzongkha has innovated an inflectional system of the stem which is not found in the other languages of the S section such as Lhoke or Choa-nga.

9.7.7.1. Case markers

The languages and dialects of S section usually distinguish at least seven case markers but if one takes into account grammaticalized relator nouns, we may consider that there are up to ten cases. Frequent cases include ergative, absolutive, dative, genitive, comparative, ablative and associative.

The ergative case is found in all the dialect groups or languages of the S section. It is used with both controllable and non-controllable verbs and in some cases also occurs
with intransitive controllable verbs. In Dzongkha it is syntactically optional. As van Driem (1998: 193) pointed out:

“The ergative suffix is in most cases homophonous with the genitive ending, they are neither formally nor semantically identical. [...] the Dzongkha ergative differ from a classical ergative, which marks the subject or agent of a transitive verb, in the Dzongkha ergative is used to highlight the agentive character of a subject which performs an activity, transitive or intransitive.”

What has been said about Dzongkha is also valid for Choča-ngača (Tournadre & Karma Rigzin 2015).

Both Dzongkha and Choča-ngača have preserved reflexes of the archaic comparative marker BAS. Outside the S section, this marker is not attested in the modern Tibetic languages except for some dialects of Kham such as Chamdo and Gyälthang.

The case systems of the three major languages of the S section, Dzongkha, Lhoke and Choča-ngača, are listed below:

Dzongkha has the following case markers: the ergative /-ki/ གིས་ GIS, the dative /-lu/ ལུ་ LU (< CT ལ་ LA), the absolutive Ø, the genitive /-ki/ གི་ GI, ablative /-lä/ བས་ LAS, the locative /-na/ ཉ་ NA, the comparative /-wa/ བ་ BA (derived from CT བས་ BAS), the associative /-tang/ དང་ DANG. Additionally, one finds three cases which correspond to the grammaticalization of relator nouns: the inessive ཉང་ NANG /-na:/ ’in’, the superessive གཁར་ KHAR /-kha:/ ’on’, ◊ ཉགུ་ GU /-gu/ ’on’. However, these forms are not entirely identical to the inherited grammatical cases because they still trigger a genitive case (when the word ends in a vowel).

Choča-ngača has seven cases: absolutive (Ø), ergative /gi/ and the allmorphs /-ki/ and /k/ (< CT གིས་ / གིས་ GIS~KYIS), genitive /-gi/ and /-yi/ (< CT གི་ / གི་ GI~YI), dative /-le/ and allmorphs /-nge/, /-ge/, /-e/ (< CT ལ་ LA), associative /dang/ (< CT དང་ DANG), ablative /leki/ and its allmorphs (may be derived from *LA-GIS ལ་གིས་), comparative /-wata/ (< CT བས་ BAS followed by a form /ta/ of unclear origin) (Tournadre & Karma Rigzin 2015).

46. Van Driem (1998: 208) provides one example of ergative used with the verb མི་འུར་ ’to fly’.
Lhoke has seven cases: the agentive གིས /-ki/, the dative ལོ /-lo/, the absolutive ན /, the genitive གི /-ki/, the locative ན /-na/, ablative ལིས /-le/ and the adessive /-tsa/ (< CT རི་ མི་). This latter case is an innovation of Lhoke.

9.7.7.2. Nominalizers

Various nominalizers (see 8.3.13) are found in the S section but some of these markers are specific for a given region of the section. The main nominalizers of the S section include the following markers:

- The nominalizer /-kän/ or /-k’an/ derived from CT མཁན་ MKHAN is used in various languages of the S section such as Lhoke and Choča-ngača, however it is not used in Dzongkha.
- The nominalizer /-mi/ derived from CT མི་ MI ‘person’ is used in Dzongkha (but not in Choča-ngača nor Lhoke).
- The nominalizer མི་ /-ni/ is very frequent in Dzongkha. It is also probably derived from CT མི་ MI ‘person’ and cognate with the Amdo nominalizer /nə/. It is used as an ‘infinitive’ marker but has other functions such as instrument of the verbal action. In its instrumental and infinitive functions, it corresponds respectively to མི་ /shäʔ/ (in Lhoke) and མི་ /sang/ (Cho).
- The nominalizer མི་ /shäʔ/ is used in Lhoke. It is derived from CT ཆིས CHAS ‘thing’ which is grammaticalized as a nominalizer in many western languages as /ches/ (La) /yä’/ (Ts) and /ya’/ in Central Tibet.
- The nominalizer མི་ /sang/ is used in Choča-ngača. Its origin is unclear.
- The nominalizer /-wa/, derived from CT བ་ BA, is attested in the languages of the S section such as Dzongkha, Lhoke and Choča-ngača. The form /wa/ has two allomorphs: /ma/ and /pa/. The nominalizer /wa/ is used among other things to as a supine marker. When the form /wa/ or its allomorphs are followed by a genitive, they are respectively pronounced as /wä/, /bi/ and /mi/.
- The form བ /po/ used in Lhoke is cognate to the above nominalizer /wa/. It

47. Tournadre’s unpublished manuscript from a fieldwork in August 2010 at the Namgyal Institute, Gangtok (Sikkim). See also Yliniemi (2015).
is derived from the CT form "PA.

- The nominalizer "sa/-sa/ is used in the languages of the section. In Lhoke, it is pronounced -so/. This nominalizer is derived from CT "sa, which means 'place' has various meaning 'the place of the verbal action'.

- The nominalizer /-tang/ derived from the noun STANGS 'way, manner, appearance, look' is attested the various languages of the S section. In Dzongkha and Choča-ngača, it is pronounced /t'ang/ STANGS. Note that /thang/ is aspirated in both languages, unlike the CT form: STANGS.

9.7.7.3. *Verbal inflections*

The inflectional forms inherited from CT or OT have generally been lost in the languages of the S section such as Dzongkha, Lhoke and Choča-ngača. However, some traces are still found. For example, there is a lengthening of the vowel in the past in Choča-ngača: /sa/ (present) < CT ZA versus /'sa/: 'to eat' (past) < CT BZAS, /'ngu/ 'to cry' < CT NGU versus /'ngu:/ 'to cry' (past) < CT NGUS. (See Tournadre & Karma Rigzin 2015.)

Dzongkha has developed innovative morphophonemic variations of the verb stem which depends on the phonological context. These stem variations depend on the type of auxiliary verb / suffix that follows the verb (see Chapter 8). For example, the forms /'-tang/ बट्टा 'to send' and /'-nang/ ग्नां 'to give' (both verbs are also used as light verbs) become respectively /'-tām/' and /'-nām/' when they are followed by some auxiliaries such as /'ing/, the auxiliary marking the 'factual present' (see van Driem 1998: 210).

Verb suppletion is also attested in the S section. For example in Dzongkha, the verb 'to go' has three forms: /'jo/ ग्त्रो (present), /'-song/ ग्त्रिं 'past' and /'-ya:/ ग्त्रा (past used with the secondary verb /so(ng)/), derived respectively from CT ग्त्रो 'GRO' 'to go' (pres), ग्त्रिं SONG 'to go' (past, imp.), and ग्त्रा YAR 'to split up, to be scattered, dispersed'; the verb 'to come' has also three forms /ong/, /'-thon/', /'-sho:/ derived respectively from CT ग्त्रो 'ONG 'to come', ग्त्रिं 'THON' 'to come out', ग्त्रिं GSHEGS 'to come' (see van Driem 1998: 280).
9.7.7.4. Linking verbs and auxiliary verbs

Copulative verbs

As in most Tibetic languages, the copulative verb of the S section is derived from the verb ཡིན་/yin/. It is pronounced /‘yin/ in Dzongkha and /‘in/ in Lhoke. In both Dzongkha and Lhoke orthography, it is written བོད་/bod/. The same copulative verb ཡིན་/yin/ or བོད་/bod/ is also used in Choča-ngača. The values of this copula may however vary in a subtle way between the various languages. In Dzongkha, Choča-ngača and Lhoke, this copula conveys essentially an authoritative or loose egophoric meaning and is perfectly compatible with the third person in declarative sentences.

The negation of this copulative verb is མིན་/min/ in CT, but the vowel is often pronounced as a /ä/ or /e/, sometimes nasalized. In Dzongkha, it is often written བོད་/bod/ and pronounced /mä/. In Lhoke, it is written བོད་/bod/ (Yliniemi 2017; 2019).

For the factual meaning one finds in Lhoke the copulative verb /‘bä/ written བོད་/bod/, which is probably cognate with the Tsang and Lhokha dialects copula /ba/ or its variants (/bä/, /bo/, etc., see above the C section and 8.3.3). This form is neither found in Dzongkha (but occurs in compound forms) nor in Choča-ngača. The negation is /‘memba/ written བོད་/bod/, Yliniemi notes that this form “is a generally asserting, evidentially neutral copula, it implies neither the personalness [egophoricity, conveyed by བོད་/bod/ and བོད་/bod/] nor the sensorialness [conveyed by བོད་/bod/]”.

Existential verbs

The old copula བོད་/yod/ is found in all the languages of the S section. It is pronounced in various ways: /‘yod/ in Dzongkha, /‘yöd/ in Lhoke and /‘yöd/ or /yet/ in Choča-ngača. In Dzongkha and Lhoke, the existential copula བོད་/yod/ conveys a loose egophoric meaning while in Choča-ngača, it has a sensory value (!), which is normally conveyed by the marker འདུག་/du/ in many Tibetic languages. This is a very original feature of Choča-ngača, without equivalent in other languages of the family (except Amdo that has བོད་/yod/ YOD.G1). In order to convey the value of the authoritative, the form བོད་/yod/ is used in Choča-ngača (see below).
The negative form of འདུག་ YOD is written མི་འདུག་ MI-’DUG in Dzongkha and Lhoke. It is pronounced respectively: /ˊme:/ and /ˊmeʔ/. In Choča-ngača, the negative existential verb is /met/.

In Dzongkha and Lhoke, the sensory forms are འདུག་ /du:/ or /duʔ/, which correspond to the reflexes of CT འདུག་ /DUG. The negative form is respectively མི་འདུག་ MI-’DUG /ˊminu/ sometimes realised as མི་འནུག་ /ˊminu/.

In Choča-ngača, the auxiliary འདུག་ ’DUG is attested but it does not function as an existential verb (see Tournadre & Karma Rigzin 2015).

**Compound linking verbs**

The combinations of various auxiliaries are frequently attested in the languages of S section.

Compound linking verbs in Dzongkha conveying evidential and epistemic meaning include: ◊ བིན་པས། ’PAS, also written ◊ བིན་མས། ’MAS /’immä/ a ‘sensory inferential’. This form has been described by van Driem (1998): “the speaker states an observed phenomenon which as such belongs to the realm of his recently acquired knowledge” (see also Tournadre 2017). The origin of the second syllable /’mas/ is unclear but could be related to the Lhoke form ◊ བིན’ ’SBAD. The epistemic compound linking verbs include the equative form ◊ བིནམ་འོང་ /’im-’ong/ and the existential form ◊ འདུགཔ་འོང་ /’yöp-’ong/.

Other frequent compound evidential forms such as ◊ འདུགཔ་བིན་ /’yöp-’ing/ and ◊ འདུགཔ་བིན་ ’MAS /’yöp-’immä/ are also attested in spoken and more frequently in written Dzongkha but they have not received much attention.

In Choča-ngača, several compound linking verbs are also attested: ◊ བིན་ཅེད་ /’yinčet/, which is a sensory inferential corresponding to /’immä/ in Dzongkha may be derived from CT བིན་ ’RGYU YOD. Other forms include the epistemic markers ◊ བིན་པི་འོང་ /’inpiong/ and ◊ འདུགཔ་པི་འོང་ /’yotpiong/.

Lhoke has also developed a set of compound equative copulas: The inferential equative བིན་ ’DUG the past factual equative བིན་ ’SBAD, as well as the compound existential copulas: the ‘factual existential’ (called ‘neutral’ by
Yliniemi 2017; 2019), ཡོད་པོ་སྦད་ / yöpobäʔ/ or its spoken equivalent ཡོད་བ་སྦད་ / yöbäʔ/ the ‘past factual existential’ ཡོད་པོ་ཨིན། / yöpoin/.

**Auxiliaries**

The main auxiliary verbs of the S section are made of the following verbs (sometimes preceded by a relator). Some of these verbs are also used as copulative verbs.

- ཨིན་ / in or YIN (Dz, Lh, Cho) also used as the copulative verb ‘to be’. The auxiliary is used in various constructions for the past, present and future.
- སྦད་ / sbad (Lho) also used as the copulative verb ‘to be’. The auxiliary is used in various constructions for the past, present and future.
- མས་ / mas / mä/ ‘sensory evidential’ (Dz). This marker is used for the present and uncompleted aspect with verbs, as well as predicative adjectives in the short form (ex. རིག་པ་ / rhig-pa ‘it is tasty’). With predicative adjectives in the long form (or invariable) the copula དུག་ / dug may also be used. རིག་པ་ / rhig-pa / ZHIM-TOG TO / DUG ‘it is tasty.’ With the verb in the present, only MAS / mä/ is possible: རིག་པ་ / rhig-pa / DRAIN-MAS (I) miss (you), but རིག་པ་ / rhig-pa / DUG ‘DUG is not attested.
With adjectives, when the two auxiliaries/endings are possible, the semantic difference between the two usages MAS / mä/ and ‘DUG / dug’ is subtle. It seems that MAS is better suited for the endopathic meaning. The origin of MAS / mä/ is unclear. A possible hypothesis is that the form /mä/ comes from the CT nominalizer PA, BA which is originally followed by the auxiliary ‘DUG’ (ex. ZHIM.PO ‘DUG) and that the latter was subsequently dropped.

- དོ་ / do/ ‘present ‘participatory sensory’. This auxiliary is attested both in Dzongkha and Chöka-nga as well as in Lhoke (combined with the auxiliary བིན / bin/). It may be used with first person as well as non-first persons. The origin is unclear but it might be derived from CT སྡོད་ཡོད་ / sdod ‘stay’ + YOD copula’.

- དུག་ / dug / du/ is used in Lhoke while རིག་ / nugg / -nu/ occurs in Dzongkha.

48. Auxiliary dropping is attested in several languages. In Lhasa Tibetan and Amdo respectively བིན་ / bin/ NGA-S-BRAN-GTI (DUG), བིན་ / bin/ DRAIN-GI. In the case of Lhasa, the auxiliary is usually dropped in affirmative sentences (but not in negative and interrogative). In Amdo, it is less clear whether an auxiliary was dropped. In Hor dialect, the negative copulative verb བིན་ / bin/ MA-BED becomes /ma/.
Both convey the ‘perfect sensory inferential’ and are derived from the CT verb འདུག་’DUG’ ‘to sit’.

• འདུག་ Yi/-yi/ and its allomorph ལེགས་/zhe/ ‘past participatory sensory’ in Dzongkha is probably related to the Lhoke form མཁན་/ce/ and its allomorph རེ་/zhe/, which conveys a similar meaning. Their etymology is not clearly established. The form འདུག་/pl used in Choća-ngača might be also cognate with these forms, and could be derived from CT PA-YIN.

• ལེགས་ SONG (Dz) secondary verb (see 8.3.5) < CT ལེགས་SONG ‘to go’. It is used in past constructions.

• རྡ་ RDA/-da/ secondary verb (Dz) < CT རྡ་BTANG ‘to send’. It is used in past constructions.

• རྡ་ CHI/-chi/ secondary verb (Dz, see 8.3.5) < CT རྡ་MCHIS ‘to exist’. It is used in past constructions. For these three secondary verbs, see Driem (1998) who provides detailed comments and many examples.

• ལེགས་ ‘ONG/-ong’, derived for CT ལེགས་ ‘ONG’ ‘to come’ is used in Dzongkha and Choća-ngača to convey an epistemic meaning.

• Finally, let’s also mention the verb ending འདུག་/-te/. This form is not an auxiliary verb but originally a connective clitic derived from the CT connective འདུག་STE and its allomorphs འདུག་ DI and འདུག་ TE. It plays an important role in the languages of the S section as in many other sections. The marker /-te/ occurs alone in Choća-ngača as a marker of the past tense. In Dzongkha, འདུག་STE /di/ and the 2 written variants (氘 DE འདུག་ TE) pronounced identically are also attested (Driem 1998: 296) to convey the perfect in combination with auxiliaries འདུག་/du:/ or འདུག་/‘ing’. This marker is also attested as འདུག་ STI/-ti/ in Lhoke (see Yliniemi 2017, 2019).

As we have seen above the S evidential/epistemic systems usually have special forms to mark sensory and sensory-participatory access to information, as well as loose egophoric, factual, inferential, epistemic and hearsay meanings. Unlike the languages of the Central and Southeast and Northeastern sections, the dialect groups or languages of the Southern section seem to lack strict egophoric marking. They also generally lack the opposition
between ‘intentional’ (or ‘volitional’) and ‘non-intentional’ (‘non-volitional’) (see van Driem 1998 for Dzongkha and Tournadre & Karma Rgyzin 2015 for Choča-ngača).

9.7.7.5. Negation

The negation has two forms in all the dialects of the S section: མ་ MA is normally used for the past and the imperative whereas མི་ MI used for the present and future. Lhoke has developed an allomorph མན་ MAN (see 8.4.11). The negation used in the various tenses and aspects comes before the lexical verb (when used without auxiliary) or before the auxiliary when it is present.

9.8. The South-western section

The South-western section (henceforth SW section) is made up of several linguistic groups spoken in the northern districts of Nepal and, marginally, in Sikkim (India) and in the Kyirong County of the Tibet Autonomous Region (China). The main groups of dialects include Sherpa ཤར་བའི་གཏམ་སྙད་ /sharwa tamnye/, Kyirong-Yolmo སྐྱིད་རོང་དང་ཡོལ་མོའིསྐད་, Lo གློ་སྐད་ (Mustang) and Dölpo དོལ་པོ་. They form a geolinguistic quasi-continuum in the northern districts of Nepal along the Sino-Nepalese border.

Nepali, an Indo-Aryan language, is the official language in the SW section and the medium in the school curriculum together with English. Thus it has a significant impact on the Tibetic languages of this section. They have integrated a number of loanwords from Nepali. More rarely, they may also borrow words from neighboring Tibeto-Burman languages such as Tamang, Manangi, Seke, Gurung, Thakali, Sunwar, Rai, Kiranti, Limbu, etc.

The languages and dialects of the SW section are not normally written down, although several orthographies have been developed. Sherpa has been written down mainly in Tibetan script and in some cases in Devanāgarī. There are a few publications, textbooks or dictionaries in Sherpa. However, the use of written Sherpa is still marginal and is not available yet on the internet. A few other languages such as Yolmo, Jirel or Lhomi have also developed a written standard in Devanāgarī, but the number of written texts remain very limited (see Hari 2004; Maibaum & Strahm 2005). For written purposes, most speakers of Tibetic languages use Nepali and in some cases English.
Literary Tibetan is still used in the Buddhist and Bönpo monasteries, as well as in the institutes of traditional medicine throughout the SW area.

In Nepal, speakers of Tibetic languages are essentially followers of Vajrayāna Buddhism, particularly of the Nyingmapa and Sakyapa sects. The Nyingmapa sect is dominant among the Sherpas and Yolmo. Sakya is dominant in Mustang. In Dölpo (Dolpa) and, to a lesser extent, in Lo Mönthang (Mustang), one finds significant Bönpo communities (Nagano & Karmay 2004).

Buddhist and Bön monasteries of the SW section include: སྟེང་པོ་ཆེ་དགོན་པ། Tengboche monastery (Nyingma) in Khumbu area, སྤང་པོ་ཆེ་དགོན་པ། Pangboche monastery (Nyingma) in Khumbu, གཞུང་དགོན་པ། Jungbesi monastery or /jung gönpa/ in the Shorong area of Solu; བྱམས་པ་དགོན་པ། Jampa Monastery in Lo Mönthang (Mustang), རྡུལ་གྲོང་གླིང་དགོན། Shel monastery and དཔེ་ལྡན་བསམ་གླིང་དགོན། Dedan Samling monastery in Dölpo (Bön), etc.

Some forms of Shamanism are also attested, for example, among the Jirels (Strahm & Maibaum 2005) and the Helambu Yolmo speakers (Gawne 2013: 25). In the case of the Jirels, one should probably define the religious practices as a form of syncretism with Shamanist and Hindu-Buddhist elements. Small Christian communities are also attested in various areas of the SW section.


9.8.1. Migration patterns, legends and historical records

Generally, Tibetic communities of Nepal have migrated from Tibet a couple of centuries ago. The details of these migrations are not known but there are some historical records. In the case of Yolmo, Kagate, Gyäsumdo and Langthang, they probably migrated from the Kyirong region in southwestern Tibet, which is corroborated by the linguistic affiliation. Other communities in northwestern Nepal such as Humla, Karmarong (Mugu), Dolpo and Lo-kä (Mustang) have also probably migrated from Tö Ngari areas. In the case of Sherpa, the situation is slightly more complicated. Sherpa communities migrated in various waves from Kham via Tö in Western Tibet. Sherpa people are said to have four greater clans of Kham origin (see Nishi 1986).

9.8.2. Linguistic groups of the SW section

Languages of the SW section are traditionally spoken by ཞིང་པ་ zhingpa ‘cultivators’ or རོང་མ་འབྲོག་ rongmadrok ‘agropastoralists’. The mutual intelligibility between Sherpa, Jirel, Lo-kä is not good, while it is undoubtedly better between Lo-kä and some dialects such as Yolmo or Langthang. The detailed description and the relationships between the dialect groups of the SW section still need further research. However, one significant geographic element is the absence of transversal roads linking the various communities. Since the Himalayan valleys of Nepal are oriented north-south, the Tibetic groups are generally isolated from each other, and used to be connected with Tibet in the north and with other other TB and Indo-Aryan communities in the south.

For the dialect classification of the SW section, we propose the following nine groups:

- Humla རུམ་ལའི་སྐད་ or Limirong;
- Karmarong (Mugu) བཀར་མརོང་དག;
  - Dolpo དོལ་པོའི་སྐད་ and Tichyurong;
- Lo-Mönthang often called Lokä (Mustang) གློ་སྐད་དམ་ཐང་གི་སྐད་;
- Kyirong-Yolmo མིར་ངོང་བཞིན་ཞིག་;

According to Hovden (2016), Humla originated from ‘OM LA’. The spelling ‘OM LA’ is also attested.
574

- Jirel རིས་ལེགས
- Sherpa ལོ་གོང་ལེགས
- Lhomi ལོ་མི
- Gola གོ་ལེགས

9.8.3. Geographic extent of the SW section

This section is located on the southern slopes of the Nepalese Himalayas in the area of the following mountain ranges: the Jomolangma ཐོ་མོ་གླང་མ, better known in English as Everest and in Nepali as 'Sagarmātha', the Jowo Yu (spelled Cho Oyu) རོ་བོ་གཡུ, the Kanchendzönga (alt. Kanchenjunga) གངས་ཆེན་མཛོད་ལྔ, the Langthang གླང་ཐང, the Dhaulagiri, the Annapurna and the Manaslu ranges.

Among the main valleys and rivers running through the eastern part of SW section, one should mention the Bumchu བུམ་ཆུ (upper course of the Arun river), the Matsang tsangpo མ་གཙང་གཙང་པོ (upper course of the Sun Koshi) and the Rongshar tsangpo རོང་ཤར་གཙང་པོ (upper course of the Bhote Koshi) which are Trans-Himalayan Rivers originating in Tibet, as well as the Dudh Koshi and the Tamur rivers. In their lower courses, they join to form the Koshi River, a large tributary of the Ganges.

In the western part of this section, among the main rivers, one should mention the Lo Mönthang river གློ་སྨོན་ཐང་ (upper course of the Kali Gandaki), which runs through Mustang and its tributary, the 'Kyirong-chu' རིང་ཐོང་ཆུ (Trishuli river). The last major river, close to the western border of Nepal, is the Kamali river which originates from the area of the Mapham Yumtsho མ་ཕམ་ཡུ་མཚོ in Tibet and runs through Humla.

The SW languages and dialects are spoken in all the districts along the Sino-Nepalese border, from West to East: Humla, Mugu, Dolpo, Mustang, Rasuwa, Sindhupalchok, Dolakha, Ramechhap, Solukhumbu, Sankhuwasabha and Taplejung, and to a lesser extent in Gorkha, Manang, Nuwakot, Lamjung. In Kyirong area, the SW section extends to the other side of the border in the Tibet Autonomous Region.

Detailed location of the dialect groups

- Jirel རིས་ལེགས 'JIREL.SKAD is locally called /jirel bat/ ('bat' is a loanword from Nepali). Jirel is spoken primarily in the Jiri and Sikri valleys and a few villages
around in Dolakha District (Janakpur Zone) at an average altitude of 2000m. No significant dialectal variation has been reported. According to Maibaum & Strahm (2005), a “small number of Jirels” inhabits some villages to the north-west of Jiri.

- **Humla** is spoken in Northwestern Nepal along the Karnali river from Hilsa to the Nyin valley east of Simikot and in the Limi valley (GLE-MI) of Humla district. It is also referred to as Limirong.

- **Karmarong** is spoken in the district of Mugum (Mugu).

- **Dölpo** is spoken in the Dölpo district, mainly in four valleys: Tsharka, Tarap, Panzang and Nangkhong. The VCDs where the Dölpo dialect is spoken are Chharka, Mukot Dho, Phoksundo, Saldang and Tinje. A neighboring dialect called Tichyurong has been reported by Honda (2018).

- **Lokä** is spoken in the district of Lo Mönthang or ‘Mustang’ (GLE-MI) is spoken in Upper Mustang, which was traditionally called the Lo Kingdom. Lower Mustang is referred to as Baragaonle (see van Driem 1997: 861).

- **Kyirong-Yolmo**. The ‘Kyirong Yolmo’ group includes the following dialects: Kyirong-mä (about the Lende variety, see Huber 2005), Langthang, Yolmo (GLE-MI), Shupa (GLE-MI), SHOG.PA’SKAD also called Kagate or Dhaibung (both Kagate Bhot in Nepali and SHOG.PA’SKAD in Tibetan mean ‘paper maker’ see van Driem 2001: 864), Gyälsumdo (GLE-MI), Nubri (GLE-MI), Tsum (GLE-MI). These dialects are in the Kyirong County (TAR, China) and in central-northern Nepal, mainly in the Manang, Gorkha, Ramechhap Sindhupalchok and Nuwakot districts.

- **Sherpa** is referred to as ‘Sharpā-ka’ by the Tibetans however the

---

51. Gyälsumdo which is located in the Manang speaking area was earlier classified as a Manangi dialect (see van Driem 2001; It is only spoken by 200 speakers. Hildebrandt & Perry (2011) propose not to group Gyälsumdo together with Kyirong on the basis of a few reflexes from CT (164 lexical items). The authors note that Gyälsumdo has been located amongst Tamangic languages for a long period of time. Gyälsumdo and Nubri “share a large amount of lexemes, but Gyälsumdo has a number of distinguishing phonological features” (Hildebrandt & Perry 2011).
Sherpas call their languageShar wi Tamnye. It has four main dialects: Khumbu, Pharak, Shorong and Rolwaling. It is one of the official languages of Sikkim. The Sherpa dialects are spoken in the Solukhumbu District. Sherpa communities are also reported in the following districts: Dolakha (Chordung-Baramji), Taplejung, east Ramechhap, Okhaldhunga, Khotang, Sangkuwa Shaba (Barun), Bhojpur (Maya Danda), Ilam and Terhatum.

One should also mention a Sherpa “inner diaspora” within Nepal, particularly in the eastern regions and a large community in Kathmandu. There is also a small Sherpa community in Dram (TAR, China) on the other side of the Sino-Nepalese Border, but most speakers have become assimilated and now speak a dialect of To Tibetan spoken in this area. Finally one encounters a Sherpa diaspora in Sikkim and West Bengal (Darjeeling and Kalimpong), as well as in Hong Kong. In Sikkim, the areas inhabited mostly by the Sherpa population are: Ribdi, Bareng, Sangkuwa, Okhrey, Seprey Nagi (Sombarey), Bega and Bermoik in the West district, Bermoik Thangasing, Palk Naya Busty, Danthang Jowbari, Perbing and Soreng in the South, Phademchen, Agamlok, Subaney Dara, Sumin, Dokchen, Thokchey and Yali in the East district and Kabi in the North district. Additionally, a dialect called Naaba has been referred to as ‘Sherpaic’ by Honda (2018). It is spoken in the villages of Kimathanka, Dangok und Pharang in Sankhuwasabha district, not far from the Lhomi speaking area.

The so called ‘Gola group’ refers to Walung-kä and Drogpä-kä spoken respectively in Walunghung Gola (alt. ‘Tokpe Gola’) and Drogpä Gola (alt. ‘Tokpe Gola’). Walunghung Gola are spoken in the north-eastern area west of Nepal near Mt Kanchenjunga. The term SGO LA (lit. ‘door pass’) refers to ‘border (mountain) pass’. Walung literally means ‘fox valley’ whereas Drogpä (alt. ‘Tokpe’) probably derives from CT GROG.PO ‘deep ravine with a torrent’. Both communities are located in the Taplejung District of north-eastern Nepal, in villages along the upper reaches of the Mewa Khola.

52. http://shodhganga.inflibnet.ac.in/bitstream/10603/137600/9/09_chapter_03.pdf
• The Lhomi and Thudam dialects are spoken in Sankhuwasabha district close to the Sino-Nepalese border which separates them from Dingri County in the TAR. Lhomi communities are located on the mountain slopes of the upper Arun River, near the Barun glacier. According to Ethnologue they are spoken in Chepuwa VDC, Chepuwa, Chyamtang, Gumba, Chhumusur, and Rukuma villages; and Hatiya VDC, Hatiya, Hungung, Pharang, Syaksila, Simbung, Namase, and Shiprung villages; the southernmost village is Seksum in Arun valley; Thudam is spoken on a slope on a tributary of the Arun river.

9.8.4. Number of speakers

The 2001 of the Central Bureau of Statistics of Nepal reported that there were 154,622 Sherpa in Nepal, 19,706 of them living in the Solukhumbu area. The 2011 census gives the figure of 112,946. Bradley (1997) estimated the number of Sherpas as 50,000. Graves (2007) gives that the number ranges between 15,000 and 70,000. There is also a Sherpa speaking community in Sikkim (India) with at least 10,000 people.

The reasons for such discrepancies in the figures are due to the fact that some Tibeto-Burman people of Nepal are sometimes assimilated with the Sherpas who have gained international recognition for their mountaineering skills. The sociolinguistic situation is also complex since some of the young Sherpas who live in Kathmandu or outside the Solukhumbu area no longer speak the language. The overall estimation of 50,000 Sherpa speakers seems reasonable.

MAP IX.7. – Linguistic area of SW section
Among the linguistic communities of the SW section, the Sherpa community is the most vigorous and has the highest number of speakers. As mentioned earlier, it has also developed some written materials both in Nepal and India.

The other communities are much smaller, as shown in the figures below, usually a few thousand people each. The numbers below just give a general idea and there are a lot of fluctuations for this data depending on the sources and the census.


Altogether, the languages and dialects of the SW section are spoken by at least one hundred thousand speakers. However, there are few monolinguals and most of these languages and dialects are endangered, with maybe the exception of Sherpa.

9.8.5. Ethnic and Sociolinguistic groups
The speakers in the SW section are predominantly རྒྱུན་པ་ zhingpa 'cultivators' and agropastoralists. There is no generic term to designate all the Tibetic people that speak SW Tibetic languages in northern Nepal. They sometimes refer to themselves as bhoṭi (བྷོ་ཊི), bhoṭe (see Chapter 2) or use the name of their community such as བོད་པ་ 'Sherpa' (locally བོད་/sharwa/), སྤོལ་ 'Yolmo sometimes also called 'Helambu Sherpa', སྤོལ་ 'Lopa' of Mustang, སྤོལ་པ་ 'Dölpa', Jirels, etc.

9.8.6. Phonological characteristics of the SW section
The phonological diversity of the SW section is rather limited. However, it is not possible to list common phonological features to the all the dialects of the S section. The phonological systems of the SW section present some similarities with those of the S section.
Suprasegmental features

All the Tibetic groups of northern Nepal have a pitch tone system usually with a two-way contrast (high and low).

Segmental features

Synchronic approach

The sound systems of the S section are characterized by the following features:

▪ Existence of voiced non-resonant sounds (b, d, ɖ, g, dz, j, z, zh, ñ).
▪ Neither prenasalization nor preaspiration are attested in this section.
▪ The set of final consonants includes n, m, ng, r, p, l, ?, k.
▪ The set of vowels is limited, usually a, i, u, e, o. The vowel length is distinctive.

Diachronic approach and reflexes of Classical Tibetan

▪ In the languages of the SW section, we find no trace of preradicals (in the initial position).
▪ Voiced non-resonant sounds (b, d, ɖ, g, dz, j) are derived from the consonants with all preradicals.
▪ There is no trace of the preradical consonants.
▪ CT final B, G, M, N, NG, L are usually preserved. The CT final t is usually dropped in most dialects but it remains in some dialects such as Jirel and in Sherpa (as a morphophonological alternance).

9.8.7. Grammatical characteristics of the SW section

9.8.7.1. Case markers

In the SW section, the grammatical cases of the languages and dialects include the ergative, the absolutive, the dative, the genitive, the ablative and the comitative.

For Kyirong, Huber (2002) provides the following list: ergative /ge/ (< གིས་GIS), absolutive (ø), genitive /-ge/ and its allomorphs (< CT ལི GI), dative /la/ derived from the CT བོ LA, ablative /-lä/ དབོ LAS.

Concerning Yolmo cases, Hari mentions apart from the usual cases similar to Kyirong the form /-ti/ པི which she labels "attributive" and function as a dative-
aesthetive marker of the experiencer (Hari 2004: 408, 774). This marker is also found in Jirel as /-te/ and also called “attributive” (Strahm & Maibaum 2005). The origin of this marker is unclear. Additionally in Yolmo, the ablative is either /-le/ or /-legi/ (Hari 2004: 242; Gawne 2016: 100).

For Sherpa, we have a similar system: ergative /gi/ and its allomorphs (< CT GIS), absolutive (Ø), genitive /gi/ (< CT GI), dative /la/ (< CT LA) and ablative /ne/ or /no/ (< CT NAS), (Grave 2005; Tournadre et al. 2008). One should also note the innovation of two markers /-ma/ and /-sur/ (respectively from CT MAR ‘down’ and TSHUR ‘hither’), which are suffixed to the ablative /no/.

As in many languages, the dative /la/ encodes in Sherpa the beneficiary, the marked patient and the “subject” of possessive constructions but it also indicates the sensory ‘ceptor’, i.e. the “subject” of perception verbs (ex. be cold) and the experiencer (ex. be happy).

For example /ngā-la-ga-wā/ “I am happy” (Grave 2005: 123). It is interesting to note that emotion verbs require in Sherpa the reverse treatment of Ü-Tsang, i.e. the experiencer (the “subject”) is marked by a case /-la/ and the stimulus (“the object”) is in the absolutive (Ø), the Ex. /ngā-la-lu-nyen-u-ga-we/ ‘I like to listen to music’ (ibid.: 155).

The Jirel case system, described by Strahm & Maibaum (2005) includes the ergative /-ki/ and its variants (/-gi/, /-i/, /-iki/) < CT GIS, the comitative /-tang/ < CT DANG, the dative /-la/ < CT LA, the locative /tu/ and its variants /du/ derived from CT purposive /DU/, the locative /-ne/ derived from CT locative NA as well as two innovative allative case markers, /-pa/, and /-ma/ conveying the specific meaning respectively of “away, at a distance” and “at a lower location.” These two markers are derived from the CT adverbs MAR “away” and MAR “down” (We adapted here Strahm & Maibaum’s transcription and provided the CT etymologies).

54. This function is also called aesthetive, see 8.1.9.
9.8.7.2. **Nominalizers**

Various nominalizers (see 8.3.13) are found in the Sherpa, and other languages along the border between Nepal and the Tibet Autonomous Region. They include the following markers:

- The nominalizer /-pa/, /-ba/ (in Yolmo and Jirel) or /p/ and the allomorph /u/ (in Sherpa) which is found in the languages of the SW section, is derived from CT ST PA/ ST BA is used in of the SW.

- The nominalizer /-ken/ or its variants /gen/ /-kan/ or /-kandi/, derived from CT SA MKHAN, is used in various languages of the SW section such as Kyirong/Yolmo and Jirel. (see Huber 2002; Gawne 2013)

- The nominalizer /-sa/, derived from CT ST SA 'place'.

- The nominalizer /-če/ is used in Kyirong and Yolmo (Huber 2002; Gawne 2013). It is probably derived from the CT SA CHAS. Similar forms are also attested in the languages of various sections (NW, W, S, C). The nominalizer usually corresponds to the CT nominalizer ST RGYU.

9.8.7.3. **Verbal inflections**

The inflectional forms inherited from CT or OT have been partly preserved in the languages of the SW section such as Sherpa, Yolmo and Kyirong. Some rare verbs still exhibit three distinctive forms for the present, the past and the imperative (see Grave 2007; Tournadre e.g. 2015). For example the Sherpa verb ‘take’ exhibits three stems: /ling/ (present-future), /la:/ (past), /lo:/ (imperative), and the verb ‘to eat’ has two stems: /sa/ (present-future) and /so:/ (past-imperative) (Grave 2007; Tournadre et al. 2015).

Concerning the Kyirong Lende dialect, Huber provides the following comment:

“...A verb can distinguish at most three stems [...] Stem alternations can consist of a change in vowel quality or vowel length, or in both. However only a few verbs make a morphological distinction between all three stems; most stems have only two forms, or even one.” (Huber 2002: 121)

Additionally, just as Dzongkha, Sherpa has developed innovative morphophonemic variations of the verb stem which depends on the phonological context. These stem variations depend on the type of auxiliary verb/suffix following the verb. They involve
the loss of the stem final consonant and the change of the vowel (see Chapter 8, section 8.3.7).

Some languages of the SW section use suppletive verbs to form various tense-aspects and modalities. For example, some frequent Sherpa verbs exhibit suppletion: to give /ter/ (present) versus /bin/ (past) derived respectively from CT तेर्न STER and बिन SBIN; /do/ (present) versus /gal/ (past) and /gyuk/ (imperative), derived respectively from CT गुर्न GRO ‘to go’, बर्ग र गु ‘to cross’ and र्गु ‘to run’: /hong/ versus /shok/ derived respectively from CT ओंग ONG ‘to come’ and ग्स ‘to cross’ (see Grave 2007). In Yolmo, suppletion is also attested. For example तेर STER /ter/ ‘to give’ has two imperatives लाङ्ग SLANG /lang/ ‘give to me/us’ versus तेर STER /ter/ ‘Give to him/her/them’ (see Hari & Chhegu Lama 2004: 835).

9.8.7.4. Linking verbs and auxiliary verbs

Copulative verbs

As in most Tibetic languages, the main copulative verb in the SW section is derived from CT यिन YIN. In Sherpa, it is realized as ए हिन hin/ but often written ए (see Graves 2007), यिन YIN /yi/ in Kyirong (Huber 2002) and यिन पा yinpa/ in Yolmo (Gawne, sketch). The copula यिन YIN has the following negations: मिन min/ in Sherpa and Yolmo, ए मांण्ड MAN /mä:/ in Kyirong. The reflexes of यिन YIN in Kyirong or Yolmo convey an authoritative meaning and not a strict egophoric meaning unlike its equivalent in Ü-Tsang, Kham and Amdo. They naturally combine in declarative sentences with third person and second person, aside from the first person.

In Kyirong the factual form is ए तिन म्षन YIN.MKHAN /yingː/ whereas in Khumbu Sherpa, it corresponds to ए न्द्वृ न्द्वृ DZA /hindza/ (this form is not mentioned by Graves 2007 in his description of Hile Sherpa).

Yolmo has a general/gnomic and factual copula ए ओंग्गे onge/ derived from the CT verb ओंग ONG ‘to come’ followed by the suffix /ge/, but यिन पा yinpa alone may also convey a general statement (see Gawne 2017).
The copulative verbs ◊ རག་/rak/\(^{55}\) and ◊ མག་/nak/ are marginally attested in the area. For example it is found in Mustang (see Kretschmar 1995). Concerning the origin of the copula མག་/nak/ see the section on CEV (8.3.3).

**Existential verbs**

Existential auxiliaries of the SW section include the following verbs:

- ◊ རེད་/we/,'/wä/ or ◊ སྲིད་/hot/ in Sherpa and /o/ in Dölpo. ◊ སྲིད་/ye/ in Yolmo and /yö:/ in Kyirong. All these forms are derived from the CT and OT form འོད་ YOD or the archaic variant འོད་ OD. In Yolmo this auxiliary conveys an authoritative meaning or a loose egophoric (see Gawne 2017), whereas as in Sherpa it has a strict egophoric meaning.

- ◊ རེད་/nok/ in Sherpa, ◊ རིག་/nu:/ in Kyirong འདུག་ /du/ in Yolmo, etc. All these forms derived from the CT verb འདུག་ 'DUG. This copula has the following negation forms: འདུག་/mindu:/ or /miduk/ in Sherpa, /mindu:/ in Yolmo, etc. These various forms indicate a sensory meaning.

**Compound linking verbs**

Compound verbs are very frequent. They involve the combination of various copulas or auxiliaries. The following combinations are attested in Sherpa:

- རེད་/hinnok/, རེད་/hotuinnok/ which derive from the combination of the existential འོད་ YOD or the equative འདུག་ 'YIN with འདུག་ 'DUG (see Graves 2007).

For Kyirong, we find:

- རིག་/yobayimbä:/, རིག་/yimbanu:/ and རིག་/yimbayö://

This latter form has a 'mnemic' function also attested in Ü-Tsang (Tournadre 1998; 2003). According to Huber (2002): the speaker indicates that he is making an assumption based on his old, personal experience. She illustrated her analysis with the

---

55. The final /k/ is normally realized as a glottal stop.
56. The form /we/ is found both in Hile Sherpa and Khumbu Sherpa (see Grave 2007; Tournadre et al. 2009). The form /hot/ is additionally mentioned by Grave (2007).
following sentence (that we reproduce in transliteration here): KHO DALTA DGERGAIN YIN, PA, YOD, ‘He is probably still a teacher.’

**Auxiliaries**

Frequent auxiliary verbs of the SW section consist of the linking verbs (sometimes preceded by a relator):

- ❁ ལིན་ /win/ or the form བ/, derived from the CT (and OT) copulative verb ིན་ YIN are used to indicate the egophoric past and the future in Sherpa; the form /-in/ is attested in Jirel (Strahm & Maibaum 2005). This auxiliary is replaced in Kyirong by the relator ལོགས་/ge-./

- ❁ ལེད་ /wä/, /we'/ in Sherpa, དེ བོ /yö/: in Kyirong, དོ བ /-ot/ or /-o/ in Jirel and དོ བ /-ye/ in Yolmo are all derived from the existential verb དོ YOD or its archaic form བ/ OD. They indicate the authoritative present and uncompleted past (together with a relator) as well as the perfect (in Kyirong and Yolmo). According to Gawne (2016), the form དཔོ /yeke/ is used specifically for the past.

- ❁ ཕ /-dza/ (used in combination with existential and equative verbs ིན་ WEDZDA, ིན་ YIN:ZDA) is used in Khumbu Sherpa to indicate a factual meaning. Its origin is unclear. It could be derived from the verb ིེར་ ZER ‘to say’ or more likely from བོད་/dak/ which is attested in Tö Ngari. Graves (2007) mentions the form བོ /g/ (with various allomorphs) to indicate a similar meaning in Hile Sherpa.

- ❁ བོ /-sung/ in Sherpa and Jirel and བོ /-so/ in Kyirong derived from CT རོང SONG indicates the sensory completed past.

- ❁ བོ /-nok/ (Sherpa) བོ /-nuk/ (Kyirong) and བོ /-du/ (Yolmo) and /-duk/ (Jirel) are derived from CT བོ DUG ‘to sit’ are used for the sensory present and perfect. This auxiliary may also be used in the future and then conveys an epistemic value (Huber 2002: 183).

- The ending /-bo/ or /-bô/ is mentioned by Huber (2002) to convey an ‘ego-receptive’ meaning, i.e. an event or action directed towards the speaker which
is usually involved (Tournadre 1996). It is derived from the CT auxiliary BYUNG’ to come, to appear’.

- བྱུང་/te/ derived from CT BSDAD ‘to sit, stay’ is used in Yolmo to indicate the durative aspect.
- ཡིན་/sin/ derived from CT ZIN’ to finish’ is used in Yolmo to indicate a past tense (see Gawne 2016).
- རྒྱུ་/dro/ derived from CT GRO ‘to go’ conveys an epistemic value in Kyirong, Yolmo (Gawne 2016).

Additionally some markers functions as verbal suffixes to indicate TAME meanings but are not derived from auxiliary verbs. They include:

- /-te/ derived from the CT connective STE/TE is used in Kyirong to indicate a generic or factual past (see Huber 2002). The suffix /-te/ is also used with a similar meaning in some languages of the S section such as Choća-ngaća.
- /-pa/ derived from the CT nominalizer P.A occurs in Kyirong to convey the intentional past (see Huber 2002: 163).

As we have seen above the SW evidential/epistemic systems usually have special forms to mark sensory access to information, as well as egophoric and/or factual, inferential, epistemic and hearsay meanings. Languages of the SW section also often exhibit an opposition between ‘intentional’ (or ‘volitional’) and ‘non-intentional’ (‘non volitional’). See Huber (2002: 162), Graves (2007: 69), Tournadre et al. (2009). Epistemic auxiliaries and suffixes are not sufficiently documented.

9.8.7.5. Negation

In the languages of SW section, the negation forms are derived from CT MA and MI as in other Tibetic languages. In Sherpa, the negation form /ma/ is used for the completed past and the imperative. The negation /ma/ is always prefixed with the imperative and usually prefixed with the past, but it may also occur after the verb and be prefixed to the auxiliary (with the progressive, see Tournadre et al. 2009: 279).

For the uncompleted present and the future, the negation is always prefixed to the verb and has several allomorphs MA/ma/, MI/mi/ MU/mu/ ME/me/ or
The vowel following the labial $M$ is homorganic with the vowel of the following lexical verb. This has been noted independently by Grave (2007) for the Sherpa Hile dialect and by Tournadre et al. (2008) for the Sherpa Khumbu dialect. Similar allomorphic variations of the negation marker are attested in Dölpo.

In Kyirong and Yolmo, as in Sherpa, the negation marker is normally prefixed to the lexical verb. The only exceptions in Kyirong are the ‘future’ and ‘aorist’ sensory markers, where the negation follows the verb and is prefixed to the auxiliary $\ddot{y}õng$/yong/. It is realized respectively as $\dddot{m}iõ$/miõ/ and $\dddot{m}ayõ$/mayõ/ (Huber 2002). Just as in CT, the negation $\dddot{M}$ /ma/ is used with completed past and prohibitive while $\dddot{M}$ /mi/ or /me/ is used with the uncompleted aspect (past and present) and the future (Gawne 2016; Huber 2002). Data is insufficient for other languages and dialects of the SW section.

Generally speaking, one of the main characteristics of the negation marking of the SW section is its prefixation to the lexical verb.

### 9.9. The Western section

The Western section (henceforth W section) corresponds to a quasi-continuum of neighboring dialects traditionally spoken by cultivators in Lahul & Spiti and Kinnaur Districts, as well as in Chamba and Kishwar districts (Himachal Pradesh, India) and by some agropastoralists of the Jangthang area of Ladakh. All the dialects of the Western section are tonal and do not have preinitial sounds which are characteristic of the dialects spoken in Ladakh (except the Jangthang area) and Baltistan.

The region of Spiti plays a leading role in the preservation of the local culture and language within the Western section. The mutual intelligibility between the Spiti, Garzha and Khunu dialects is rather good. The intelligibility with the other dialects of Paldar, Pangi and those spoken in the Jangthang is probably more limited due their isolation.

From the linguistic point of view, the W section occupies an intermediary position between the western dialects of the Central section, especially the neighboring Tö dialects of Ngari and the dialects of Zangskar and Ladakh of the North-western section. One could even consider that there is a quasi-continuum of dialects from Sumdo (at the border between upper Kinnaur and Spiti) to Rongchung, which is located in Tsanda.
County on the Tibetan side. The intelligibility is reportedly very good between these dialects. However, for some linguistic specificities as well as geopolitical, historical and sociolinguistic reasons, it is clear that Spiti, Lahul and upper Kinnaur dialects constitute a separate linguistic entity from Ngari, and that they cannot be grouped together with the dialects of Zangskar.

In the W section, the official languages are Hindi and English. Both languages are taught in the schools. The dialects of Lahul, Spiti and upper Kinnaur are not written down. Most people write in Hindi and to a much lesser extent in English. Pahari an Indo-Aryan language is also spoken in the area of Kinnaur and Lahul. Some people of upper Kinnaur can speak various Kinnauric languages and most people can speak Hindi. One should also mention a substantial number of Nepalese migrants, some of whom are native speakers of Tibetic languages, such as Yolmo.

A growing number of people speaking Tibetic dialects of Spiti, Lahul or upper Kinnaur are now bilingual with Common Tibetan. This is due to the strong influence of the Tibetan community under the spiritual leadership of the Dalai Lama and the Central Tibetan administration whose seat is in Dharamsala, a town also located in Himachal Pradesh, about 200 km from Lahul and Spiti. Literary Tibetan is essentially confined to the Buddhist monasteries and to some rare private schools such as Serkhang School in Tabo, Samten Chöling in Jispa (Lahul) or Keylong Central School (Lahul).

The Tibetic dialects of the W section are in contact with Kinnauri proper and various Kinnauric languages:

- In Kinnaur District: Chitkuli (spoken in the Sangla valley), Jangshung (spoken in Morang Tahsil, Jangi, Lippa); Tukpa (spoken in Nesang, Tsaorang, and Kunnu). Shumcho (spoken in Puh Tehsil: Kanam, Labrang, Spilo), Sunnan (spoken in Puh Tehsil: Sunam).
- In Lahul: Pattani (alt. Manchati) in the Myar valley and at the jonction of the Bhaga and Chandra valleys; Tinani in the Chandra valley; Gahri (alt. Bunai)

57. Alternative names are: Thebarskad (Ethnologue) or Tibarskad. The Kinnauri dialects of Puh and Morang tehsil are also known as Orisanuskad, Tamiriyskad and Yamskad.
in Keylang and Kardang area.

They are also in contact with various Indo-Aryan languages such as Pahari, Garhwali and Chiniali.

A few villages of this Tibetic speaking area particularly in upper Kinnaur and in Spiti have also a Kinnauri origin. For example, a lot of toponyms in Lower and Upper Kinnaur bear the name /rang/ (i.e. Rarang, Dirang, Sera Rang), which means ‘peak’ in Kinnauri. The mountain name Pomarang situated in Lower Spiti literally means ‘snowy mountain’ in Kinnauri. A village of upper Kinnaur is called Khab, which means ‘gorge, ravine’ (and not ‘needle’ as it would be in Tibetan!).

The term Spiti itself སྤི་ཏི་SPI. TI, which is sometimes spelled སྤྱི་ཏི་SPYI. TI in old texts, is probably of Kinnauri or Zhangzhung origin and could be translated as the ‘water from above’. The term ti means ‘water’ in Kinnauri and, although in Spiti the normal term for ‘water’ is chu as in the other Tibetic languages, a few expressions related to water still bear the word ti in the Spiti dialect such as niti ‘remaining irrigation water’ and zagti ‘smaller canal for water irrigation’ as pointed out by Spiti Gelong Dorje (2011). A few words containing the root ti are also found in Ladakh and Baltistan e.g. khati ‘saliva’ (La). The term ti is also found in Western Tibet in the toponym རྩེ་སྦས་TL. SE, which refers to the Tibetan name of Mount Kailash, as well as in the word རིད་TL. If the above hypothesis is correct, Balti could thus refer to the ‘water from below’. The region of Spiti and Khunu in the upper Sutlej is in fact much higher than the region of Batistan which corresponds to the lower course of the Indus river. A Ladakhi scholar A.G. Sheikh (2010: 164) independently proposed the following explanation:

“Balti means ‘watery ravine’: bal means ‘ravine’ or ‘valley’, and -ti means ‘water’. The name ‘Balti’ is derived from the dialect of Zangzhung, […] There are also several words in Ladakh with the suffix -ti, such as chage ‘water leaking through a roof’, natì (a running nose) and ratì (the rising or flooding of water in a stream, etc.).”

58. SPYI is used among other things to refer to the crown or top of the head (as in SPYI.GTSUG).
The existence of many toponyms of Zhangzhung or Kinnauri origin throughout the Western regions (Ngari, Khunu, Spini, Garzha, Ladakh and Baltistan) points towards the existence of a non-Tibeti substrate.

Many people, particularly in upper Kinnaur, in Garzha (Lahul), Pangi, Paldar and to a lesser extent in Spiti, have recently adopted Indian names such as Raj, Ram, Devi, Kumar, Rakesh, Baldev Singh, etc., or mixed names such as Angyal Ram (Wanggyal Ram) instead of the traditional Tibetan names Karma, Tshering, Dikiyt, etc (the name are usually given in the school). Moreover, an Indian ‘family name’ or surname, ‘Negi’, is used for all the people of Kinnaur. Thakur (ཐཱཀུར་) is given to many inhabitants of Kaza (Karze), but this practice is not pervasive in Spiti and Garzha, where the traditional names of Tibetan origin are generally used.

In the region of Spiti, Upper Kinnaur, Garzha, Pangi and Paldar, speakers of Tibetic dialects are followers of Vajrayāna Buddhism, but in many areas, particularly in lower Kinnaur and Lahul, one can witness some syncretism with Hinduism, which is the dominant religion in Himachal Pradesh. The syncretism is stronger and obvious among the non-Tibetic ethnic groups (Kinauri, Lahuli, etc.), than among the speakers of Tibetic languages. It is interesting to note that people who practice Buddhism are often referred to as ‘Pot’ (i.e. Tibetan) whereas the Hindu are referred to as ‘Mon’ (i.e. Ethnic Kinnauris).

In Spiti, the majority are followers of the Gelugpa sect, but Nyingma and Sakya monasteries are also present in the area. The five great monasteries of Spiti are: Kyilgön Norbu Gephel (Gelug), Thegchen Chöling (Gelug), Drangkhar Trashi Chöling (Gelug), Tringön or Gungri Orgyen Sangngak Chöling (Nyingma) and Tenggyü Gön Chönkhor Ling (Sakya). One should also mention the very old temple Lhalung serkhang in middle Spiti, which is allledgedly older than that the monasteries of Tapo and Drangkhar.

59. Also called Lavot Gompa བ་འོད་. 60. Kardze Chökhorling is a branch of Tenggyü monastery.
In Lahul, the monasteries are mostly Drugpa Kagyü and Nyingma. They include Kardang Gonpa (Keylong), Drânphuk Lhakhang (near Khoksar), Samdrup Chökhörling (Teling), Gemon Samtän Chöling (Gemur) and Shashur Gompa (Stingri).

One should also mention the famous pilgrimage site near Uddaipur called ‘Triloknath’ or Garzha phagpa, which is a case of Hindu-Buddhist syncretism. The statue of the shrine is revered by both Buddhists and Hindus either as Avalokiteshvara or as an avatar of Shiva.

The main monasteries of upper Kinnaur are located in Chang (alt. Chango): Kamtshang Thubtän Özer Rabgyäling (Kagyu nunnery) and Trashi Yangmo previously called Kharkhok.

Menri Monastery, a major Bön monastery built by the Tibetans in exile is found not far from the W section, in Dolanji, Himachal Pradesh. In Khunu most of the monasteries are Drugpa Kagyü but there are also Nyingma and one Gelugpa monastery.

Most dialects of Lahul, Spiti and Upper Kinnaur have not been described or have received little attention. The dialects of Upper Kinnaur and Garzha are particularly threatened to disappear since Literary Tibetan is not taught in schools (contrary to the situation in Spiti) and the local authorities of Himachal Pradesh or the Indian Central government do not provide sufficient help to support the traditional culture and education of the Upper Kinnaur people and Lahul & Spiti.

Most of the publications are devoted to Spiti: Grierson (1909) and Hein (2001, 2007), Kato (2001), and to a lesser extent to Lahul: Roerich (1933). A dictionary of the Spiti dialect was compiled by Matthew Thomas, Drimet Lodrö and Tsewang Dorje but has not been published.

9.9.1. Migration patterns, legends and historical records

The region of Lahul, Spiti and upper Kinnaur has been part of various kingdoms in the course of history. Before the Tibetan empire it was under the kingdom of Zhangzhung (Bellezza 2008; Vitali 1996), whose capital was reportedly Khyunglung,
in the upper Sutlej valley (Langchen Khabap), not far from the present region of Spiti and upper Kinnaur. The Zhangzhung language was probably a west Himalayan language related to Kinnauri. After the annexation of Zhangzhung by the Tibetan Empire in the seventh century, the region fell under Tibetan control. After the fall of the Empire, it became part of the Guge-Purang Kingdom who also conquered a large part of Ladakh. After the fall of the Guge-Purang Kingdom and the emergence of the Namgyal dynasty (sixteenth-nineteenth centuries) in Ladakh, Spiti was from time to time under Ladakhi control. Thus, it is clear that from the linguistic point of view, Lahul, Spiti and upper Kinnaur have been under the influence of both Tibetan and other west Himalayan languages such as Zhangzhung and Kinnauri. As pointed out earlier, the languages of the W section are in contact with a number of Indo-Aryan languages.

Migrations for Western Tibet and contact with the Tibetans have continued until the annexation of Tibet by the People’s Republic of China in 1959.

9.9.2. Linguistic groups of the W section

Further research is needed to determine precisely the degree of mutual intelligibility but from the recent data show that the intelligibility is rather good. According to Veronika Hein (pers. comm.), Khunu and Lahul-Spiti dialects are fully intelligible:

“In spite of the phonological differences, Piti-kat speakers refer to Khunu-kat as ‘ngui kat’ ie. ‘our language’. [...] The area around Darcha and Jispa, where a Tibetan dialect [called Töt-kat or Kolong] is spoken, is located in the Bhaga Valley and geographically separated from Khoksar [in the Chandra valley] by other Tibeto-Burman languages of Lahaul like Gahri [also called Bunan, a Kinnauric language], in the Keylong area.”

Hein concludes from the geographic situation that the dialect spoken in the Darcha and Jispa might be linguistically more related to the dialects of the Zangskar valley. N. Roerich (1933) already mentions this point suggesting an affiliation of Töt-kat with the northern dialects of Zangskar and Ladakh:

“The first [Kolong] is commonly designated by the name of To̱k-kag [Töt-kat] and is related to the subdialect of Ladaki spoken in the upper Indus Valley above Sheh [...] and to that of Zangskar spoken around sPadum.”
From the data collected during our own field work (2013) in Garzha-Töt, we can say that some Garzha words are similar to those of Zanhar. However, the Garzha phonology is not closely related to Zanghar phonology and differs mainly by the presence of tones and the fact that preradicals are not pronounced and do not trigger fricativization as in Zanhar.

Here are some basic words which illustrate the difference between the Garzha dialect and those of Spiti and Khunu-Töt and the proximity with Zanhar བཅོ་BCO (Ga, Za) versus ལེེ་BYED (Sp, Khu) 'to do', ཇོ་CHA (Ga, Za) versus ལེེ་GRO (Sp, Khu) 'to go', རོ་STOR (Ga, Za) versus རོ་SKYAL 'to lose', བཀོར་ZACHAS (Ga), ZAN (Za) versus བཀོར་LOTOCHAS (Sp, Khu) 'food, meal', མོ་THORE (Ga, Za) versus བཀོར་NANGMO (Sp, Khu) 'tomorrow'. Despite these types of lexical convergences, we maintain the grouping of Garzha with Spiti and upper Kinnaur for phonetic, grammatical as well as geographic and cultural reasons.

For the dialect classification of the W section, we propose the following eight groups:

- Spiti སྤི་ཏི
- Khunu-Töt ཁུ་ནུ
- Garzha གར་ཞྭ
- Pangi བང་གི
- Paldar དཔལ་འདར
- Durbuk Jangpa dialect དུར་བུག་ས་ཁུལ་གྱི་བྱང་པའི་སྐད་
- Nyoma Jangpa dialect ཉོ་མ་ས་ཁུལ་བྱང་པའི་སྐད་
- Jadang (or Dzathang) dialect རྫ་ཐང

The dialects of Spiti and Khunu-Töt are very closely related, whereas the distance is a little greater between Garzha and Spiti dialects. The Pangi and Paldar have yet to be described. However, it is possible to say that there are grammatical differences even concerning the basic existential auxiliary which is བཀོར་SNANG in Pangi (as in Balti: Turtuk and Nubra) instead of ལེེ་DUG in Spiti, Garzha, Khunu and Ladakh Jangthang.
Additionally a dialect usually referred to as Jadang is spoken by about 400 people in Uttarkhand Pradesh. One dialect of the SW section, Mugum, is spoken in Himachal not far from the Garzha speaking area.

9.9.3. Geographic extent of the W section

The area of the W section is located in the Himalayas between Zangskar and Nepal along the Sino-Indian border, in the Indian States of Himachal Pradesh and Uttarkhand Pradesh. It extends over three districts: Lahul & Spiti, Kinnaur (both in Himachal Pradesh) and marginally in Uttarkashi (Uttar Pradesh). Additionally two dialect groups are spoken in the Ladakh Jangthang area in the State of Jammu and Kashmir.

The region of Lahul-Spiti and upper Kinnaur used to be part of the Guge Kingdom of Western Tibet. Later it fell under the power of Ladakhi kings. The famous Tabo monastery of Spiti was built by Rinchen Zangpo in the end of the tenth century.

The region is drained in the west by the Chandra River (upper course of the Chenab) and the Bhaga river as well as by the Langchen Khabap river (upper course of the Sutlej) which runs westward to join the Indus river. The Spiti river is one of the tributaries of the upper Sutlej and receives as its affluents the Pin, Lingti (གླིང་ཏི་ཆུ་) and Pare (པ་རེ་ཆུ་) rivers. The Beas River, which is part of the hydrological basin of the Indus River is one of the five rivers of Panjab. It joins the Sutlej river but is located outside the Tibetic-speaking area.

In the east, the area of Jadang is drained by the Bhagirathi river (CHU BO SKAL IDAN SHING RTA) and its tributary the Jadh Ganga river. The Bhagirathi is one of the two headstreams of the Ganges.

61. Since the upper Chandra river is located in a Tibetic-speaking area, the term Chandra ‘moon’ in Sanskrit is most probably a translation of the Tibetan བོད་ ‘moon river’.

62. In Persian, Panjab literally means ‘five waters’ (panj ab): it refers to the Jhelum, Chenab (Chandrabhāga), Ravi, Sutlej and Beas.
Detailed location of the dialect groups

- **Spiti** སྤི་ཏི་སྐད་

  The dialect of Spiti སྤི་ཏི་སྐད་*Piti-kä* locally pronounced /piti-ket/ is spoken in the Spiti Tehsil and to a lesser extend in Lahul Tehsil of Lahul and Spiti district in the Spiti valley and adjacent valleys.

  This dialect includes four varieties: Tabo ལོ་གསར་, which is also referred to as Sham ཀྲ་'lower region', Kaza referred as Bar བར་'middle region', Kyil རི་'higher region' and the Pin valley བོད་, whose name is derived from དཔྲིན་'SPRIN' 'cloud'.

  The main villages of Spiti are: Losar ལོ་སྒར་, Kyomo དཀྱིལ་བར་, Hänsa ལྷ་ནས་, Trotro དྲོ་, Pangmo རང་མོ་, Morang རོང་, Khurik སུམ་གླིང་, Ririk སོ་, Sumling སུམ་གླིང་, Ririk སོ་, Khurik སུམ་གླིང་, Sangnam རང་མོ་, Guling སུམ་གླིང་ and Tabo རྟ་པོ་.

  The honorific language is more used in lower Spiti and Pin valley than in upper Spiti. For example, the term for 'you' is generally བོད་/‘nyöt’ (Hon.) ‘you’ < OT ཝེད་NYED in Karze and lower Spiti, while in the upper area, the ordinary form བོད་/‘khyöt’ ‘you’ is still used. The humilific for ‘I’ is བོད་/‘mā’/ in the Pin valley and a cognate form /mā/, is used in Tabo and upper Kinnaur. The general humilific form in Spiti is བོད་/‘ngo/ but the ordinary form is བོད་/‘nga/.

  The main village is Kaza, which is currently pronounced in Hindi as Kaza or even Kaja, is derived from Classical Tibetan བཀྲ་ཤིས་ DKA R, MDZES.

- **Khunu Töt** གུན་ནུ་སྟོད་སྐད་

  The dialect of Upper Kinnaur གུན་ནུ་སྟོད་སྐད་‘Khunu-Tökä’, locally pronounced /Khunu-Töktä/ is spoken in Upper Kinnaur District (Himachal Pradesh). It is sometimes called ཁྲ་ཏོད་ ‘Rong-kä’ (locally /rong-kät/) i.e. ‘the gorge language’, by the inhabitants of Spiti. This dialect of upper Kinnaur used to be called Nyamkat or Sangyas (according to Ethnologue) in earlier publications, but now local people do not use this.

63. Kyil is spelled in different ways as 'Key', 'Ki', etc., on Indian maps.
term, which is an exonym. The term Nyam is used by the lower Kinnauri people to refer to the Buddhist inhabitant of upper Kinnaur. 'Bhotia of upper Kinnaur' is sometimes used as alternative name for the Khunu dialect. However, the term Bhotia may induce confusion between Kinnauri (a Himalayan TB language) which in Tibetan is referred as མོན་སྐད་ 'lower Khunu language', རོང་ཆོས་ 'the gorge language' or even མོན-ཆོས་ 'Mön language' and the Tibetic dialect of upper Kinnaur. Most of the area between Rampur (རམ་པུར) and Rekong Peo (རེ་ཀོང་པེའོ) near Morang village), the district capital, is a Kinnauri and Pahari speaking area, whereas the Tibetic speaking area essentially begins at village of Puh.

The Khunu-Töt dialect speaking area was closed for foreigners until 1993 and an inner line permit is still required after Rekong Peo from Akpa and Jangi checkposts to Sumdo checkpost, which is located at the border between Upper Kinnaur and Spiti.

The Khunu-Töt dialect is spoken in three tehsils. Hangrang tehsil is located in the north along the lower Spiti river, Pooh and Morang tehsil in the valley of the upper Sutlej. The two first tehsils correspond essentially to a Tibetic speaking area (with some pockets of Kinnauri Shumcho speakers), whereas in Moorang tehsil, the Khunu-Töt dialect is spoken only in the remote villages in the valleys towards the Sino-Indian border.

The Khunu-Töt dialect is essentially spoken in the following main villages: Puh (alt. Pooh) བུའུ་, Chang གང་ (usually spelled Changgo), Nako (or Nao) གནས་ཞིག་, Trashi Gang གནས་སྒོ་, Kah འ, Khab བཐེ, Dubling དུ་བླིང་, Nesang རེ་སང་, Li (or Leo) རྗེ་, Namgya རྣམ་རྒྱལ་, Kyakhar (alt. Shelkhar) སྐྱ་མཁར་, Hang བཞ, Tsuling བཙུ་གླིང་, Sumra འསྱམ་རག་, Naling (or Meling) གནས་བཟང་, and in some villages of Morang: Lambar ཁམ་, Kun རུན་ and Tsarang རྩ་རང་ (generally written as 'Charang').

64. Term which is applied by the inhabitants of upper Kinnaur to the ones of Lower Kinnaur.
65. The name Rekong Peo has a Kinnauri origin. It means according to one of our consultants: “place to collect fruits.”
Garzha

The dialect group of Garzha is also sometimes referred to as Lahuli, for which the Tibetan etymology of TIBETAN: LHA YUL, i.e. ‘land of deities’ or ‘paradise’ has been proposed. The term Lahuli also designates a Tibeto-Burman language spoken in Lahul (Sharma 2007), so in order not to confuse these two languages, it is preferable to avoid the glossonym Lahuli to designate the Tibetic language and use only Garzha. This dialect group is spoken along the Chandra and the Bhaga rivers, in the Lahul Tahsil of Lahul & Spiti. It has three dialects: Töt-kat, Khoksar and Patanam.

The ‘Töt’ dialect is spoken in a few villages located along the Bhaga River. They include the following villages: Dartse, Sumdo, Jispa locally called Zhi (sometimes Zhipa), Gemur locally known as Gemön, Kolong and Khangsar. The ‘Töt’ dialect is spoken in a few villages located along the Bhaga River. They include the following villages: Dartse, Sumdo, Jispa locally called Zhi (sometimes Zhipa), Gemur locally known as Gemön, Kolong and Khangsar.

The ‘Patanam’ dialect (alt. Patnam) includes some villages located further west in Udaipur subtehsil, in the Patanam valley which is a tributary of the Chandra river.

The ‘Khoksar’ dialect is spoken in the upper course of the Chandra river in the villages of Khoksar, Dränphuk and ‘Teling’, whose local Garzha name is /samling/. The Khoksar variety is isolated from the Spiti area by the Kündzom pass and a desertic mountainous area of about 50 km. Sissu, the next village in the west,

66. Since the etymology of Garzha is not clear, there are a number of spellings which have been proposed but none are really convincing.

67. Dartse is located north of the Garzha area at the confluence of three rivers: Zangskar chu (not to confuse with the tributary of the Indus also known as other Zangskar chu), Zhung chu and Yotse chu, which form the Bhaga river.

68. In the village of Khangsar is found an old fortress (AKHILAR) where a king of Garzha had once its residence.
less than 10 km away from Teling, is a Lahuli-speaking village. Not only does the language change radically, but there are also significant religious differences. Although the inhabitants of Sissu are predominantly Vajrayāna Buddhists, the temple of Sissu, which has the shape of a Hindu mandir, is devoted to the deity Ghepan (sometimes considered as a form of Pehar Gyälpo), clearly a form of Hindu-Buddhist syncretism.

The dialects of Dartse, Khoksar and Patanam are not in direct contact. These Tibetic-speaking pockets are separated by various villages who speak Himalayish and Indo-Aryan languages. In the north, Dartse is separated from the Zanhar-speaking area by a desertic region with several high passes such as the Bara Lacha la and the Shingo la. Given their isolation, the dialects of Garzha are endangered and could disappear within one or two generations. Another aggravating factor is related to the climatic conditions: in winter, the Garzha area is entirely isolated from Manali and inlocked in the snow. This partly explains why many inhabitants of Garzha spend the winter months from October to April in the southern regions of Himachal Pradesh or further south. For example, in Khoksar, only a few residents stay during the winter to watch the houses. If the infrastructure is not improved, this will turn this area into a summer resort.

About the Khoksar and Kolong (Tōt-kāt) dialects, Roerich (1933) gave the following commentary:

“The Koksar [sic] sub-dialect in many details agrees with Spiti, but its phonetic structure is influenced by the neighboring Himalayan dialects, such as Tinān and Manchāṭī. At first one is tempted to class it as a branch of the Spiti dialect, but a closer investigation of its phonology shows its close affinity with the sub-dialect of Kolong in the upper Bhāga. [...] The Kolong sub-dialect has its immediate neighbor in the Bunan dialect [a Kinnauri language]. It is noteworthy that the latter has been strongly influenced by Tibetan in phonetic structure, noun inflexion and vocabulary, but its influence on Tibetan is almost negligible, being limited to a few loan-words.”

The main town of Lahul and Spiti district is Keylong, sometimes spelled ༽དབྲང་ 'Kyelang' in Tibetan script. Its population includes not only native speakers of Indo-Aryan languages (Hindi, Pahari, Nepali, etc.) but also of the dozens of TB languages (Himalayish and Tibetic) spoken in Lahul. Kardang, the ancient capital of Lahul, is
located near Keylong on the other side of the Bhaga River, where the eponym Kardang Drugpa Kagyü monastery བཀྲ་དྭངས་དགོན་ is found.

**Pangi བང་གི**

Pangi is an administrative sub-division or Tehsil of Chamba district in Himachal Pradesh, with its headquarters at Killar. The main town of Killar is located at a distance of 170 km from Chamba and 260 km from Manali. Hindu (95.25%) religion is predominant amongst the Pangwals, followed by Buddhism with about 1,012 followers or 4.7% of the total population.

The Pangi region is located in some valleys, tributaries of the Chandrabhāga river. In the lower villages, most people speak an Indo-Aryan language called Pangvali which is similar to Pahari and are followers of Hinduism. The Pangi speaking area is generally Buddhist but there has been some syncretism with Hinduism as in the case of Garzha. Most people are followers of the Drugpa Kagyu lineage. According to our consultant, the Tibetic migration in Pangi and Paldar area is fairly recent and dates back only five generations. People are said to have migrated from the adjacent areas of Garzha, Kinnaur and Zanskar. The main settlements are found in the tributary valleys of the Chandrabhāga river, in Chamba district, less than 50 km from Udhampur. The main Pangi speaking villages are located in the upper Saichu valley in Pangi Pot བོད་ (alt. Chasak bhatori) and Hill Tuan དྲ་གུན་ and in the next valley in Parmar Bhatori ཁང་རྒྱུས. Pangi is also spoken in another valley, tributary of the Chandrabhāga river in Hudan Batori རྒྱུན་, above the villages of Killar and Takwas. Another village is located in the upper Sural valley, in Sural Bhatori མུ་རུལ. The last village of between Pangi and Paldar is Ganir village གནོད་, but from an administration point of view, it belongs to Padder block in Kishwar District.

From a linguistic point of view the intelligibility between Pangi and Paldar is not very good. However as in most other Tibetic speaking areas, one is in the presence of a geolinguistic continuum. Pangi and Paldar dialects may soon undergo a rapid evolution since many children from the area are sent to Tibetan schools in Dharamsala and learn Common Tibetan.
• **Paldar**

The main villages of Paldar valley include Kabon village, the new settlement of Gulabgarh which is located at the banks of the Chenab river (also called Chandrabhāga). Most people who live in this settlement come from the higher valleys and speak different varieties of Paldar. The other villages are located in the upper Danglong valley in Hangu, Halo, Dranga, Jashel, Machail, Losseni and Sumtsham. Paldar-speaking people are in contact with speakers of Indo-Aryan languages. In summer, Gujjar and Gaddi ethnic groups, who are sheep and goat herders, bring their flocks into the area. The Paldar area is also located at the crossroad between Islam, Hinduism, and Buddhism. The first Muslim villages are found just after Atholi, in Kishtwar District.

In 2010, the Dalai Lama visited the town of Gulabgarh on the request of Himalayan Buddhist and Cultural Association. Since then Buddhism has found a new popularity in the region. Many people who had adopted Hindu names are now using Tibetan Buddhist names.

• **Duruk Jangpa dialect**

This Jangpa dialect (alt. Changpa) is spoken in the Ladakh Jangthang, in the area of Pangong lake region and in the upper Shayok river. It includes the villages of Duruk Block, Duruk, Trangse, Chushul, Man-Merak, Shachukul, Thakhung, Phobrang. The people of this area are called Lalokpas ‘from behind the pass’.

• **Nyoma Jangpa dialect**

Nyoma Jangthang dialect is spoken in the Ladakh Jangthang, in the Tsomoriri lake region, in the upper Indus valley and on the Rupshu plateau. It includes the villages of the Nyoma Block: Chumathang, Puga Sumdo, Nyoma, Loma, Hanle, Chumur (including other hamlets of the Pare valley), Kharnak, Tshokar (Thukje) and Tsomoriri (Korzok region).
Jadang dialect also called Dzad or Jad is spoken in Jadang and Nilang villages in Harsil Tehsil of Uttarkashi district (Uttarakhand). According to Sharma (2001), “the name Jad seems to be derived from the village name ‘Jadang’, which is the summer village of the Jad speakers.” The name of the village Jadang might itself be influenced by the Indian pronunciation of the Tibetan word རྫ་ཐང་ Dzathang, which means ‘Slate plain’. Jadang is located in the eponym valley of the Jad Ganga. As noted by X. Becker (pers. comm.), this valley is separated from the Khunu area by the Dzarong valley རྫ་རོང་ located on the Chinese side of the border, in Tsanda County (Tibet Autonomous Region).

Sharma gives the following precision: “During the winter, Jad speakers migrate to Dunda Sub-division, just 17 kilometers below the Uttarkashi district town on the banks of the river Bhagirathi [upper course of the Ganges].” G. van Driem (2001) explains that “the Jad people were resettled after the Indo-Chinese conflict of 1962. Some settlements are also found in Purola, Raiga Jhi, and Bhatvarl sub-divisions. Their original homes lay on the Indo-Tibetan border.” There is no recent data about the situation of the Jadang dialect. Grierson (1909) states that Jad is closely related to Spiti dialect. So far no linguistic data are available on Jadang and thus fieldwork is needed to establish the proximity between this language and the Spiti-Garsha-Khunu group of dialects. However, given the isolation of Jadang within a Hindi-speaking area, one can say this dialect spoken by only a few hundred speakers is seriously threatened of extinction.

69. Jad is mentioned by Ethnologue as well as Nishi (1986).
9.9.4. Number of speakers

According to the 1981 census, the population in Spiti is 10,383. Lahul Tehsil has 10,414 inhabitants (1971 census). According to the 1981 census, the total population of Khunu (or upper Kinnaur) area is 15,576. There are also 400 Jadang speakers in Uttarkashi District. The number of speakers of the Jangthang dialects as well as those of Pangi and Paldar are not known. It is difficult to establish precisely the number of speakers of the Western section due to the complex sociolinguistic situation and the impact of Indian languages in the area. However, the total number of speakers most likely does not exceed 35,000. Given this relatively small number with few monolingual

70. 3,099 in Hangrang, 5,086 in Pooh, 73,91 in Moorang.
The phonological diversity of the W section is rather limited. However, it is not possible to list common phonological features to the all the dialects of the W section. The phonological features of this section are very similar to those of the Western Tö Ngari dialects of the C section, but are slightly more conservative.

**Suprasegmental features**

The dialects of the Western section, including the two Jangthang dialects spoken in the state of Jammu Kashmir have a pitch tone system generally with a two-way contrast (high and low). This is a major distinction with the neighboring dialects of Ladakh (except the Jangthang dialects) and Zangskar. Concerning Pangi and Paldar, additional research is needed.

**Segmental features**

*Synchonic approach*

The sound systems of the W section are characterized by the following features:

- Existence of voiced non-resonant sounds (b, d, q, g, dz, j, z, zh, ḥ).
- All the dialects lack the voiceless /l/.
- Prenasalization ("d,"b,"g, etc.) is present but is often very light and barely audible.
Preaspiration is not attested.

The set of final consonants includes n, m, ng, p, ?, l, t. However /l/ and /t/ are not always clearly audible (except for the Jangthang dialects) and may be realized respectively as /:/ and /ʔ/. The set of vowels is limited.

Diachronic approach and reflexes of Classical Tibetan

In the languages of the W section, the preradical sounds are no longer present except for some prenasals.

In Spiti and Khunu-Töt, voiced non resonant sounds (b, d, ɖ, g, dz, j) are derived from the consonants with preradicals (except m and ’), however in the Töt and Patanam varieties of Garzha, the consonants without preradicals also yield voiced sounds.

SR corresponds to /ʂ/ in Spiti and Garzha and to /ʈ/ in Khunu-Töt: ལན་མ་/ʂ änma/ or /ʂatma/ (Ga, Sp) versus /تانما/ (Khu) ‘peas’. སུང/ /ʂung/ (Ga, Sp) versus /ʧʊŋ/ (Khu) ‘to keep’, སའ/ /ʂaʔ/ (Ga, Sp) versus /ʦʔ/ (Khu) ‘to burn’, སོ/ /ʂoʔ/ (Ga, Sp) versus /ʦoʔ/ (Khu) ‘life breath’.

The reflexes of SPR and SKR are identical to those of SR (see above).

The final consonants P, M, N and NG are usually preserved. However, in some contexts NG leaves no trace. The final consonant G is either maintained or changes into a glottal stop; L is preserved as /l/ or trigger a lengthening of the vowel /:/ and d is realized as /t:/ or as /ʔ/. CT LH is always realized as /l/.

The reflex of RTS is usually /s/.
9.9.7. Grammatical characteristics of the W section

9.9.7.1. Case markers

We do not have sufficient data concerning the Paldar and Pangi dialects. The case system of the Spiti dialect (which is closely related to Garzah and Khunu) includes the following cases (see Hein 2007):

- The absolutive Ø.
- The dative /-la/ (< CT case ལ་ LA).
- The ergative corresponds to /-su/ སུ་. It is rare and only used for an emphatic function. (Hein pers. comm.).
- The genitive /-i/ ཀི་ and the variant /-ki/, both derived from CT case གི་ GI. In some dialects of Spiti, the form /di/ is also attested after /n/: Ex. /gegen-dip/ “the teacher’s book”, and in some cases the /-i/ spreads to the preceding syllable: /sakhang/ ‘restaurant’ > /sakhip/ ‘restaurant’s boss’, /balang/ cattle > /balip/ ‘cattle’s tail’.
- The ablative is འན་ /na/ འན་སུ་ /nesu/ (see Hein 2007). The latter form is attested in CT.
- The instrumental is སུ་ /-su/. However, Hein (pers. comm.) mentions some variants such as /-nai/, /-nakisu/ or /-nasu/.
- The comparative གང་ /-sang/. Unlike Ladaks, this marker is not preceded by the genitive and directly attached to the preceding word.

9.9.7.2. Nominalizers

Various nominalizers (see 8.3.13) are found in the W section. They include the following markers:

- The nominalizer /-kan/ derived from CT གམ་ MKHAN.
- The nominalizer /-she/ and its variants /-če/ and /-zhe/ are derived from CT ཇེས་ CHAS ‘thing’. It is also grammaticalized as a nominalizer in many Western and southern languages: Ladaks /-čes/71, Kyirong as /-če/, Lhoke as /-shāʔ/ and

71. Reminder: in our phonological transcription, /č/ notes a non-aspirated consonant (see 7.1). The aspirated is noted as /čʰ/.
in Central Tibet, as /-yä/ and /-ya'/.

- As in other section the nominalizer /-sa/, derived from CT सा, is used to convey the place of the verbal action.
- The nominalizer /-na/ is used to indicate the instrument of the verbal action.

It is probably from a connective function of the locative न ‘when, if.’

9.9.7.3. Verbal inflections

The various dialects of Spiti, Khunu and Garzha have preserved some verb stem distinctions found in CT. For example in Spiti, the verb /'sa/ (present) becomes /'sö:/ in the past (Hein 2001) corresponding respectively to the reflexes of ZA (present) and ZOS (past). Just as Sherpa, Dzongkha and the neighboring dialects of Tö Ngari, the dialects of the SW section exhibit morphophonemic variation of the stem. However, available data are scarce.

In the W section some frequent verbs also exhibit suppletive forms to mark various tense-aspect and modalities. For example, in the Garzha dialect, the verb 'to go' has four forms: /'ča/ (present), /'song/ (past, first person, imperative), /'put/ (past, non first person), /'qul/ (imp), respectively derived from CT चस ‘to set for a trip’, सौं ‘to go’ (past), बु ‘to go out’ (past), ग्रु ‘to travel’.

9.9.7.4. Linking verbs and auxiliary verbs

We will illustrate the S section with data from Spiti. They were collected by Hein in Tabo as well as by Tournadre in Kaza (Kardze). However, the dialects of Spiti, Khunu and Garzha have a lot of similarities.

Copulative verbs

As in most Tibetic languages, the main copulative verb /'yin/ in Spiti2 is derived from CT यिन. Another copulative verb यिनः /yinnok/ is derived from the combination of CT यिन and दु ‘else’. Their respective negations are /'min/ and /'mak/.

---

2. The materials are from Hein (2001; 2000), Bielmeier (2000) as well as Tournadre’s field work in 2013.
Existential verbs

Existential auxiliaries of Spiti include the following verbs:

▪ ཡོད་/yöt/ pronounced [yöt] or [yö], conveys an egophoric meaning. The factual existential verb is ཡོད་ཀག་/yökak/. Both are derived from the CT existential verb ཡོད་/YOD/. The negation form is /met/ (< CT མེད་/MED/).

▪ འདུག་/duk/ (< CT verb ཡོད་/YOD/ 'to stay, to sit') conveys a visual sensory meaning. This existential verb has the following negation forms: མི་འདུག་/mi-duk/.

▪ གྲག་/ʈak/ or /rak/ is derived from the CT verb གྲག /GRAG/ 'to sound'. It is used for a non-visual sensory access to information. This copula has the following negation forms: མི་གྲག /mi-GRAG/.

Compound linking verbs

Compound verbs are also attested in the W section: this is the case for example of the above equative copula ཡིན་ནོག /yinnok/ (see also Hein 2007).

Auxiliaries

Frequent auxiliary verbs of the W section consist of the linking verbs (sometimes preceded by a relator) as well as some lexical verbs. They include:

▪ ཡིན་/in/ is used for the future and བེན /wen/ is used in the completed past. Both forms are derived from the CT copulative verb ཡིན /YIN/ and are used to indicate egophoric past and the future.

▪ བཏ /ct/ is used for the present and the uncompleted past. It is derived from the CT existential verb ཡོད /YOD/. According to Hein (2001), it indicates an egophoric meaning.

▪ སོང་/song/ < CT སོང་/SONG/ indicates a visual sensory marking used with the completed past. A variant སང /sang/ is analyzed by Hein (2007) as a 'mirative morpheme', but we consider that it can be interpreted as a connotation of the sensory meaning.

▪ The auxiliary བྱུང་/čung/ is mentioned by Hein (2001). It is derived from the
CT auxiliary BYUNG 'to come, to appear'. It conveys an 'ego-receptive' meaning, i.e. an event or action directed towards the speaker which is usually involved.

- /rak/ is derived from the CT verb GRAG 'to be heard of'. The possibility that it is derived from the CT verb REG 'to touch' was proposed by Bielmeier (2000) but it seems less plausible (for phonological and comparative reasons). First, we can argue that for the copulative function, both forms /ṭak/ (in Spiti) and /rak/ (in Garzha) are found in the W section as well as in Ladakh. Second, we have seen that the CT verb GRAG has also been grammaticalized in languages of Kham in eastern Tibet. It is used for a non-visual sensory access to information. The auxiliary /rak/ is used with the present and, together with the marker /de/ /derak/ (or as /perak/), occurs in the perfect. The auxiliary /rak/ conveys a 'non-visual sensory' meaning (Hein 2001). The endopathic function is also typically encoded by this marker and it is used to convey inner sensations, feelings and intuitions.

- /kak/ or /ak/ is used for the habitual present as well as the future. It is a kind of factual or assertive marker. Together with the marker /de/ /dekkak/ (or as /pekak/) occurs in the perfect. Hein (2001: 43) suggests that "/-ak/ is most likely a contraction of some particle to which the morpheme –ak is added as the negative form implies [ka-m(a)-ak]." For the suffix /-ak/, which is cognate with /kyak/ (in Leh Ladaks), we propose the following reconstruction: /kak/ < /kanak/ < * /khanak/ /ka"dak/ MKHAN."DAG. For the suffix /pekak/ < /pet-kak/ < * /payotka"dak/ PAYOD.MKHAN."DAG /peyotka"dak/. If our hypothesis is correct, it is thus cognate with the form /'odmkhan."dag/ /okande/ attested in Tö Ngari. Hein (2001) considers that /-kak/ "expresses the speaker's knowledge of the verbal action without specifying how this knowledge is/was gained" and called this marking "speaker's

73 The markers /de/ and /pe/ are probably reflexes of the connective DE and the nominalizer PA as suggested by B. Zeisler (pers. comm.). An alternative candidate for the marker /de/ would be the verb BSDAD 'to stay'.
unspecified knowledge." This interpretation has been confirmed by our own
fieldwork in Spiti. The auxiliary /kak/ is also found in the languages of the NW
section such as the Kenhat dialects of Ladakh.

- ུ/tuk/ and the allomorphs ཨ/ruk/ or ཨ/uk/ are derived from CT
  ས/DUG ‘to sit’. These endings are used for the present. Together with the
  marker /de/ ཨ/deruk/ (or as /peruk/) this auxiliary also occurs in the
  perfect. It conveys a ‘visual sensory’ access to information.

- ཨ/tang/ or its variant ཨ/rang/ is analyzed by Hein (2007: 7) as an
  ‘intentional morpheme’ and has a mirative tone. ཨ/ /tang/ and its variant are
derived from the CT verb ཨ/BTANG ‘to send’. As we have seen in Chapter 8,
this verb is grammaticalized as a secondary verb in Amdo and Dzongkha. In
Amdo, ཨ/ /tang/ also conveys among other functions an intentional meaning.
Thus in Spiti, this verb also functions as a secondary verb.

- ཧ/ɖo/ is used as an epistemic auxiliary (< CT ཧ/GRO ‘to go’. It is also used
  with the copula and evidential verbs /yöt-tro/, /yin-tro/. Other epistemic
auxiliaries are attested such as /-sere/ but epistemic auxiliaries and suffix
categories are not sufficiently documented.

As we have seen above the W evidential/epistemic systems usually have special
forms to mark visual and non-visual sensory access to information, as well as egophoric
and/or authoritative, factual or ‘unspecified knowledge’, inferential and epistemic
meanings. Among the specificities of the languages/dialects of this section, we find the
 distinction between visual sensory and non-visual sensory marking.

The grammatical marking of ‘intentionality’ is also present in the dialects of Spiti,
Khunu and Garzha (see also Hein 2001; 2007).

9.9.7.5. Negation

In the languages of W section, the negation forms are derived from CT ཨ/MA and
ཨ/MI as in other Tibetic languages. The negation marker are usually postponed to the
lexical and prefixed to the auxiliary.
9.10. The North-western section

The north-western section (henceforth NW section) includes the traditional regions of Ladakh བླན་དྭགས་, in India, as well as Baltistan བལ་ཏི་ཡུལ་ 'Baltiyul', which is situated in the Northern territories of Pakistan. Ladakh is itself made of several distinct regions with strong identities: Central Ladakh or Leh area གླེ (alt. སླེ SLE), Zangskar ཟངས་དཀར་, Purik པུ་རིག་, Nubra ནུབ་ར་ (alt. ལྡུམ་ར་ LDUM.RA) the Jangthang area བྱང་ཐང་ and Broqyul འབྲོག་ཡུལ་.

From an administrative point of view, one finds two Ladakh Autonomous Development Hill Councils (LADHC) for the two districts of Ladakh, Leh and Kargil.

The region of Ladakh བླན་དྭགས་ is pronounced /ladaks/ in the Central region and /ladaχ/ in Balti (hence the spelling 'Ladakhi'). We will keep the traditional spelling Ladakhi to designate the region and the people but we adopt for the language the spelling Ladaks which reflects the endonym instead of the Indian designation of Ladakhi. In the same way, we will maintain the name Zangskar and Zangskari for the region and its inhabitants but will adopt for the language the spelling Zanhar which reflects the local pronunciation /zãhar-hat/ (< CT ZANG.SDKAR-SKA.I)

The linguistic diversity in the NW is relatively high but one is confronted with a geolinguistic continuum of dialects. However, from a political and cultural point of view, it is convenient to speak of three closely related languages: Balti (in Baltistan), Purik (in Kargil area) and Ladaks (in Leh area). There is a fairly good intelligibility between neighboring dialects Purik and Balti, Purik and Sham, Sham and Central Ladak, Central Ladakh and Zangskar, etc. The division of the geolinguistic continuum into three major ‘languages’, Ladaks, Purik and Balti, is convenient, but it should not overshadow the dialectal diversity. A large database on the dialects of Ladakh has been established by B. Zeisler.

74. Ladakh was part of the state of Jammu & Kashmir, but since 2019, Ladakh has obtained the Union Territory status and is no longer part of Jammu Kashmir.
75. The English forms 'Zangskar' (and the alternative Zanskar) roughly correspond to the pronunciation of this word in Ladaks language.
One should note here that the term Ladaks or its Hindi-English equivalent Ladakhi is rather fuzzy. Although it is commonly used in Ladakh, and has been used for linguistic description by various authors such as Koshal (1979; 1982) and Norman (2001; 2019), it does not designate a precise dialect nor language. In most cases, it refers to the dialects of Central Ladakh. However, in a loose interpretation, it may include the dialect of the Nubra region, the Sham dialects, and even the dialects of Zanskar, etc. Whenever we use in this book the term Ladaks without further explication, it refers to the dialects of Central Ladakh spoken in Leh and around the regional capital.

In the NW section, the official languages are Hindi-Urdu and English. They are taught in the schools, both in Ladakh (on the Indian side) and in Baltistan (on the Pakistani side). Ladaks and Balti languages have been developing written forms during the last decades but they do not have an official status. Classical Tibetan is taught in many Buddhist monasteries of Ladakh but is not taught in the Muslim area (except for some marginal cases).

The Tibetic languages of Ladakh are in contact with Hindi-Urdu and various other Indo-Aryan languages such as Brokspat and Kashmiri which belong to the Dardic branch of Indo-Aryan (Indo-European). Brokpa (Shina) are traditionally Buddhists but a number of villages have converted to Islam. They live in the lower Indus valley, the Hanu valley and the Dras valley. They have been largely under the cultural influence of the surrounding Tibetic languages and speak Purik as a second or third language (with Hindi-Urdu) and, in some cases, as their native language.

Just as in the case of Spiti, a number of Ladakhi Buddhists (but also some Muslims) send their children to the Tibetan Children village of Choklamsar (Ladakh) or even to Dharamsala (the seat of the Tibetan administration in exile located in Himachal Pradesh) to study Literary Tibetan. Thus, Common Tibetan and Literary Tibetan are still considered as relatively prestigious. In the case of Balti and Ladakhi Muslims, Persian was also known by scholars and akbon or mulei (mullah) but now has a limited impact.

At soon as the end of the nineteenth century, Ladaks (གྲ་དབང་སི་སྐད་, Ladakṣi-skat) has been transcribed in Tibetan script, but there are still many discussions about the standardization. Some textbooks and publications have appeared in written Ladaks,
but their implementations in the school are still marginal. Programs in Ladak's language are broadcasted on the radio station in Leh and on the local television (Doordarshan \(^{76}\) Leh) and All India Radio (Kargil). There is also a local cable network 'Kargil Today' which broadcasts in Purik and Urdu languages. Written Ladak is so far not broadcasted on the internet, but as mentioned above, written Tibetan is still used in the Buddhist monasteries.

For Baltistan, the situation is somehow different, since the Tibetan script has been abandoned in the 15\(^{th}\) century after the conversion of Baltis to Islam. After the 15\(^{th}\) century Persian and then Urdu have been used as written languages. In the recent years, several transcriptions of Balti have been developed in Tibetan script (see chapter 5), Urdu script (a variety of Arabo-Persian) or even in Romanization. However, in the absence of standardization or official language policy, the development of written Balti is still very marginal. It is not available on the internet.

Concerning Purik (Kargil district in Western Ladakh), the situation is more complex. The majority of the population is Muslim, but there is a Buddhist minority. The Tibetan script is no longer used by the Muslims because they consider that the Sambhota script is associated with Buddhism. However, this script has been used by the Muslim rulers of Purik until the end of eighteenth century, beginning of the nineteenth century, when the country was invaded by the Dogras. As a testimony one finds a lot of manuscripts and inscriptions on rocks in Tibetan script signed by Muslim kings, Sultans or Maliks, of the Purik area.

As mentioned by Devers:

“Purig [Purik] became part of Western Tibet [Ngari Korsum] only in the early 11\(^{th}\) century it was overtaken by the armies of ‘Odde [the King ‘Od.de Kung rgyal]’ [...] At the time when Tibet was still practicing pre-Buddhist religion, Purig might already have been exposed to forms of Buddhism coming from the east” (Devers 2018: 10-11).

The NW section is characterized by a religious diversity with Muslims, Buddhists and a small community of Christians. Bönpos are no longer found among Ladakhis.

76. Doordarshan means television in formal Hindi but the Ladak's name རྒྱང་མཐོང་/RGYANG.MTHONG ‘television’ (see Norman 2019) is also used.
but the tradition existed and the monastery of Lamayuru བླ་མ་གཡུ་རུ་དགོན་ still bears the name Yungdrung འབྲླུང་རུང་ (Skt: svastika), which is the main symbol of Bön.77 The Tibetic area within Jammu Kashmir also includes a lot of Hindu and Sikh settlers and migrants, but they are generally not speakers of Tibetic languages.

Among the Muslims, Nurbakhshia and Shiʿah, are found in Baltistan (Pakistan) and in Ladakh (India) particularly in the Kargil district and to a much lesser extent in other areas of Ladakh. Nurbakhshia and Shiʿah Muslims are called འབྲལ་ཏི་ Balti whereas the Sunni Muslims are usually referred to as ཅེས་ Khache (< Kashmiri) or རྩེ་ Arghon. This community was formed merchants who originally came from Kashmir or Turkestan (Yarkand, etc.) and intermarried with Ladakhis. The term Arghon is probably derived from the Uighur word Argun which means ‘mixed’ (Norman 2017). This Sunni population is found in Dras, Panikhar (Suru valley), Leh, Shey and Thiksey. Padum the capital of Zangskar has a majority of Muslims whereas the rest of Zangskar is entirely Buddhist.

Concerning the Muslims of Ladakh, Abdul Ghani Sheikh wrote:

“Muslims have inherited many things from the Buddhists of Ladakh but they have themselves made significant contributions to the Ladakhi culture, art, music, language, literature, food and social life. There are many Urdu, Persian, Kashmiri, Turki and Uzbek words in the Ladakhi language introduced by Muslims. Muslims have enriched Ladakhi culture with the introduction of Ghazal and Qawalli [qasida and marsia]?” (Sheikh 2010)

Nangpa or Buddhists are essentially found in Leh district, i.e. in the region of Leh, the capital, and in Zangskar, Rupshu and the Shayok upper valley, but also in the Purik area of Mulbek and in the Brogpa (Dard) area of the lower Indus valley (in Dha, Dartsik, Garkon, etc.).

Ladakhi Buddhists primarily belong to the Kagyü school (both Drugpa and Drigung lineages) and Gelugpa school of the Tibetan Buddhism. The major and oldest monasteries of Ladakh are: གྲིག་རྩེ་དགོན་པ Thikse /thriktsa/ monastery (Gelugpa), འབྲལ་ཏི་ རྩེ་དགོན་པ Arghon monastery.
Alchi monastery, Hemis monastery (Drugpa), Phyang monastery (Drigung Kagyu), Pethup monastery usually called ‘Spituk’ (Gelugpa), Mangtro monastery also called ‘Mathro’ (Sakya), Lamayuru monastery (Drigung Kagyu), Lukyil monastery lit. ‘nāga encircled’ also known as Liker, located near Bazgo, Ridzong monastery (Gelugpa), (Nyingmapa), etc. In some villages of the Purik area (Kargil district), such as Mulbek, Apati and Khartse Khar, one finds well preserved giant statues of Buddhas and Bodhisattvas carved in the rock. Although the Purik area is largely Muslim, some villages such as Mulbek have both Buddhist and Muslim populations.

There are also many mosques in Ladakh, both Shi‘ah, and Sunni, particularly in Kargil district (رة). A few mosques are also found in Leh district (رة) and the major mosques in Leh are the Jama Masjid (Sunni) and the Matamsara mosque (Balti Shi‘ah) and also the Shey Masjid (Shi‘ah) in Shey village. In the Kargil area, one finds a famous khang in Baru. The Lartse Masjid (Shi‘ah), is located in Pashkyum village which historically used to be an important place on the road towards Leh. The most famous aishanas (graves of famous saints) are in Karpo Khar (رة) ‘lit. the white fort’ and Pharona in the Suru valley.

Finally, throughout Ladakh, one finds many ancient fortresses or khar (路口), both in Leh and Kargil districts (see Devers 2017). The names of some of these forts tell a lot about the rulers and the historical migrations of the region:  (BOD MKHAR.0U) Bod Kharpu ‘the small Tibetan fort’ (previously known as MKHAR.PO.CHE lit. ‘the big fort’),  (BROG.PAT MKHAR) Broqpai Khar (Dardic Fort),  (MON.MKHAR) the Mon Fort (probably referring to Kinnari-like populations),  the Turkic Fort and  (BAL.TLMKHAR) ‘the Balti fort’ (see also next section 9.10.1).

Most Ladakhi people in Leh District bear names of Tibetan origin when they are followers of Buddhism, whereas in Kargil District, most people have names of Arabo-Persian origin, except for those who are Buddhist and bear Tibetan names. The same holds for the Broqpas who have either Tibetan names or Arabo-Persian names depending
on their religion. In some marginal cases, one also encounters hybrid names made of both Tibetan and Persian or Arabic roots.


9.10.1. Migration patterns, legends and historical records

Ladakh is located at the strategic crossroad of nearly unavoidable routes leading to the Tibetan high plateau, which was once the production center of a variety of precious commodities such as gold, pashmina, musk, salt and iron. These resources were circulated along routes passing through Ladakh in direction of major trading centers, such as Khotan and Yarkand, located on the so-called southern Silk Road, the regions along the northwestern corridor of the Pamir and Hindukush ranges such as Gilgit, the Swat valley, Gandhāra and the Wakhan corridor and regions of the northern subcontinent such as Kashmir, Jammu, Kushiwar, Mandy, Lahul and Kinnaur (Rizvi 1999; Devers 2017). Unsurprisingly, various populations have coexisted and succeeded one another in Ladakh. Among the groups that we are able to trace, in addition to the Tibetans, there have been Dardic settlers, predominantly Brogpa, and also Shina. There have also been Turkic groups (Horpa), coming primarily from the Pamir corridor and from the Tarim basin. In some parts of the Dras and Suru valleys, there have additionally been Kashmiri and other Dardic populations. Finally, there used to be so-called ‘Mon’ communities, locally said to have come from Himachal Pradesh. As can be expected from such situation, numerous toponyms in Ladakh are not of Tibetan origin – examples are villages like Hemis, regions like Purik, and most probably Ladakh itself. As a testimony of this critical commercial location and diverse populations, Ladakh is the area of the Tibetan plateau with the greatest variety of
scripts observed in rock inscriptions. In addition to the use of Tibetan script, we find historical inscriptions written in Kharoṣṭī, Brāhmī, Chinese, Sogdian, Tokharian, Arabic and Śāradā scripts. It is also the area with the highest density of archaeological remains, being for fortifications, rock art, temple ruins, pre-Buddhist funerary sites, etc.

In the past, we can observe material differentiation between the eastern and western Indus valley: remains from the eastern parts, sometime associated by oral tradition to past Mon populations, display significant shared traits with the Tibetan Plateau; remains from the western parts, commonly associated by oral tradition to Brogpa, Horpa and Balti populations, display stronger kinks with the Pamir corridor. A transitional area of mixed influences lies in the central belt that extends from the Indus-Zanskar confluence to the confluence with the Gya brook. Parallel territorial variations can be outlined in Zangskar, Nubra and Purik, whereas in the westernmost fringes of the latter, Kashmiri and Shina influences are further noticeable (Devers 2018).

As can be expected in a country of considerable trading and mining wealth, these territories, as delineated by considerable material remains, evolved over the centuries, often considerably. Whereas ancient fortifications or toponyms relatable to the Brogpa can be found in all parts of northern Purig as well as in the Indus valley from Baltistan to the confluence with the Gya brook, Brogpa populations are now strictly confined to the lowest reaches of the Indus river, downstream of Hanu. More strikingly, the Mons have entirely vanished from the country.

Prior to the tenth century, the history of the territory corresponding to present-day Ladakh is largely unknown. We know that to the east, on the Tibetan high plateau, was Zhangzhung and to the west, along the Pamir corridor, was Balur (alt. Bolor). The exact borders of these two ancient kingdoms are unknown. In the light of the geographical distribution of the material remains that have been found, we may

postulate that a frontier between them lay in central Ladakh. In other words, the 
Ladakh as we know it now may have been divided between two larger kingdoms. A 
domain covering both eastern and western Ladakh seems to have emerged only in the 
tenth and eleventh centuries under the name Maryul, when the West Tibetan kingdom 
of Ngari Skorsum progressively brought these territories under its control (Devers et al. 2016; Devers forthcoming b). An alternative name for Maryul was Mangyul, the 
two names being apparently interchangeable. An alternative name for Maryul was Mangyul, the 
two names being apparently interchangeable.79 Ladakh, on the other hand, appears to 
have then been only the name of a subdivision of Maryul around Alchi, from where 
the Namgyal dynasty originated.80 When this dynasty took control of Maryul in the 
mid-sixteenth century, the name Ladakh seems to have progressively been applied to 
the country as a whole.

9.10.2. Linguistic groups of the NW section

In 2011, Zeisler proposed a classification of the NW languages into two major 
groups: གཤམ་སྐད་ GSHAM.SKAD sham-kā locally pronounced /Shamskat/ i.e. the ‘language 
of Lower Ladakh’ and གྱེན་སྐད་ GYEN.SKAD Gyen-kā, locally pronounced /Kenhat/ i.e. 
‘language of Upper Ladakh’. Unlike most other dialects groups of the Tibetic area, the 
Highlanders speak ‘innovative dialects’ while Lowlanders speak ‘archaic dialects’. Zeisler 
(2011) proposed the following description:

“According to phonetic features alone, the various dialects spoken in Ladakh are 
presently classified [by Bielmeier 2004 and others] roughly in two main groups:

• Western Archaic Tibetan: the non-tonal ‘conservative’ dialects of the north-eastern 
and central areas: Baltistan, Purik, Lower Ladakh, Nubra, and Leh, showing initial and 
final consonant clusters.

• Western Innovative Tibetan: the ‘innovative’ dialects of the south-eastern areas: 
Upper Indus, Changthang, and Zanskar, where the clusters have been reduced and 
tonal features can be found.”

79. For instance, in the seals of documents emitted by Nyima Namgyal found in Purig during 
surveys with Nils Martin, the Namgals present themselves as the “rajas of Mangyul” (rNal.mRgyal. 
Mang.yul.RaJa).

80. Howard was probably the first scholar to understand that Maryul and Ladakh were not 
The author in the same article proposed a more refined definition of the first group as: “The group of historically younger, but lexico-phonetically conservative Shamskat dialects” and the second group as: “The historically older, lexico-phonetically partly conservative, partly innovative Kenhat dialects” (ibid.).

Among the grammatical differences between Kenhat and Shamskat groups of dialects one can mention the definite marker which is དེ་DE in the former group and བོ་PO in the latter. The archaic form བོ་PO / བོ་BO is also attested in Amdo.

A detailed account of the sound changes and morphophonemic dialectal variation in Kenhat is presented in Zeisler (2011).

While Zeisler’s general distinction between ‘Shamskat’ and ‘Kenhat’ is certainly valid and relevant, one needs to propose subgroupings in order to describe the linguistic diversity of the NW section. If we consider ‘Shamskat’ (as defined above) for example, it is clear that Balti speakers do not easily communicate with Leh dialect speakers or the people of the Nubra region.

It is interesting to note that from a linguistic point of view two regions may occupy an intermediary position between the Western section and the Northern section: Zangskar and Kharu areas. Both dialect groups have developed a frequent fricativisation of initial consonants (see Zeisler 2011; Norman 2019), reflexes of r and s preinitials. However, these two dialects have preserved traces of other preinitials and thus are phonologically closer to the other dialect groups of Ladakh.

The Zangskar region was traditionally isolated from the rest of Ladakh at least six to seven months in winter and early spring (even if some people could walk on the frozen Zangskar River). A few mountain passes at an altitude of 4,500–5,000 meters prevent any transportation and communication to Zangskar and out: the Pensi-la82 in the north-west, Sengge La in the north, Charcha-la in west and Shingku-la in the south. However, southern Zangskar had traditionally more relations with Garzha.

81. Just as the difference between pastoralists’ and cultivators’ dialects is valid but insufficient to describe Amdo dialects.
82. A few hamlets of Zangskar are located on the northern side of the Pensila up to the village of Rongdum.
Finally, two groups of dialects of Ladakh are spoken in the Jangthang in the Pangong lake and in the Tsomoriri lake areas as well as the upper Indus valley. From a linguistic point of view, these two Jangthang dialects groups are more closely related to the dialects of Spiti and Western Tibet than to the other groups of Ladakh. They have developed tones and have not preserved initial consonantic clusters. They have also maintained prenasals a feature also found in Spiti and Tö Ngari but not elsewhere in Ladakh. All these phonological features are found in the neighboring dialects of the Western section and therefore we classify the Jangthang dialects together with these dialects (see the Western Section). A similar classification has also been adopted by the CTDT which has labeled these dialects as 'Western innovative dialects' (just as those of Spiti, Garzha and Khunu). However, from a cultural and political point of view, these two dialects are now heavily influenced by the other dialects of Ladakh and a growing number of people are bilingual with the Leh dialect.

For the dialect classification of the NW section, we propose the following seven groups:

- Balti འབལ་ཏི་སྐད་
- Purik བུ་རིག་སྐད་
- Nubra ཉུབ་ར་སྐད་
- Sham གཤམ་སྐད་
- Leh (Central Ladakh) གླེ་སྐད་
- Zanhar བཟངས་དཀར་སྐད་
- Kharu མཁར་རུ་སྐད་

9.10.3. Geographic extent of the NW section

From the administrative point of view, Balti is spoken in Baltistan (Pakistan) which is divided into two administrative districts: Baltistan District གལླེ་སྐྱེ་ཐང་ where the capital Skardo སྐར་མདོ་ is located, and Ganche District གངས་ཆེ་རྫོང་, as well as a few

83. With the exception of some varieties of Zangskar and Gya-Miru, as noted by B. Zeisler (p. c.).
villages on the Indian side: Turtuk and Bogdang in the Shayok valley as well as Hardas and a few hamlets in the Shingo valley.

Ladaks (central Ladakh, according to our terminology), Purik, Sham, Zanhar and Kharu dialects are spoken in the Ladakh Union Territory in Leh District གླེ་རྫོང་ and Kargil District དཀར་དཀྱིལ་རྫོང་. Leh District has consists of six blocks or 'tehsils': Nubra, Khalsi (Khalatse), Leh, Kharu, Durbuk and Nyoma. Kargil District consists of three tehsils: Kargil, Zangskar and Sanku.

The main valleys and rivers of this area are the Indus river, known as སེང་གེ་ཆུ་ 'Sengge-chu' i.e. 'Lion river' in Baltistan and in Ladakh, སེང་གེ་གཙང་པོ་ SENG, GE GTSA NG, PO pronounced 'Sengge Itsangpo'. In runs from the south-east to the north-west in Ladakh (upper course of the Indus) and then continues its course in Baltistan (lower course of the Indus). In the north-east of Ladakh, the main waters are the Shayok river སྦྱ་ཡོག་གཙང་པོ་ and its tributary the Siachen river སེ་བ་ཅན་གཙང་པོ་ (often called Nubra river on Indian maps). In the north-west region of Purik, the main river is the Suru river སུ་རུ་གཙང་པོ་ a tributary of the Indus, called Shingo in its lower course after its confluence with the Dras river. The Zangskar river (alt. Zanskar) ཟངས་དཀར་གཙང་པོ་ runs through Zangskar in the south to Central Ladakh in the north where it joins the Indus. The main tributaries of the Zangskar river in the east are the Lungnak-chu སྦུང་ནག་ཆུ་, also called སྦུ་རུ་གཙང་པོ་ སྦུ་རུ་གཙང་པོ་ སྦུ་རུ་གཙང་པོ་ in it upper course as well as the Töpe-chu སྦོད་པའི་ཆུ་ in the west. In the region of Ladakh, one finds two great lakes: the Panggong lake སྤང་གོང་མཚོ་, which is located in eastern Ladakh on the Sino-Indian border and extends over 130 km and the Tshomo Riri lake མཚོ་མོ་རི་རི་ located in the Rupshu plateau area.

Zangskar is made of four large valleys: the Zangskar valley, in the north-east, called Sham 'the lower (valley)' and the Tö valley, locally pronounced /tot/, lit. 'the higher valley', in the north-west which meet in the capital Padum /paðum/. The region around Padum is called Zhung, lit. the middle (valley). Finally, we have the Lungnak valley in the south-east after Padum. The Lungnak valley is separated from Garzha by a high pass, the Shingo-la. Additionally a few Zangskari villages are found on the north-western side of the Pensi-la in the upper Suru valley.
The region located in the lower Indus valley after Leh is traditionally called ‘Sham’ (lit. the lower [valley]); The territory from Nimu to Nurla (before Khalatse) corresponds to ‘upper Sham’ (STOD.GSHAM). ‘Lower Sham’ (SMLAD.GSHAM) starts at Khalatse and extends to the villages below, on the banks of the Indus river such as Domkhar and Achinathang as well as the village of Lamayuru situated in another valley which joins the Indus. The neighboring language, Purik, is also divided into two dialects: Eastern Purik and Western Purik. The former starts after the Photola pass and includes the villages of Henaskut, Bod Kharbu, Mundik, etc., as well as the main villages of Kangral, Chiktan and Yogma Kharbu in the Sengge Lungma valley. Eastern Purik extends to the villages of Wakha, Mulbek and Sharkol in the Wakha valley. After a long gorge, Darket is the first village of Western Purik, which includes the district capital Kargil as well as the whole Suru valley and its tributaries up to the border with Zangskar at Rongdum. Lower Sham and Eastern Purik, constitutes a geolinguistic continuum between Western Purik (Kargil area) and Central Ladakh (Leh area). During our fieldwork, we have found that there are even some fluctuations in the designation of the languages in some villages of Eastern Purik such as Henaskut, Bod Kharbu and Mundik which have both Buddhist and Muslim populations. The spoken language is either designated as Sham-skat or Purik-skat. It seems that the choice of the former or latter depends to a certain extent on the religion: Buddhists call their language Sham-skat or Purik-skat whereas Muslims prefer to call it Purik-skat.

The Nubra region, locally called /dumra/, is situated on the banks of the Siachen river (a tributary of the Shayok) and the lower Shayok river. As the Sham and Purik areas, Nubra has a lot of archaic features. The upper Shayok valley is entirely barren and not inhabited.

Finally, one ought to mention the Broqpas’ area. Broqpas originally speak a Dard language (called Brokskat) but nowadays a number of Broqpas speak Purik as their native language. The Brokpas are settled in the Dras valley, in the lower Indus valley and in the Hanu valley. The main villages include Dras, Hanu, Dha, Bema, Dartshik, Batalik and Culican. This last village is located right at Indo-Pakistanese border. Some
Brokpas are also located on the Pakistani side of the border on the Deosai plateau in Baltistan in some villages such as Ganoaks, Morol, Dananusar and Chechethang.

**Detailed location of the dialect groups**

- **Balti** 
  Skardo སྐར་མདོ, Rongyul or Rongdo རོང་ཡུལ་, Shigar ཤི་གར་, Khapalu མཁ་ལུ, Kharmang མཁར་མང་;^84^ and on the Indian side of Border: Turtuk ཤུ་ཏུག་ and Bogdang རོང་མདོ་ in the Shayok valley, as well as, Hardas, Karkichu and Latu in the Shingo valley.

- **Purik** གུ་རིག་ 
  *West Purik*: Kargil, རྒྱལ་, Baru, ཤང་, Sanku, ལྕུར་ལུ, Panikhar, རུམ་གྱུར་, Parkachik, དར་མྱིུ་ Dras,^85^ Darker འབྲལ་ and Olthinthang གཉིས་ལེགས་.

  *East Purik*: Mulbek རུམ་མྱུར་, Chiktan ནགུད་མྱུར་ and Bod Kharbu འབྲས་དམོ་.

- **Nubra** རུབ་ར་ 
  Khardong རྐྱ་དྲོང་, Khalsar པོ་མཆོག་, Sumur དུམ་པོ་, Kyagar རྦ་གར་ (called Tiger by the Indian army), Panamik འདོད་མེད་, Aranu རུས་, Yarma རྜྷར་, Deskit རྒྱུན་པོ་, Hundar ཐུངས་, Thois (alt. Thoise) ངུ་ཚེ་, Tirit ངུ་ཚེ་, Rongdo རོང་མདོ་, Agyam རྣལ་པོ་ and Digar རྡོ་རྗེ་. Some Balti villages are located in Nubra: Turtuk and Bogdang (see Balti).

- **Sham** གཤམ་ 
  *Upper Sham*: Nyemo (Snyemo) སྙེ་མོ་, Bazgo སྲེ་, Nyc སྲེ་, Lukil (alt. Likhir) སྲེ་, Saspol འཛུམ་, Alchi འཛུམ་, Uledkpo ལྡེན་པོ་, Nyurla (Nurla) སུ་རྒྱུས་ and Teya (Tia) སྲུད་.

  *Lower Sham*: Khalatse རྒྱ་ཁྲེ་, Skyindiang ནུ་རི་འཛུམ་, Tingmosang རུ་འཛུམ་, Domkhar རྒྱུན་, Achinathang རུ་འཛུམ་, Skyurbuchan སྲུ་འདོད་, Takmachik སྲུ་བཅུ་.

---

84. Balti communities are also found in Gultari, Kargil, Leh, etc., and in the Pakistani diaspora around the world. Ladaks speakers are found in Delhi (Buddhavihar, etc.), Jammu, Chandigar, Dehradun, Manali, Dharamsala and in south India, in Mysore, Bylakuppe, Bengalore, Mumbai, etc. and in the Indian diaspora around the world.

85. Purik communities are found in Leh and Gar (TAR).
Lamayuru བལ་མ་ཡུ་ (maybe derived from བལ་མ་ཡུང་དྲུང་), Wanla བན་པུ་, Hanu བན་པུ་and Da ིུ (Brokskat speakers have become speakers of Sham).

- **Leh གླེ་** (Central Ladakh)

Leh གླེ་, Spituk གླིང་སྐ།, Choklamsar གླིང་སྐྱ་མྱེ, Sabu བན་པུ་ (or བན་པུ་དྲུང་), Shey སྐད་ (or སྐད་སྲིད་མཚན་), Thikse སྐྱ་མྱེ, Stagna སྐྱ་མྱེ, Stok སྐྱ་མྱེ, Matho སྐྱ་མྱེ་ and Chemre སྐྱ་མྱེ. Some specific varieties of Central Ladakh are spoken in Lingshet གླིང་ཚེ, Gongma-Skyunpata གླིང་སྐྱ་མྱེ-སྟོག་, Yulchung-Nyeraks ཡུལ་ཆུང་ཉེ་རོ་, Photoksar དཔོ་ཏོགས་རྟེ་, Hanupata སྐྱ་མྱེ་ and the Markha valley, Chiling རྒྱ་མཚོ་ and Sumda རྒྱ་མཚོ་."}

- **Kharu རྒྱ་མཚོ་**

Kharu (Block) རྒྱ་མཚོ་, Igu རྒྱ་མཚོ་, Upshi རྒྱ་མཚོ་, Shara (Shera) རྒྱ་མཚོ་, Gya-Miru རྒྱ་མཚོ་, Kyere རྒྱ་མཚོ་ and Hemnya རྒྱ་མཚོ་.

- **Zangskar རྒྱ་མཚོ་**

Padum རྒྱ་མཚོ་ locally pronounced /fadum/or /faðum/, Rangdum རང་འདུམ་, Phe རྒྱ་མཚོ་, Sani རྒྱ་མཚོ་, Karsha རྒྱ་མཚོ་, Stongde རྒྱ་མཚོ་, Tsazar རྒྱ་མཚོ་, Zangla རྒྱ་མཚོ་, Rare རྒྱ་མཚོ་, Kargyak རྒྱ་མཚོ་, Phuktal (alt. Phukthar), རྒྱ་མཚོ་ and Shade རྒྱ་མཚོ་.

86. If the spelling *SPA.LDUM* is correct (it matches perfectly the pronunciation of Zanhar) and the pronunciation in Leh dialect also points towards this spelling, then the etymology would be ‘the juniper garden’. *SPA.MA* is a kind of juniper tree.
According to the 1971 census, Ladaks is spoken by 60,272 persons, but Zeisler (2012a) mentions a figure of 180,000 (including Purikpa and Zangskarpa).

For Balti, the 1981 census gives 223,296 inhabitants. Ethnologue mentions the figure of 270,000. The total number of speakers in the NW section probably does not

87. The figure of 400,000 proposed by Abbas Khazmi is probably an overestimation (see Zeisler 2006a; 2012a).
exceed 400,000. However, a precise number is hard to ascertain because of the impact of both Hindi-Urdu and English in this NW section.

9.10.5. Ethnic and sociolinguistic groups

The Tibetic or Bhoti people of the region usually refer to themselves as ‘Ladakspa’ (ལ་དྭགས་པ་) locally /Ladakpa/, Zangskarwa (ཟངས་དཀར་པ་) /zāharwa/, Balti (བོལ་ཏི་) and Purikpa (པུ་རིག་པ་). Additionally, the religion plays a significant role in the identity. Buddhists are referred to as ཨང་པ་ Nangpa, lit. ‘the insiders’. Muslim communities are referred to as Shī‘ah Balti or simply Balti where as the Sunni communities are called Khache (For example Khache of Padum in Zangskar).

The languages and dialects of the NW sections are predominantly spoken by བུ་དྭགས་པུ་ zhungpa ‘cultivators’ and agropastoralists. This is true for Balti, Purik, Ladaks and Zangskar. There are also some small pastoralists’ communities called གྱང་པ་ jangpa alt. ‘changpa’ settled in the Jangthang area of Ladakh (however, from the linguistic point, these communities are grouped together with the Spiti dialect, see the Western section) As mentioned in Chapter 2, the term བྲོག་པ་ BROG.PA locally pronounced /Broqpa/ or /Bloqpa/ designates native speakers of བྲོག་སྐད་ Brokskat, an Indo-Aryan Dardic language, closely related to Shina. Thus, in this region, unlike other Tibetic regions, the term does not designate herdsmen or cattle-breeders but an Indo-Aryan group. The Brokskat speaking region is called བྲོག་ཡུལ་ BROG.YUL, and is located in the Indus valley and the Hanu chu valley near Kargil as well as in some villages on the other side of the Indo-Pakistani border in Baltistan.

9.10.6. Phonological characteristics of the NW section

Although the phonological diversity of the NW section is rather limited, it is not possible to list common phonological features to the all the dialects of the NW section. The phonological characteristics are usually valid at the level of the two main groups Shamskad and Kenhat, but things are more complicated (see Zeisler 2011).

88. The name of the language is of course a loanword from Tibetan. It is sometimes written as Brokskat.
Suprasegmental features

It is well-known that the Tibetic languages in the NW section do not have distinctive suprasegmental features. However, in Balti for example, the position of stress may affect some phonetic realizations.

Segmental features

Synchronic approach

The sound systems of the NW section are characterized by the following frequent features:

- Multiple combinations of initials with a preinitial (see Chapter 7) are a widespread feature particularly in Balti, Purik, Sham, Leh region and Nubra. Labial, alveolar, velar, uvular preinitials are attested in most dialects, however prenasals are not found.
- In Zanhar and Kharu initial fricatives are attested.
- Existence of voiced non-resonant initial sounds (b, d, q, dz, j, z, zh, ɣ, ɦ).
- Some dialects have uvular consonants: q (final), χ.
- Some dialects have a rich set of final consonants, particularly Purik, Balti and Sham. They exhibit final combinations such as /ks/ and /ts/, ɣ /ngs/ and /ns/.
  However, this is not the case in Zanhar and other southern dialects. Zanhar has preserved only the final /l, r, t, n, ng, m/ but has no longer /s/.
- One characteristic feature of Balti dialects is the dissimilation strategy. When two nasals are in contact, one is often denasalized: GRANG.MO’cold’ > /graŋmo/, LCANG.MA’tree’ > /lchaŋma/, MDONG.MA’churn’ > /doŋma/, etc.
- A limited set of vowels. Zanhar has developed diphtongues: CHOS /čœ/ ‘dharma’/’religion’, SPOS /foe/’incense’, etc.

Diachronic approach and reflexes of Classical Tibetan

- In the languages of the NW section, the reflexes of preradical sounds are realized as segmental features (Balti, Purik, Bazgo, Nubra, Leh), yield a fricativisation as 89. The final s corresponds to the past form of the verb.
PART 2 – CHAP 9. Inner classification of the Tibetan languages 627

in Zanhar (and sometimes Leh).

- Voiced non-resonant sounds (b, d, ḍ, g, dz, j) are derived from the consonants with preradicals. In some dialects, b, d, ḍ, g, dz, j may also be derived from the consonants without preradicals.

- The combinations KR, KHR, GR, PR, PHR, BR as well as DR yield either retroflex sounds (tʃ, tʃ', d) in East Purik, Leh, Nubra and Zangskar or preserve the consonant clusters (in West Purik and Balti): /kr/, /tʃr/, /pr/, etc.

- Preinitial (labial, retroflex, uvular, glottal, etc.) is triggered by all the CT preradicals: RK, LK, SK, BK, LG, RT, LT, ST, etc. Some dialects (Pur, Ba, Sham) still make a distinction between the preradicals R, S, and L, while other dialects, which have retained the preradicals, such as Leh, do not usually make a distinction between R, S, and L. In this dialect, the preradicals tend to merge with S, but in other dialects, they merge with L (notably in the Nubra dialect, Zeisler, pers. comm.).

- All CT final consonants G, NG, D, N, B, M, R, L, S are well preserved in most dialects except for the S which is dropped in Zanhar and for the G which is realized as a glottal stop (Za). The second final S (YANG-JUG) is preserved in many dialects, notably Balti, Purik, Sham and Leh.

- In the Western dialects (Balti, Purik and Western Sham), the combination SL yields: /ts/ or /ʃ/ and and the combination zl: /dz/, /d/. This is the case in words like ཕྲལ་ Kangri 'easy', ཕྲལ་གྲི་Zla Gir 'moon'.

- In Zanhar, all the initial clusters with an s preinitial such as SK, SKY, SG, SD as well as a R: RK, RG, RD, etc., yield fricative sounds: མོ་ /mo/ 'door', རཾལ་ /hampo/ 'dry', རོང་ /hangpa/ 'foot', མཅོག་ /çetpa/ 'back (body)', མཆོག་ /çitpo/ 'pleasant'.

- In Zanhar and Kharu, the phonetic change from affricate to fricative: རྩ་ RTSI 'vein' > /sa/, རྩི RTSI > /si/, རྩི་གུ་RTSI GU 'apricot kernel' > /siwu/. This feature is also found in Spiti and Garzha and Ladakh Jangthang.

9.10.7. Grammatical characteristics of the NW section

The dialects of the NW section exhibit some grammatical differences particularly in their verbal morphology and syntax. For example, concerning the verb morphology,
we find the following basic CEV. Most dialects of Ladakh, such as Leh, Sham, Nubra, Zanhar and Eastern Purik (up to Wakha-Mulbek and Sharkol) have a special auxiliary for non-visual sensory and endopathic: རག /rak/ or འདུག /dag/ or /dag/ derived from the CT verb འདིག گراغ. This auxiliary verb is also found in Spiti, Garzha, Khunu, Ladakh Jangthang, western Tibet (Ngari), Dölpo (Nepal) and some Kham dialects (see SAK LA ZANG GYUR MED and SAK LA ZANG DBYANGS CAN 2002). This auxiliary is not used in Western Purik (Kargil area), where it is replaced by འདུག 'DUG and འདུག 'YOD nor in Balti where it is replaced by ལུང 'SNANG and འདུག 'YOD.

Another significant difference concerns the CEV ཀྲིན YIN and འདུག 'YOD. In Central Ladaks, Zanhar and Sham, these two verbs convey an egophoric meaning, whereas in Purik and Balti, they convey a neutral or authoritative meaning.90

Another difference between the languages and dialects of the NE section concerns the nominalizers (see below) and infinitive markers: Ladaks has /-čes/ or /-ches/ whereas Purik has /-pa/, /-ba/, /-ma/, and /-a/ depending on the context. Among the syntactic differences, we may mention the position of the adjectives. In Balti and Purik, the adjective is normally placed in front of the head noun as in English: གྷྲི་མ་ཤོག་བུ་ 'old book', རེས་མ་ཟམ་པ་ 'new bridge', རྒྲུལ་པོ་རྒྱུ་མ་ 'black intestine', རྡེ་མོ་ཡུལ་ 'a nice village', རྡོ་ཐོབ་ 'blue eye'. This is not the case in Ladaks and Zanhar and other Tibetic languages, where the adjective follows the head noun.

9.10.7.1. Case markers

The languages and dialects of the NW section distinguish up to seven cases (see Zeisler 2007; 2011). Frequent cases include ergative, absolutive, dative, ablative, genitive and locative, comparative and associative.

For example, the Leh dialect includes the following cases:

90. The replacement of the auxiliaries YIN NOG, YOD KYAG and گراغ respectively by YIN, YOD and 'DUG has a significant impact on the evidential-epistemic grammatical system since both the distinction between egophoric and factual as well as the distinction between non-visual and visual sensory non longer exist.

91. Depending on the languages, the semantic meaning may be a general locative or an inessive, an illative.
The ergative /-e/ or its variants /i/ or /yi/ (Koshal 1979) is derived from the CT ergative case ང་. It is identical to the genitive. It is worth noting the ergative and genitive cases are still distinct in Balti, Purik, Nubra and Sham: these dialects have preserved the final -s of the CT ergative marker ང་ (or its allomorphs): -(ɨ)s/ in Sham, /-ze/ in Nubra (see Zeisler 2007), /-(ɨ)s/ in Purik (see Zemp 2018), /-sil/, /-s/ in Balti (see Read 1934; Bielmeier 1985).

- The absolutive Ø.
- The dative /-a/ ར་ derived from the CT case ང་ LA.
- The genitive /-e/ variants /i/ or /yi/ (< CT case གི་ GI). This case is morphologically identical to the ergative in the oral language but written differently. See the ergative above.
- The ablative /-ne/ ས་. It is pronounced /-na/ in Sham dialects.
- The comitative /-tang/ ཞེ or its usual variant /nang/ ཞེ.
- The comparative /-sang/ བཞི. This marker is usually preceded by the genitive. Additionally Zeisler (2011) mentions the existence of a locative marker /ru/, /roa/ and other allomorphs in some dialects of Shamskat and Kenhat. This case which is marginal in the system is derived from CT ང་ RU.

Thus all the cases found in Ladaks are derived from forms which are cognates with CT. The only exception is the case /-sang/, which is also found in other western areas such as the Western section and the Tö Ngari area (Central Section).

In written LADWAGS-SISKAD, the ergative is written down as ང་ 's and thus is different from the genitive ང་ 't. Another characteristic features of the written system is that the ergative, the genitive and the dative when attached to a word ending in a

92 For example in the Ladagi Melong, or in the texts of Khanpo Konchok Phanday and Bakula Bakula Rangdol Nima, ང་ ལོང་ "S/he-ERG" (Konchok Phande 2017: 19) versus ང་ ལོང་ "S/he-GEN" (ibid.: 23), ལོང་ NYE.RANG-NGIS "you(H)-ERG" versus ལོང་ NYE.RANG-NGI "you(H)-GEN."
consonant are formed phonologically by adding a homorganic consonant followed by -I for the genitive, -IS for the ergative and -A for the dative.93

The dialects of Sham and Purik as well as Balti have a similar case system as the Leh dialect. There are however some discrepancies particularly in the differential marking of the ergative and the genitive. In these dialects, the ergative has a clear reflex of the final 'IS, as mentioned above. Depending on the dialects, the ergative form is /-s/, /-is/, or even /-ze/ (Zeisler 2007; 2011; Zemp 2018), whereas the genitive is generally /-i/.

9.10.7.2. Nominalizers
Various nominalizers are found in the NW section (see 8.3.13). They include the following markers:

- The nominalizer /-k’an/ or /-kan/ derived from CT མཁན་ MKHAN is used in Ladaks and Balti. Jäschke (1881) notes that “in colloquial language, esp. in Western Tibet, it has […] entirely displaced the proper participle termination in PA.”

- The nominalizer བྱེས་ BYES /-čes/ (sometimes written བྱེས་ BYES) or its variant བྱེས་ /-shes/ derived from CT ཆས་ CHAS ‘thing, tool’. It is the main citation form for the infinitive in the Ladaks dictionaries. The word CHAS is also grammaticalized as a nominalizer in the Western, South Western, Southern and Central Sections. In other words, this nominalizer is found under various forms in most Tibetic languages except for Dzongkha, Sherpa and the languages of Eastern Tibet, i.e. Kham and Amdo.

- The nominalizer ག་ PA /-pa/ or its variant ག་ /-ma/ derived from CT ག་ PA are used in Purik and Balti.

- As in other section the nominalizer /-sa/, derived from CT ག་ SA is used to convey the place of the verbal action.

93. For example ག་པ་ SKAD-DI “language-GEN”, བྱེས་ NANG-NGA “inside-DAT.” (The homorganic consonant is in bold).
9.10.7.3. Verbal inflections

The northwestern languages have not inherited the irregular verb morphology of OT and CT (see the discussion on this issue in Chapter 8). Zeisler (2004: 620) describes the Ladaks verb morphology in the following way:

“In Ladakhi [Ladaks], the imperative stem is identical with the present stem when the syllable of the present stem is closed, and identical with the past stem when the present stem ends in a vowel. A separate imperative stem showing ablaut a ⇒ o does only exist when the present stem has a vowel a.”

The verb in the ‘past’ (conveying the completed aspect) normally takes a suffix –S and this suffix corresponds to the generalization of the suffix –S attested for many verbs in OT and CT.

For example in Ladaks: སེར་ ZERS /zer-s/ ‘said’ (cf. CT: སེར་ ZER), གུན་མ་ ’DUGS /duk-s/ ‘stayed’ (cf. CT: གུན་ DUG) and in Balti: མིན་ SHIS /shi-s/ ‘died’ (cf. CT: མིན་ SHI). The suffix –S is however not entirely systematic as noted by Zeisler (2004: 620):

“There is a certain tendency that only controlled action verbs show the past tense suffix –s or show post-final suffix –s, whereas accidental [non-controllable] event verbs either do not take the past tense suffix –s or show post-final –s, in the present or, rather neutral stem, but there are a few exceptions.” Moreover, the suffixation of the –s varies a lot according to the various languages and dialects: “Despite the tendency in West Tibetan [our Northwestern section] to generalise the suffix –s, for verbs that were lacking it in Old Tibetan and Classical Tibetan […], there is an opposite tendency in Balti and Purik to generally delete it in the Simple Past, i.e. the mere past stem not followed by a morpheme.” (Zeisler ibid.: 624)

In the NW section, just as in all other sections, verb suppletion to mark various tense-aspect and modalities is also attested. For example, in most dialects, the verb ‘to go’ has two suppletive forms: /ča/ (present), /song/ (past, imperative), respectively derived from CT ངེས་ CHAS ‘to set for a trip’, སོང་ SONG ‘to go’ (past) (see Norman 2019; Hoshi & Tondup Tsering 1978).

9.10.7.4. Linking verbs and auxiliary verbs

Copulative verbs

The main copulative verb corresponds to ཡིན་ YIN or བི་མིན་ /in/ in Balti, Purik and Ladaks. The negation is བི་མིན་ /men/. These forms are derived from CT ཡིན་ YIN and
This copula has different meanings in the various languages. In Balti and Purik, it lacks egophoricity (see Bielmeier 2000; Zemp 2018) whereas in Ladaks, it has definitely an egophoric meaning.

In Ladaks, the form ནིན་གུང་ /inok/ is used to convey an inferential meaning (Bielmeier 2000, and section 8.4.3.4) but also a factual meaning (see 8.4.3.6). The copulative verb ནིན་ཤུལ་ /intsuk/ is used in Balti and Purik to convey an inferential meaning (Zemp 2018).

**Existential verbs**

Existential auxiliaries of the NW section include the following verbs:

- ི་ /yot/ < CT existential verb ི YOD. The negation form is /met/ (< CT མེད MED). In Balti and Purik, it conveys an authoritative or neutral/unmarked existential meaning, whereas in Sham, Zanhar and Ladaks (Central Ladakh), it conveys an egophoric meaning.
- ས་ /duk/ is attested in Purik and Ladaks. Its use is normally related to visual perception. This verb is derived from the CT verb ས DUG 'to sit'. it has the following negation forms: བ་ནུད་ MI-DUG /minduk/ or བ་ནུད་ /minuk/.
- བ /nang/ is attested in Turtuk (Balti) where it conveys a sensory meaning and the neighboring Nubra dialect where it is related to visual perception versus /(<r)-/ak/ indicating non-visual perception.
- ག /rag /rak/ or the variant ཕ GRAG is used in Ladaks, Zanhar, Sham and eastern Purik (Mulbek, Chitkan, etc.). It is derived from the CT verb ཁ GRAG 'to be heard of'. It is used for auditory and more generally non-visual sensory and endopathic access to information. This copula has the following negation form: བ་བ་ /mirak/ (even shortened as /miak/ in the Nubra valley) or བ་བ་ MI-GRAG.

**Compound linking verbs**

Compound linking verbs are frequent in the NW section. They include for example ི་པིན་ /yotpin/ derived from CT ིཔིན YOD.PAYIN or the inferential
Auxiliaries

Frequent auxiliary verbs of the NW section consist of the linking verbs (sometimes preceded by a relator) are some lexical verbs. They include:

- ◊ འིན་ IN /-in/ or ◊ འིན་ YIN are derived from the CT copulative verb ◊ འིན་ YIN is used for the future and the completed past. It indicates an authoritative meaning.
- ◊ འིན་ནོག་ YIN.NOG or the variants ◊ འིན་ནག་ YIN.NAG, ◊ འིན་འདག་ YIN.DAG has two entirely distinct functions: it conveys an authoritative (factual) meaning and a visual sensory inferential meaning. These auxiliaries are attested in Ladaks (Leh, Sham) and in the dialects of Rongkat and Jangkat along the upper Indus valley and in the Jangthang. The form found in Chumathang and Nyoma is ◊ འིན་འདག་ YIN.DAG which is the most archaic. Following the Indus river down towards Leh, the auxiliary changes to ◊ འིན་ནག་ YIN.NAG and then to ◊ འིན་ནོག་ YIN.NOG (in Leh). The auxiliary ◊ འིན་འདག་ YIN.DAG is also attested in Tö Njari on the Tibetan side of the border. It is important to distinguish clearly ◊ འིན་འདག་ YIN.DAG, the authoritative/visual sensory inferential marker and ◊ འིན་གྲག་ YIN.GRAG, the non-visual sensory despite the fact that in some dialects such as Jangkat, the two pronunciations are very similar: ◊ འིན་གྲག་ YIN.GRAG /yin $\ddot{d}$ ak/ (with a retroflex) and ◊ འིན་འདག་ YIN.DAG /yindak/ (with a dental sound). In Leh, the same opposition is very clear: ◊ འིན་འདག་ YIN.NOG versus ◊ འིན་གྲག་ YIN.GRAG /yindak/.
- ◊ འིན་ PIN /-pin/ derived from CT འཇལ་ PAYIN has two distinct functions: 'simple past' and the 'remoteness marker'. The former is auxiliary is only compatible with controllable verbs and normally used with the first person. It indicates generally an egophoric meaning. As a marker of 'remoteness', it occurs after a first auxiliary and simply indicates a past reference.
- ◊ འིན་ YOD /-yot/ or its allomorphs འིན་ /-et/ འིན་ /-at/ is used for the 'simple present' and the 'present continuous' (Koshal 1982/2005: 18, 69) and the uncompleted past.
▪  བུག་ /-buk/, བོག་ /-bok/, འོག་ /ok/ are probably derived from CT བ་འདུག་ BA 'nominalizer' + 'DUG' 'to sit' (see Zeisler 2017). They are used to mark the inferential future.

▪  བུག་ /-buk/ and its variants བོག་ /-bok/ , འོག་ /ok/ are used to convey a past inferential meaning. The etymology of /-buk/ (-/dok/, /-rok/) is rather clear. As shown by Zeisler, it is derived from CT ཀུན་ DUG.

▪  སུག་ /-suk/ and སུག་ TSUG sometimes spelled སུག་ TSHUG in the written transcriptions (Koshal 1979). There has been some debates about the origin of this marker. Tournadre and Konchok Jiatso (2001) and Denwood (2007) have proposed to derive these forms from the related CT verbs སུགས་ TSHUGS 'to be steady, firm, to be established' or རྒྱ རི རུ་ ZUG 'to be planted, to get pricked' and pointed out that the auxiliary རི རུ་ ZUG was also found in Amdo and Kham. However, Zeisler (2017) proposed for the Shamskat dialect an alternative explanation. The form results from the development: -s-duk > -suk ~ -sok or -se-duk > *-se’uk > -suk ~ -sok. For Purik, Zemp (2018) has also proposed to derive སུག་ TSUG from the ‘conjunctive participial suffix དེ་ STE + ཀུན་ DUG. Even if there are some phonological problems with this hypothesist, their argumentation is quite convincing. Particularly the parallelisms between ཨོ་ YOD and ཀུན་ DUG in the resultative constructions: ཨོ་ YOD, ཨོ་ YOD, ཨོ་ YOD, ཨོ་ YOD as well as the negation of སུག་ TSUG which is རྒྱ རི རུ་ ZUG.

▪  རག་ /rak/ and its variant རཀ /trak/ is derived from the CT verb རཀ རི རུ་ GRAG ‘to be heard of’ is used in a ‘non-visual sensory’ meaning (see Tournadre, 2022). The verb GRAG is grammaticalized as an auxiliary in Ladakh, Spiti-Garzha-Khunu, Tö Ngari and some Kham dialects. The endopathic function is also typically

94. Namely the fact, that after the auxiliary གཉིས་ YIN, the conjunctive participial should be /-c/ and not /-se/ according to Zemp’s own data. Thus it should yield གཉིས་ DUG, which is unlikely to change into གཉིས་ (T)DUG.
encoded by this marker and it is used to convey inner sensations, feelings and intuitions. This copula has the following negation forms: ◊ བོད་ /mirak/.

- ◊ ཁྲ /-kak/ or ཉོ /kyak/ is used with the various tenses and aspects. According to Koshal (1982: 590), it refers to 'general statement' or 'habitual actions'. It is also attested in Spiti-Khunu-Garzha. For a discussion about the etymology of /-kak/, see the W section.

- ◊ བོད་ /-rNaN/ or its variant ◊ ཏོ /-ang/ is used in the Nubra, Balti and western Shamskat dialects (see e.g. Ebihara 2014). It is derived from CT བོད་ SNANG 'to appear'. Jones (2009) describes this form as a 'present mirative' (see also Zeisler 2017).

The NW evidential/epistemic systems usually have special forms to mark visual and non-visual sensory access to information, as well as authoritative or factual, inferential, hearsay and epistemic meanings. Among the specificities of the languages of this section, we find the distinction between visual sensory and non-visual sensory marking.

9.10.7.5. Negation

In the languages of NW section, as in other Tibetic languages the negation forms are derived from CT བོད་ MA and བོད་ MI. The negation markers may occur after the verb and prefixed to the auxiliary, or before the lexical verb.
10. Contact languages

10.1. The various language families in contact with Tibetic

From a historical point of view, the Tibetan scholars have been exposed to various literary languages such as Sanskrit and some Prakrits, Classical Chinese, Persian (and otherIranic languages such as Khotanese). All these languages had some limited impact on literary Tibetan. Sanskrit had a more significant impact because it was chosen as the model language for the Tibetan grammar (see Chapter 5), and also because a great number of texts were translated from Sanskrit. In this section, we will concentrate on the current contact between the Tibetic languages and the spoken languages.

The Tibetic languages are in contact with many other languages spoken on the Tibetan Plateau or in the Himalayas and the Karakoram. Most of them belong to the Sino-Tibetan macrofamily but some are affiliated to other families such as the Mongolic, Turkic, Indo-European and Burushaski. See Endo et al. (2021) for their geographical distribution with linguistic maps under the Asian geolinguistic perspective.

The ST languages in contact with Tibetic essentially belong to three branches: "Tibeto-Himalayan," Qiangic and rGyalrongic. Tac Tibetic languages are also in touch with Lolo-Burmese, Tani, Naic languages and a few other languages, whose classification is currently under debate.

It is important to note that the contact languages are not only found at the margins of the Tibetic speaking area but are also located in Tibet itself. Within the Tibetan Autonomous administrative units, particularly within Sichuan, Qinghai, Gansu and Yunnan, one finds many non-Tibetic languages, which are spoken either by Tibetans or other ethnic groups (see Chapter 2) who have settled in Tibet. According to Roche and Suzuki (2018), there are fifty-two non-Tibetic languages spoken within Tibet. (For detail, see the various sections of this chapter; see also Sonntag & Turin 2019.)

Some of these non-Tibetic languages, such as Oirat, Bai, Drung, Lisu, Nosu, Naxi, etc. are also spoken in other territories outside Tibet, while others are only found in

---

1. Several scholars consider that rGyalrongic languages constitute a well-established subgroup of Qiangic (see Sun 2000).
Tibet. The Chinese authorities have given the political status of “nationality” to some of these communities who speak non-Tibetic languages, such as the Salar, Tu, Western Yughur, Bonan, Mönpa [Chin: Menba], Lopa [Chin: Luoba], etc. (see Chapter 2), but the majority of the communities who speak non-Tibetic languages are officially considered as belonging to the Tibetan Nationality.

10.1.1. Contact with Tibeto-Himalayan languages

The Tibetic languages are in contact with many Tibeto-Himalayan languages which belong to the following subgroups: Bodish, Tshangla, Tamangic and Western Himalayan.

The Bodish languages and Tshangla are spoken in the eastern part of the TAR (China) as well as in Bhutan and Arunachal Pradesh (India).

The Tamangic languages are located in Nepal and the Western Himalayan languages in Himalachal Pradesh and Uttarakhand (India).

Some Kiranti languages spoken in the Eastern Himalayas such as Limbu are marginally in contact with the Tibetic languages of Sikkim. For some information about the Tibeto-Himalayan languages and their relation to Tibetic, see 10.4. For the list of Tibeto-Himalayan languages, see 10.7.1-5.

10.1.2. Contact with rGyalrongic languages

rGyalrongic languages are also genetically related to Tibetic languages and in close contact particularly with Amdo and Kham dialects. They are spoken in Sichuan province and marginally in the TAR. The rGyalrongic languages have preserved an archaic morphology. The various rGyalrongic languages have borrowed a great deal of vocabulary from Tibetan at least since the seventh century and have preserved many archaic forms. In the rGyalrongic speaking area the traditional written language is Literary Tibetan.

For more detailed information about the rGyalrongic languages see 10.4 and 10.7.6.

2. Limbu, Yakkha, Chiling, Athpahariya, Lohorung, Yamphu, Mewahang, Kulung, Nachiring, Sampang, Sam, Chamling, Puma, Bantawa, Chintang, Dungmali, Hayu, Thulung, Ombule, Dumi, Bahing, Sunwar.
10.1.3. Contact with Qiangic languages

Qiangic languages are spoken in south-western China in the provinces of Sichuan, Yunnan, Gansu (mostly in Tibetan autonomous prefectures) as well as in TAR. None of the Qiangic languages, except Tangut, are written and Literary Tibetan is used as a written language.

For more detailed information about the Qiangic languages see 10.4 and 10.7.7.

10.1.4. Contact with Tani, Lolo-Burmese, Naic and Nungish languages

A few Tani languages are in contact with the Tibetic languages. They are spoken in the TAR (China) as well as in Arunachal Pradesh (India). Lolo-Burmese and Naic languages are spoken essentially in Yunnan but also in Sichuan. Nungish languages such as Trung and Rawang are spoken respectively in Yunnan (China) and the Kachin state (Myanmar).

All these languages are in contact with Kham Tibetan. In Sichuan, the speakers of Kham generally do not learn Nosu even though they live in the area close to the Nosu-speaking area. In contrast, the speakers of Kham in Yunnan learn some Nosu if they live in the same community. In Weixi County, the elder people of various ethnic minorities used to speak several ethnic languages such as Tibetan, Naxi, and Lisu rather than Chinese, and the mutual influence of each language can be found especially in the vocabulary. In Myanmar, Burmese is an intrusive language and Tibetan speakers also speak it.

10.1.5. Contact with Sinitic languages

A few dialectal groups of Mandarin are in contact with Tibetic languages. Among them, one can mention 兰银 Lanyin (Gansu), 中原 Zhongyuan (Gansu, Qinghai), 西南官话 Xinan-Guanhua “Southwestern Mandarin” (Sichuanese and Yunnanese). 普通话 Putonghua or standard Mandarin is the official language of the People’s Republic of China. As such, it is spoken in the main towns and villages

---

throughout the Tibetan speaking area in China and it is widely used in the schools and media.

There are three so-called “mixed languages” reported and described so far. They are as follows:

- **Daohua** (A-tshogs 2004): a mixed language of Tibetan and Chinese (Southwestern Mandarin) is spoken and it is locally called Daohua. The variety spoken in Nyagchukha County in Kandze Prefecture (Sichuan) is well described.

- **Wutun** (Sandman 2016): a mixed language which has a Sinitic substratum but has incorporated many Tibetan and Mongolic (Bonan) elements; spoken in Rebgong County (Tongren) in Malho Prefecture (Qinghai).

- **Selibu** (Zhou 2018; Zhou & Suzuki 2020, 2022): a mixed language which has a Sinitic substratum (Southwestern Mandarin but Zhongyuan Mandarin remaining) and has incorporated many Tibetan elements; spoken mainly by Hui Nationality living in Gyalthang Municipality in Dechen Prefecture (Yunnan).

### 10.1.6 Contact with Indo-Iranian languages

Indo-Iranian languages spoken within the Tibetic area, particularly on the southern slope of the Himalayas and the Karakoram comprise mainly Indic (or Indo-Aryan) languages. They also include one Iranian language. Here is a list of the main languages in contact with the Tibetic languages:

- Hindi-Urdu in Ladakh, Baltistan, Spiti, Garzha, Upper Kinnaur and Sikkim (and Indo-Aryan dialects such as Pahari in Himachal Pradesh).

- Nepali in Northern Nepal, Sikkim (India) and Bhutan.

- Bengali in the Darjeeling and Kalimpong areas of West Bengal and Sikkim (India).

- Assamese in Arunachal Pradesh (India).

- Dardic languages such as Brokskat or Shina in Baltistan, Ladakh and Kohistan. (See Yoshioka 2015: 209.)
• One Iranian language, Wakhi, is found near the Balti-speaking area in Pakistan.

10.1.7. Contact with Mongolic languages
Southeastern Mongolic are in contact with Tibetic languages in the north of the zone, mainly in Amdo. The various Mongolic languages spoken in Qinghai are Oirat (Wulân, Tuulán, and Sogvo counties), Mangghuer (Minhe County), Mongghul (Huzhu County), Bonan (Tongren or Rebgong County). In Wulân County, Mongolian speakers do not speak Tibetan, whereas in Sogvo, Tibetan is dominant and only elders can still speak some Mongolian. In addition, there is another language, Wutun, which is considered a Chinese-Tibetan-Mongolian creole (see above). It is spoken in Amdo in Rebgong County (Chin: Tongren), Qinghai Province.

10.1.8. Contact with Turkic languages
Uyghur is mainly spoken in the neighboring province of Xinjiang, but Tibetic languages are no longer in direct contact with Uyghur. However, historically, at the time of the Tibetan empire, Tibetan and Uyghur were in contact.

Salar and Yughur languages are the only Turkic languages spoken today within the Amdo region (see Simon 2016). The Salar people are traditionally Muslims. A few of them speak Amdo Tibetan. Kazakh language, another Turkic language, is spoken only at the periphery of the Tibetan area and only marginally in contact with Amdo.

10.1.9. Contact with Burushaski
Burushaski is a language isolate spoken in the Hunza valley at the border with Baltistan. Burushaski is mentioned in Tibetan annals of the empire as Drusha. Some Burushos have also settled in Baltistan, and this language has been in contact with Balti for many centuries. Many articles have been devoted to this language but Burushaski remains a riddle for linguistics. Concerning the linguistic descriptions of Burushaski and its possible linguistic affiliation, see van Driem (2001).

10.1.10. Contact with intrusive languages
Within China, Putonghua is now used as a second language by a majority of Tibetans. The number of Tibetan speakers who master Chinese has increased during the last decades. For this reason, Chinese has a growing influence on some Tibetic languages.
and dialects, especially in the eastern regions. Even in Central Tibet, many people mix in their everyday speech Chinese vocabulary with Tibetan vocabulary and grammar. This has given rise to the so called ར་མ་ལུག་སྐད་Ramaluk-ka lit. 'half sheep-half goat language.'

Within the Tibetic areas in India and Pakistan, Hindi and Urdu have also become intrusive languages. They are dominant in the school system and there is a tendency to mix the Tibetic languages such as Spiti and Ladakhi, etc. with Hindi-Urdu particularly in the district headquarters. Nepali has also gradually become a dominant language not only in Nepal but also in the Indian State of Sikkim and is known among the Bhutanese elite.

Since the end of the twentieth century, the "hyper language" English, another Indo-European language, has become widespread among the elites of the Tibetic speaking communities in India, Pakistan, Nepal and Bhutan. It is also used as education medium in many schools of the area.

We have no precise data about the situation in Myanmar, but it is likely that the National language Burmese as well as languages spoken in the Kachin state such as Rawang put pressure on the Kham dialect spoken in some villages of northern Myanmar.

10.2. Tibeto-Burman and Sinitic

Tibeto-Burman and Sinitic form the "Sino-Tibetan" macrofamily. However, many debates have taken place among the linguists about the structure and the extension of ST. (See e.g. Thurgood & LaPolla, 2017.) Recent arguments by Zhang

---

4. Elites in Central Tibet, Bhutan, Sikkim and Ladakh began learning English to various degrees even in the beginning of the twentieth century.

5. Most scholars use the term 'Sino-Tibetan'. It would also be relevant to call it Sino-Tibeto-Burman. Such a label would be based on the three main literary languages that have been used to reconstruct this macrofamily: Old Chinese, Literary Tibetan and Literary Burman. Some authors have proposed alternative appellations to Sino-Tibetan such as Sino-Bodic (van Driem 1997), Trans-Himalayan (van Driem 2014) or even greater groupings such as Sino-Tibetan-Austronesian (Sagart 2001).

6. Most authors consider as Thurgood and LaPolla (2003) that "the Sino-Tibetan family consists of two major sub-groups: Sinitic and Tibeto-Burman." However, a minority of authors such as van Driem (2015) consider Sinitic as branch of Tibeto-Burman.
et al. (2019), Sagart et al. (2019), and Zhang et al. (2020) support the establishment of the Sino-Tibetan family. Recent research about the genome of populations in East Asia also seems to confirm the Sino-Tibetan hypothesis (Wang et al. 2021).

The Sinitic family only includes a dozen of mutually non-intelligible languages, which used to be called “Chinese dialects,” all derived from Old Chinese: Mandarin (Northern Chinese), Jin, Wu, Gan, Xiang, Hakka, Min-bei (Northern Min), Mindong (Eastern Min), Min-nan (Southern Min) and Yue (Cantonese). There are more specific dialects reported such as Pinghwa. (See e.g. Chappell 2006.)

On the other hand, the Tibeto-Burman family comprises about 400 languages. The TB family includes the following main branches: Tibeto-Himalayan, Bodo-Garo, Lolo-Burmese and Karenic. However, during the last thirty years, some scholars have proposed to add other branches such as Naga-Kuki-Chin, Qiangic, Tani (also called Mirish), Bai, Asakian and Nungic languages. There are many questions still regarding the classification of the Tibeto-Burman branch, especially at the level of the groups and subgroups.

Modern Tibetic languages are only remotely related to most other groups of the ST macrofamily such as Sinitic, Baric, Lolo-Burmese, Karenic, etc. but they are obviously more closely related to Qiangic, rGyalrongic and especially to Bodic languages. We can compare this situation with the position of English within the Indo-European macrofamily. English, as a Germanic language, is only very remotely related to Indo-Iranian, it is genetically (and geographically) slightly closer to Slavonic or Romance languages, but it is very closely related to the other Germanic languages, such as Dutch, German, Swedish, etc. As mentioned earlier, the subfamily of Tibetic languages is comparable in size diversity to the Romance languages. They also share with the latter the existence of an old written language which is closely related to their ancestor (such as Latin for Romance and Old Tibetan for Tibetic).

7. Some scholars mention a slightly lower number. Michailovsky gives the figure of 250 languages, while Matisoff mentions 250–300 languages. If we take into account the diversity of the Tibetic family, it is necessary to add at least fifty “languages” without mutual intelligibility.

Let us now illustrate in Chart X.1 below the degree of remoteness between Tibetic and Sinitic languages (see also the STEDT database). Their genetic kinship, which is accepted by most scholars, is not obvious for non-linguists. Indeed, modern Sinitic and Tibetic languages seem so different that it is hard to imagine any genetic relationship between the two groups. However, when we compare Old Chinese and Classical Literary Tibetan, some roots are obviously cognate as shown in the chart below. These roots are also cognates in most TB languages. We can also clearly see that modern Chinese (in this case Mandarin) has undergone an evolution which makes it look even more different from Literary Tibetan.

In the chart below, the reconstruction of Old Chinese and Middle Old Chinese below is based on Baxter & Sagart (2014).9

---

9. As noted by Guillaume Jacques (pers. comm. 2021): “although several scholars have proposed reconstructions of proto-Sino-Tibetan or Proto-Tibeto-Burman (Peiros & Starostin 1996; Matisoff 2003), they are highly controversial.” See also Fellner & Hill (2019).
### CHART X.1. – Comparison of cognates between Old Chinese, Middle Old Chinese and Literary Tibetan

<table>
<thead>
<tr>
<th>Lexicon</th>
<th>Old Chinese</th>
<th>Middle Old Chinese</th>
<th>Literary Tibetan</th>
<th>Modern Mandarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>fish</td>
<td>ṅa 魚</td>
<td>ŋo</td>
<td>ṅ NYA</td>
<td>魚 yú</td>
</tr>
<tr>
<td>five</td>
<td>C.ɲ'a? 五</td>
<td>ɲuX</td>
<td>ṅ LNGA</td>
<td>五 wǔ</td>
</tr>
<tr>
<td>I</td>
<td>ɲ'aj? 吾</td>
<td>ɲaX</td>
<td>ṅ'NGA</td>
<td>我 wǒ</td>
</tr>
<tr>
<td>four</td>
<td>s.ɦ[j]-s 四</td>
<td>sịH</td>
<td>ṣ BZHI</td>
<td>四 sì</td>
</tr>
<tr>
<td>dog</td>
<td>kroʔ, [k]ʷʰˤuʔ 犬</td>
<td>kuwX, khwenX</td>
<td>ṡ KHYI</td>
<td>狗 gǒu, 犬 quán</td>
</tr>
<tr>
<td>three</td>
<td>s[ɬ]um 三</td>
<td>sam</td>
<td>ṡ GSUM</td>
<td>三 sān</td>
</tr>
<tr>
<td>grass</td>
<td>[tsʰ]uʔ 草</td>
<td>tshawX</td>
<td>ṡ RTSWA</td>
<td>草 cāo</td>
</tr>
<tr>
<td>mouth</td>
<td>kʰ(r)uʔ 口</td>
<td>kuwX</td>
<td>ṡ KHA</td>
<td>口 kǒu</td>
</tr>
<tr>
<td>nine</td>
<td>k[j]uʔ 九</td>
<td>kjuwX</td>
<td>ṡ DGU</td>
<td>九 jiǔ</td>
</tr>
<tr>
<td>name</td>
<td>C.meŋ 名</td>
<td>mjien</td>
<td>ṡ MING</td>
<td>名 míngzì</td>
</tr>
<tr>
<td>eye</td>
<td>ɲ'ɪ&lt;ɾ&gt;ɲ[n]ʔ, C.m[r][u]k 目</td>
<td>ṅenX, mjuk</td>
<td>ṡ MYIG</td>
<td>眼睛 yǎnjìng</td>
</tr>
<tr>
<td>cold</td>
<td>C.ranŋ 涼</td>
<td>læŋX, ljæŋ</td>
<td>ṡ GRANG</td>
<td>冷 lèng</td>
</tr>
<tr>
<td>tongue</td>
<td>m₃.lat 舌</td>
<td>zet</td>
<td>ṡ LCE, LLIAGS</td>
<td>舌 shé</td>
</tr>
<tr>
<td>two</td>
<td>ni-s 二</td>
<td>juiH¹⁰</td>
<td>ṡ GNYIS</td>
<td>二 èr</td>
</tr>
<tr>
<td>poison</td>
<td>[d]ʼuk 毒</td>
<td>dowk</td>
<td>ṡ DUG</td>
<td>毒 dú</td>
</tr>
<tr>
<td>die</td>
<td>sịʔ 死</td>
<td>sịX</td>
<td>ṡ SHI</td>
<td>死 sǐ</td>
</tr>
</tbody>
</table>

10. Karlgren provides the following reconstruction for initial segment ʰiːz.
Apart from a handful of lexical cognates, the phonology and the grammar of Old Chinese and Classical Literary Tibetan are very distinct. The same could be said if we compare Karenic and Tibetic groups.

Classical Tibetan is more closely related with Classical Burmese and even with spoken Burmese than with modern Sinitic languages, as we can see from the following comparative chart.

**CHART X.2. – Comparison of cognates between Burmese and Tibetan in contrast with Chinese**

<table>
<thead>
<tr>
<th>Lexicon</th>
<th>Classical Burmese</th>
<th>Classical Tibetan</th>
<th>Chinese (Mandarin)</th>
</tr>
</thead>
<tbody>
<tr>
<td>fire</td>
<td>/mi/</td>
<td>མེ་ ME</td>
<td>火 huō</td>
</tr>
<tr>
<td>moon</td>
<td>/la/</td>
<td>བླ་ ZLA¹¹</td>
<td>月 yuè</td>
</tr>
<tr>
<td>fish</td>
<td>/nga:/</td>
<td>སྦ YA</td>
<td>鱼 yú</td>
</tr>
<tr>
<td>salt</td>
<td>/s'ɑ/</td>
<td>སྦ TSHWA</td>
<td>盐 yán</td>
</tr>
<tr>
<td>hand</td>
<td>/lak/ (lak)</td>
<td>ཕག་ PHAG</td>
<td>手 shōu</td>
</tr>
<tr>
<td>pig</td>
<td>/wak/</td>
<td>སྦ LAM</td>
<td>猪 zhū</td>
</tr>
<tr>
<td>road</td>
<td>/lam/ (lam)</td>
<td>སྦ LAM</td>
<td>路 lù</td>
</tr>
<tr>
<td>mother</td>
<td>/ame/</td>
<td>ཀ འ མ ཡ A.MA</td>
<td>母亲 mǔqīn</td>
</tr>
<tr>
<td>father</td>
<td>/aphe/</td>
<td>ཀ ཁ འ ཡ A.PHA</td>
<td>父亲 fūqīn</td>
</tr>
</tbody>
</table>

Despite these similarities in a few basic words, Burmese and Burmic languages are only remotely related to Tibetic and they also exhibit a lot of differences in their grammar. As we will see, within the TB family, Tibeto-Himalayan, rGyalrongic and Qiangic

11. In modern Tibetic languages the reflex of ZLA is /da/ or /dza/ but the form /la/ is also found.
12. Many modern Tibetic languages have now a suffix LAG.PA. However some languages have preserved the classical form without suffix.
languages share more features with the Tibetic languages. These three groups of languages have been directly in contact with the Tibetic languages for many centuries.

10.3. **Tibetic and Tibeto-Himalayan**

The term "Bodic"\(^\text{13}\) is a traditional label coined by R. Shafer to refer to a group of languages that is also called "Tibeto-Himalayan." It lumps together Tibetic languages with a number of languages essentially spoken in the southern Himalayas, in Nepal, Bhutan and India. Apart from Tibetic, Tibeto-Himalayan (hence TH) includes the following subgroups from West to East: *Western Himalayan* \(^\text{14}\) or Kinnauri (Uttarakhand and Himachal Pradesh, India; Western Nepal); *Central Himalayan* \(^\text{15}\) (Mid-western Nepal); *Tanangic* (Central Nepal); *Eastern Himalayan* or Kiranti (Eastern Nepal) as well as *Far eastern Himalayan* languages spoken in Bhutan, Sikkim, Arunachal Pradesh (India) and on the Tibetan Plateau.\(^\text{16}\)

"Tibeto-Himalayan" or "Bodic" is a rather "heterogenous and impressionistic" group (van Driem 2001) and some scholars such as R. LaPolla or G. Jacques (pers. comm.) think that it should be abandoned altogether. The validity of the TH grouping has not been proven so far. Many authors such as Matisoff, LaPolla, van Driem or more recently Hyslop (2013) have called for a bottom up approach in the classification before making hypothetical groupings.

The term "Bodish," also coined by Shafer, was intended to designate the "Tibetan dialects" as well as the languages most closely affiliated with "Tibetan," within the TH branch. In its original meaning, Bodish comprises of the Central Bodish (which corres-

---

\(^\text{13}\) The label "Bodic" and "Bodish" derive from the root *bod* meaning 'Tibet' in Classical Tibetan. Both terms are misnomers since the Bodic and Bodish languages include a lot of languages which do not use a word derived from the root *bod* to designate Tibet. For example, in Byangshi the word for Tibet is /kiday/, in Rongpo (or Garhwal) /byan/, etc. Thus the label "Tibeto-Himalayan" is a preferable.

\(^\text{14}\) This group is often referred to as West Himalayan (see van Driem 2001).

\(^\text{15}\) It includes languages such as Kham, Magar and possibly Newari. These languages are sometimes grouped together with the Kiranti languages in a "Mahākiranti" group.

\(^\text{16}\) E.g. Tshangla, Dukpa, Lhokpu, Gonduk, Lepcha, Bhumthang, Dakpa, etc.
ponds to our Tibetic], South Bodish and West Bodish,\textsuperscript{17} and a series of languages spoken in Bhutan and Arunachal Pradesh called East Bodish. According to van Driem (2001: 828), “Bodish is a more well-defined subgroup whereas Bodic is intended as a more tentative and loosely formulated heuristic set.”

Although “Bodish” is a “more well-defined subgroup” and is less controversial than “Bodic” (or “Tibeto-Himalayan”), it is also problematic. First, various authors have proposed different definitions of Bodish. For example, Bradley (1997) incorporates West Himalayish and Tamangic as branches of “Bodish.” Michailovsky and Mazaudon (1994) consider that the Tamangic languages or TGTM (Tamang-Gurung, Thakali, Manangha group) “belongs in the Bodish section but outside the Bodish Branch.” For van Driem (2001), the label “Bodish” does not include the Tamangic languages.

As was shown by Shafer (1955) as well as Michailovsky and Mazaudon (1994), East Bodish languages are clearly not “Tibetan dialects” (see also Hyslop 2011, 2013). They exhibit more conservative features than the “Tibetan dialects.” Hyslop (2013: 3) noted correctly that

“a confident placement of the East Bodish languages with regard to Tibetan [Tibetic] and within Tibeto-Burman is impossible at the present time. There has obviously been considerable influence of Chöke (Classical Tibetan, the liturgical language of Bhutan) and Dzongkha on the East Bodish languages in Bhutan and probably in Arunachal Pradesh as well.”

Thus in its restricted sense, “Bodish” is equivalent to the grouping of “Tibetic” (in our sense) with “East Bodish” languages.

If we look at the languages that have the closest genetic affiliations with Tibetan, we find not only the East Bodish languages, but also the Tamangic languages, spoken in the southern Himalayas, as well as some languages whose classification is not entirely clear such as Basum (spoken in Kongpo, TAR) or Tshona Monpa, both spoken on the Tibetan Plateau.

\textsuperscript{17} Shafer’s Central Bodish comprises Ü-Tsang, Amdo, Khams but also some “dialects” of the southern Himalayas such as Sherpa and Jirel; South Bodish refers to Dzongkha, Lhoke, Dromo; West Bodish corresponds to Balti, Ladakhii (West archaic group) as well as Spiti, Garsha and the Western part of Ngari.
Other TH subgroups such as West Himalayish or Kiranti are not as closely related to Tibetic. Although they do share a lot of cognates, they differ in their basic vocabulary and grammar. For example, both West Himalayish and Kiranti languages are pronominalized languages, i.e. they exhibit a verb "agreement" or rather an argument indexation, which is absent in all the Tibetic languages.

**Illustration of some basic lexical discrepancies within “Tibeto-Himalayan”**

TH languages share a great deal of cognates as we show below with examples taken from Kiranti, West Himalayish languages and Classical Tibetan (CT). We provide below the correspondences in the various languages as well as the Proto-Tibeto-Burman (PTB) reconstructed forms. Let us first look at some lexical items, which at first glance would advocate for a close relationship:

- 'eye': /mik/ (Limbu), /mig/ (Rongpo), མིག་MIG (CT), *s-myak (PTB)
- 'die, to': /si(-ma)/ (Limbu), /hi(-ci-mo)/ (Byangsi), ཤི་SHI (CT), *søy (PTB)
- 'drink, to': /thung/ (Limbu), /tun(-mo)/ (Byangsi), ཐུང་THUNG (CT), *doŋ (PTB)
- 'fire': /mi/ (Limbu), /mhe/ (Rongpo), མེ་ME (CT), *mey (PTB)
- 'fish': /ŋa/ (Limbu), /hnya/ (Darma), ཉ་NYA (CT), *ŋya (PTB)
- 'five': /ŋasi/ (Limbu), /ŋe/ (Rongpo), ཐིང་SHING (CT), *ŋ-bŋasi (PTB)
- 'house, home': /him/ (Limbu), /cim/ (Byangsi), dim (Tamang), ཨོི་KHYIM (CT), *kyim (PTB)
- 'louse': /siɁ/ (Limbu), /rhi:g/ (Rongpo), སྒི་SHIG (CT), *s-r(y)ik (PTB)
- 'meat': /sa/ (Limbu), /ʃa/ (Rongpo), སྤི་SHI (CT), *a(ya) (PTB)
- 'wood, tree': /siŋ/ (Limbu), /ciŋ/ (Darma), གཅིག་GCIG (CT), *g-t(y)ik (PTB)

18. Concerning Limbu (a Kiranti language), the data are taken from Michailovsky (2002). Concerning Rongpo, Byangsi, Darma and Chaundangsi the data are taken respectively from Sharma (2001a-b) and Shree Krishan (2001a-b).

19. Limbu is a Kiranti language while Rongpo, Byangsi, Chaundangsi and Darma are West Himalayish languages.
‘month, moon’: /laba/ (Limbu), hla (Chaudangsi), གྲ་བ་ ZLA (CT), *s/-ła (PTB)
‘two’: /ni/ (Limbu), /nhis/ (Rongpo), གཉིས་ GNIS (CT), *g/-nis (PTB)
‘three’: /sum/ (Limbu), /sum/ (Rongpo), གསུམ་ GSUM (CT), *g/-sum (PTB)
‘four’: /lisi/ (Limbu), /bi/ (Basum), /‘bli/ (Tamang) Kheng /ble/, བཞི་ BZHI (CT), *b/-lay (PTB)
‘silver’: /mul/ (Rongpo), /mal/ (Byangsi), དངུལ་ DNGUL (CT)

Even when the words are obviously cognate, their form does not comply with the phonological evolution of the Tibetic languages. When we look at the cognates mentioned in the list above, we could already determine that Limbu is not a Tibetic language because in some cases, the cognates are not reflexes of Classical Tibetan but bear more archaic features. As we showed in Chapter 4, one of the important innovative features of Old Tibetan was the palatalization of dental and sibilant consonants: *ty > SH and *ty > CH. The Limbu words /siʔ/ ‘louse’, /si(-ma)/ ‘die’, /siŋ/ ‘tree’ (wood), /sa/ ‘flesh’, /thik/ ‘one’ have not undergone the palatalization that occurred in CT: SHIG ‘louse’, SHI ‘die, to’, SHING ‘wood, tree’, SHA ‘meat’ and GCIG ‘one’ and in all the languages derived from it. Concerning the fricative SH, Amdo and some Kham dialects have undergone a change into a velar and pronounce SH /x/ or /ɬ/, but no Tibetic language has preserved a dental fricative /s/.

However, the series of cognates listed above could indeed give the wrong impression of a close relationship between the above TH languages (Kiranti or West Himalayish) and Tibetic, whereas it is in fact a relatively remote relationship. First, some of the words listed above such as ‘eye’, ‘louse’, ‘one’, ‘wood’ are not specific cognates and are found beyond the TH branch in many TB languages (as shown by the comparison with Proto-Tibeto-Burman) and even Old Chinese (Matisoff 2003: 347). Second, the vast majority (up to 90%) of basic lexical items found in Kiranti and West Himalayish are in fact not cognate with CT. Let us illustrate this point with basic lexical items:

20. However, we find a sound correspondence of SH with a dental apical fricative in a marginal area of Kham.
'blood': /makkhi/ (Limbu), /vi/ (Darma), KHRAI (CT)
'brain': /ñhi/ (Limbu), /tanu/ (Byangsi), KLADPA (CT)
'flower': /phun/ (Limbu), /ce/ (Byangsi), /búr/ (Lepcha), METOG (CT)
'gold': /sammyaŋ/ (Limbu), /jan/ (Darma), GSER (CT)
'horse': /sn/ (Limbu), /hran/ (Byangsi), RTA (CT)
'leg': /lan/ (Limbu), /lude/ (Darma), RKANGPA (CT)
'medicine': /sidaŋ/ (Limbu), /wo-so/ (Chaudangsi), SMIAN (CT)
'stone': /luŋ/ (Limbu), /un/ (Rongpo), RDO (CT)
'iron': /phendže/ (Limbu), /na-jhán/ (Chaudangsi), LCAGS (CT)
'milk': /nu/ (Limbu), /nü/ (Byangsi), /cer/ (Lepcha), OMA (CT)
'red': /he/ (Limbu), /man-ju/ (Darma), DMAR (CT)
'yellow': /hik/ (Limbu), /hleda/ (Chaudangsi), SER (CT)
'white': /phö/ (Limbu), /áída/ (Rongpo), DKAR (CT)
'seven': /nus/ (Limbu), /hnis/ (Chaudangsi), /nif/ (Byangsi), BDUN (CT)

What has been told about Kiranti and West-Himalayish is also true for other TH languages such as Lepcha. Let us give a few other examples:

'throne': /bhi/ (Lep), KHR (CT)
'law': /him/ (Lep), KHRIMS (CT)
'market': /hom/ (Lep), KHRM (CT)
'iron': /punjeng/ (Lep), LCAGS (CT)
'leg': /pán/ (Lep), RKANGPA (CT)
'red': /phuyur/ (Lep), DMARP (CT)
'white': /págü/ (Lep), DKAR (CT)
'blood': /vi/ (Lep), KHRAI (CT)

---

21. The classification of Lepcha is still not entirely clear and there is no consensus about its precise genetic affiliation within Tibeto-Himalayan. The data are based on Plaisier (2007).
‘flower’: /búr/ (Lep), རྨེ་ཏོག་ (CT)

When we examine the Lepcha, we can see that the first words are very similar to their equivalents in Classical Tibetan. This is due to the fact that they have been borrowed many centuries ago. In fact, words such as ‘throne’, ‘market’ and ‘law’ are good candidates for loanwords, because they are related to power or have an economic, cultural or religious value. All the other words (‘iron’, ‘leg’, ‘red’, ‘yellow’, ‘white’, ‘blood’, ‘flower’) are clearly not cognate with CT. This situation is typical of many non-Tibetic languages spoken on the Tibetan Plateau and in the Himalayas, where we find loanwords with a pronunciation reflecting the old Tibetan pronunciation and indigenous basic vocabulary which is entirely different.

As we can easily see, the words of the various TH languages are not cognate with CT words. In Limbu, about 90% of words are in fact not cognate with their CT equivalent.

On the other hand, all the modern Tibetic languages have direct reflexes of the above CT words.

It is obvious from the above list that the basic vocabulary of these TH languages is not related to Classical Tibetan. To the speakers of Tibetic languages, these TH languages are perceived as alien and relatively difficult to learn since they also exhibit significant differences in their morphology and syntax.

Languages closely related to the Tibetic family

The Tamangic and East Bodish languages, spoken in Nepal and Bhutan, as well as a few other languages spoken in Tibet, such as Dakpa and Basum, show a greater proximity to the Tibetic languages and Classical Literary Tibetan than the rest of the TH languages.

This proximity is not only genetic but due to the fact that they have borrowed a great number of words to the neighboring Tibetic languages and to Classical Tibetan which is in many cases used as a liturgical language of Tibetan Buddhism and Bön.

22. The Tamangic languages include various Tamang “dialects,” Gurung, Thakali, Manangi, Nar-Phu, Chantyal and Kaike (the affiliation of Kaike to the Tamangic branch is debated).
Here are some examples of the lexical proximity between Kurtö (East Bodish) and Tibetic. We provide here both CT and Lhasa correspondences for the East Bodish examples below, adapted from Hyslop (2011: 252).

'tiger': /tɑ:/ (Kurtö), /taʔ/ (Lhasa), བླ་ STAG (CT)
'dragon': /dju:/ (Kurtö), /djuʔ/ (Lhasa), ཇུ་ BRUG (CT)
'pig': /p’a:/ (Kurtö), /’p’aþpa/ (Lhasa), སྒྲུབ PHAG.(PA) (CT)
'hammer': /t’éfono/ (Kurtö), /t’ówe/ (Lhasa), སུལ་ THO.BA (CT)
'bridge': /zām/ (Kurtö), /sampa/ (Lhasa), བླ་ ZAM.(PA) (CT)
'saddle': /ga/ (Kurtö), /’ga/ (Lhasa), བླུ་ SGA (CT)
'drum': /nèa/ (Kurtö), /’nþa/ (Lhasa), བླུ་ RNGA (CT)
'brain': /klatpa/ (Kheng), /klatpa/ (Kurtö), /’läþpa/ (Lhasa), བླུ་ KLA.D.PA (CT)

The reflexes of CT in Kurtö are sometimes pronounced in a very similar way to the Lhasa or Dzongkha pronunciation, but in some words, they reflect an ancient pronunciation of CT or OT, which probably indicates that they are loanwords. This is for example the case for the word 'brain'.

The lexical similarities extend to the honorific vocabulary as well:

<table>
<thead>
<tr>
<th>Kurtö Ordinary</th>
<th>Kurtö Honorific</th>
<th>CT Ordinary</th>
<th>CT Honorific</th>
</tr>
</thead>
<tbody>
<tr>
<td>/luspu/</td>
<td>/kuzu/</td>
<td>བླུ་ LUS.PO</td>
<td>བླུ་ SKU.GZUGS</td>
</tr>
<tr>
<td>/mi/</td>
<td>/cen/</td>
<td>བླུ་ MIG</td>
<td>བླུ་ SPYAN</td>
</tr>
<tr>
<td>/sha/</td>
<td>/kusha/</td>
<td>བླུ་ SHA</td>
<td>བླུ་ SKU.SHA</td>
</tr>
<tr>
<td>/neng/</td>
<td>/thuk/</td>
<td>བླུ་ SNYING</td>
<td>བླུ་ THUGS</td>
</tr>
</tbody>
</table>

24. The data is from Hyslop (2011: 252).
25. The phonetic realization is usually devoiced or half devoiced.
It should be noted that the existence of numerous loanwords and common etyma does not entail that the basic terms are always cognates. In fact, as we will see, there is still a significant linguistic gap between these TH languages (Tamangic, East Bodish, Basum, Dakpa, etc.) and the Tibetic languages not only in their vocabulary but also in their phonology and grammar.

Let us illustrate some of the lexical differences in the basic vocabulary:

'red': /zhinti/ (Kurtö), /dənde/ (Basum), བདེ་ནད་ DMAR.PO (CT)
'to come': /ra/ (Kurtö), /rɔ/ (Bumthang), /rɔ/ (Basum), རོ་ YONG’ONG (CT)
'milk': /ʒu/ (Kurtö), /dʒu/ (Kheng), ཇུ་ OLMA (CT)
'tooth': /kwa/ (Kurtö), གྲེས་ SO (CT)
'leg': /tawa/ (Kurtö), /ci/ (Basum) གནས་ RKANG.PA (CT)
'to know': /khan/27 (Kheng), མེ་ཤེས་ SHES (CT)
'seven': /ni/ (Dakpa), /nɔ/ (Bumthang), /ni/ (Basum), ཉི་ BDUN (CT)
'water': /khwe/ (Kurtö), /khwi/ (Bumthang), ཇུ་ CHU (CT)

In all the modern Tibetic languages, the basic words listed above are regular reflexes of the CT forms, whereas it is not the case in the East Bodish and Tamangic languages.

Finally, let us mention another "Bodish" language very close to Tibetic, Basum, which is spoken by 3,000 ethnic Tibetans in various villages around Basum Lake, in the Kongpo region of the TAR, about 400 km east of Lhasa. This language was classified by Qu (1996) as a "central Tibetan dialect," but we will show that according to our definition, Basum is not even a "Tibetan dialect" (or a Tibetic language). Basum-speaking people are well aware that their language is different from the surrounding Tibetan dialects, and call their language a Dākinī’s language (for more information about Basum, see section 10.6 as well as Wang 2020). Most of Basum modern vocabulary is cognate with Tibetan or borrowed from it. Let us compare a list of basic words in Basum and CT:

27. This word may be cognate with MKHAN’ expert’ and MKHYEN’ to know’ (Honorific), but the pandialectal word for the verb ‘to know’ is SHES or *SHE in Tibetan.
<table>
<thead>
<tr>
<th>Basum</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>'die, to'</td>
<td>'SHI'</td>
</tr>
<tr>
<td>'drink, to'</td>
<td>'THUNG'</td>
</tr>
<tr>
<td>'eye'</td>
<td>'MIG'</td>
</tr>
<tr>
<td>'fire'</td>
<td>'ME'</td>
</tr>
<tr>
<td>'fish'</td>
<td>'NYA'</td>
</tr>
<tr>
<td>'five'</td>
<td>'LNGA'</td>
</tr>
<tr>
<td>'four'</td>
<td>'KZHI'</td>
</tr>
<tr>
<td>'house'</td>
<td>'KHYIM'</td>
</tr>
<tr>
<td>'louse'</td>
<td>'SHIG'</td>
</tr>
<tr>
<td>'pig'</td>
<td>'PHAG'</td>
</tr>
<tr>
<td>'one'</td>
<td>'GCIG'</td>
</tr>
<tr>
<td>'road'</td>
<td>'LAM'</td>
</tr>
<tr>
<td>'wood'</td>
<td>'SHING'</td>
</tr>
<tr>
<td>'two'</td>
<td>'GNYIS'</td>
</tr>
<tr>
<td>'needle'</td>
<td>'KHAB'</td>
</tr>
<tr>
<td>'iron'</td>
<td>'LCAGS'</td>
</tr>
<tr>
<td>'moon'</td>
<td>'ZLA(BA)'</td>
</tr>
<tr>
<td>'black'</td>
<td>'NAG.PO'</td>
</tr>
<tr>
<td>'yellow'</td>
<td>'SER.PO'</td>
</tr>
</tbody>
</table>

Two observations: first, in the list above, Basum and CT words are clearly cognates. However, Basum words are not regularly derived from CT and do not follow the
phonological rules that normally apply to Tibetic languages. For example, they do not exhibit a palatalization of dental and sibilant consonants in the above words 'to die', 'one', 'wood', 'louse' or 'iron'. Another striking feature is the lack of aspiration for certain words such as 'pig', 'house' and 'needle'. This aspiration is an important phonological feature of the Tibetic languages, and although some rare words such as ཟྭ མྱི་ KHYI ‘dog’ may have lost the aspiration in some languages, this phenomenon is rarely attested. Three words of the list, 'black', 'yellow' and 'moon' could be considered as problematic, but they actually perfectly match their equivalents in the surrounding Tibetic languages. The words / naːnæ/ and / sese:/ are not directly derived from CT NAG.PO and SER.PO but from the reduplicated stem NAG.NAG, and SER.SER, a phenomenon which is frequently attested in the neighboring Kham languages, particularly for color adjectives.

The word /daːɡa/ is also not directly derived from CT ZLA.BA but from a frequent compound used in Tibet ZLA.DKA.R ‘white moon’, and the Basum term for ‘moon’ perfectly matches the regular reflexes of this compound.

In brief, Basum equivalents are all clearly cognates of CT or derived from original compounds but their phonology does not always correspond to the expected regular reflexes. This alone suggests that there is a problem with the affiliation of Basum as a Tibetic language.

The numbers ‘seven’ and ‘four’ often allow a quick checking of the Tibetic affiliation. In Basum, both numbers do not correspond to Tibetic reflexes. As we will see, a more careful examination of the vocabulary and grammar confirms this hypothesis. Some basic words clearly indicate the existence of a distinctive substratum.

---

29. E.g., Dingri (Tö), Sherpa and Chagthreng (Kham).
In the above list, none of the Basum words are cognate with CT, whereas all these words are pandialectal in the Tibetic family. Moreover, if we consider the grammatical words, we can also realize the distance between Basum and the Tibetic languages.

Again, all these grammatical words are not attested in the Tibetic languages.

The case of Dakpa (Tshona Mönpa), also spoken in the TAR, bears resemblance with the case of Basum. It is interesting to note that Basum speakers are ethnic Tibetans while Dakpa speakers are ethnic Mönpa. This shows once again that language, ethnicity and nationality do not necessarily match, but within the Chinese political system, the determination of the "nationality" is often arbitrary and Dakpa could have been classified as Tibetan, and conversely Basum could have been treated as a separate minzu "nationality."
Both languages are so closely related to Tibetic that some sentences sound entirely Tibetan and could easily be written in Tibetan script following in many cases the Classical orthography. Let us give some examples of Dakpa (Lu Shaozun 1986), with our transcription in Tibetan script and transliteration:

(368) བཀྲ་ཤིས་ཡི་གེ་བྲིས་བོ་ནེད་
**BKRA-SHIS YIG BRIS-BO.NED**

\[
\text{[tsA l jícI pìl wo ne̞M]} \quad \text{Pn} \quad \text{letter write-SUFF AUX}
\]

‘Trashi has written the letter’

(369) བུ་ཚ་བོད་ཡིག་མ་ཁན་ནི་
**BU.TSA BOD.YIG MA-KHAN-NI.NED**

\[
\text{[prIcA pøji mA khanl ni ne̞M]} \quad \text{child} \quad \text{Tibetan script} \quad \text{NEG can AUX}
\]

‘The child can not (read) the Tibetan script’

(370) དོ་བུ་ཚ་ཀོ་ཡི་གེ་ཡིན་ཏེ་
**DZO BU.TSHA-KO YIG YIN.TE**

\[
\text{[tsO prIcA-køl jícI yinheI]} \quad \text{this child REL book AUX}
\]

‘This is the child’s book’

(371) སོ་ག་ཏོ་ཀི་ར་བོ་ནེད་
**PE GA-TO-KI RA-BO.NED**

\[
\text{[peI kA hoI kI ra-wøI ne̞M]} \quad \text{3SG where REL come-suff AUX}
\]

‘Where does he come from?’

The first sentence (368) really corresponds to a sentence of a Tibetic language (central Tibet), since all the lexical items are derived from CT and only the origin of the auxiliary [ne̞?] is not clear. The second sentence (369) has a form [khanl] ‘to be able’ which is not found in CT. In (370), the demonstrative [tsO] (also attested in Basum /tsu/) is used instead of the Tibetic ‘DI, DE or GAN, and the genitive for [ko] which is not a regular reflex of the CT genitive. In the last sentence (371), all the words, whether lexical or grammatical do not have obvious Tibetic cognates: [pe] the third
person singular (3s) corresponds to KHO in CT, [ra] 'to come' instead of YONG/ONG ‘to come’ in CT.

We will examine more in details in the section 10.5 various criteria to identify Tibetic languages and distinguish them from other closely related languages.

10.4. Tibetic, rGyalrongic and Qiangic

rGyalrongic and Qiangic languages have preserved a more archaic morphology and have maintained nominal and verbal prefixes (instead of suffixes) as well as verb agreements which makes the genetic relationship to Tibetic less obvious than in the case of Tamangic or East Bodish.

Here are examples of obvious cognates in Japhug and bTshanlha, some of them exhibiting a very archaic pronunciation close to Classical Tibetan, which shows that in some cases, these may be loanwords.

- ‘meat, flesh’: [ɕa] (Japhug), [ʃa] (bTshanlha),SHA (CT)
- ‘moon, month’: [sla] (Japhug), [zlawa] (bTshanlha), ZLA (BA) (CT)
- ‘eye’: [ry-mɲaɾ] (Japhug), [ry-mɲak] (bTshanlha), MIG (CT), DMYIG (Proto-Tibetic)
- ‘pig’: [paɾ] (Japhug), [pak] (bTshanlha), PHAG (CT)
- ‘three’: [xum] (Japhug), [ka-som] (bTshanlha), GSUM (CT)
- ‘two’: [ɾnus] (Japhug), [ka-ɲis] (bTshanlha), GNYIS (CT)
- ‘black’: [ku-ɲaɾ] (Japhug), NAG (PO) (CT)

30. Qiangic languages include Rmaic (Qiang), rGyalrongic, Minyag (Darmdo Minyag and Shimian Minyag), Prinmi (Pumi), nDrapa, Choyu (Queyu), Lhagang Choyu, Namnyi/Namzi, nGochang (Guiqiong), Shihing (Shixing), Ersu, Donu, Luzu, Lamo, Larong sMar, and Drag-yab sMar as well as an extinct language, Tangut. According to most authors (Sun Hongkai 1983; Sun T.S. 2000; Jacques 2004a). Qiangic also comprises rGyalrongic languages (see Sun T.S. 2000); Situ, bTshanlha, Japhug, Tshobdun, Zbu, sTau (RTA), Geshita, Nyagrong Minyag, sTodsde (Shangzhai), and Khroskyabs (Lavrung). The Japhug examples are taken from Jacques (2004a), the Basum examples are Tournadre’s data and the reconstructed PTB (Proto Tibeto-Birman) are taken from Matisoff (2003).

31. Two rGyalrongic languages spoken in Sichuan, Japhug is based on the Kannyo dialect, while bTshanlha is based on the Bago dialect.
rGyalrongic and Qiangic languages have preserved ancient pronunciations derived from Old Tibetan that are lost in the modern Tibetic languages. For example, in the above bTsanlha words (rGyalrongic), the words for 'iron' /ltʃaks/ and 'rice' /mbras/ are still pronounced as they were pronounced in Old Tibetan and it exactly matches the classical Tibetan orthography, respectively ལྕགས་ LCAGS and འབྲས་ 'BRAS. 32 rGyalrongic languages such as bTsanlha are also the only languages which have preserved Tibetan words with four initial consonant clusters such as བརྒྱུད་བསྒྲགས་ BRGYUD-BSGRAGS /brɟt-bzgraks/ 'to pass on, transmit' or བསྒྲགས་ཡིག་ BSGRAGS-YIG /bzgravg-yik/ 'handbill, flyer'. There are a great number of old Tibetan loanwords in rGyalrongic languages and in Qiangic languages. As in the case of East Bodish languages, it is not always easy to distinguish between loanwords and inherited vocabulary.

In many cases, the roots are similar or identical to CT but they differ in the affixes: rGyalrongic languages make use of prefixes (such as /ka/, /ta/, etc.) whereas CT and Tibetic languages use suffixes (such as PA/BA, MA, PO/BO, MO, etc.):

'to die': /kɤ-si/ (Japhug), /ka-fi/ (bTsanlha), ཤི་བ་ SHI-BA (CT)
'to arise': /ka-far/ (bTsanlha), ཤི་བ་ SHAR-BA (CT)
'to pass away' (H): /ka-ʃaks/ (bTsanlha), བཤེགས་པ་ GSHEGS-PA (CT)
'to stand up' (H): /ka-bʒes/ (bTsanlha), བཞེངས་པ་ BZHENG-PA (CT)
'food': /ta-zə/ (bTsanlha), རང་ ZALMA (CT)
'mind': /ta-šems/ (bTsanlha), སེམས་ SEMS-PA (CT)

32. In bTsanlha, 'iron' /ltʃaks/ is clearly a loanword. The native word for 'iron' is /ʃom/.
A significant part of the basic lexicon in rGyalrongic and Qiangic languages is not cognate with its Tibetic equivalents as we see from the following examples:

- 'well, nice': /kɯ βdi/ (Japhug), BDE.PO (CT)
- 'year': /tə pa/ (bTsanlha), LO (CT)
- 'brain': /tərnok/ (bTsanlha), SER.BA (CT)
- 'milk': /təlu/ (bTsanlha), OMA (CT)
- 'leg': /tame/ (bTsanlha), RKANG.PA (CT)
- 'fish': /qa-ɟy/ (Japhug), ηʃho (bTshanlha), NYA (CT)
- 'flower': /tapat/ (bTsanlha), ME.TOG (CT)
- 'tongue': /teʃme/ (bTshanlha), LCE (CT)
- 'red': /kəwɔrm/ (bTsanlha), kɯ-ɱumi (Japhug), DMAR.PO (CT)
- 'yellow': /kɯ-qarŋe/ (Japhug), SER.PO (CT)
- 'sand': /kawe/ (bTsanlha), BYE.MA (CT)
- 'sheep': /qa-zo/ (Japhug), kəjo/ (bTsanlha), LUG (CT)
- 'horse': /məro/ (Japhug), məro/ (bTsanlha), RTA (CT)

Again, for all the above words, the Tibetic languages have regular reflexes of the CT forms.

Recent studies on rGyalrongic and Qiangic languages have shown that these languages differ drastically in their grammars from the Tibetic languages (see e.g. Sun 2014b; Jacques 2004a; LaPolla 2003).

To illustrate the lexical and grammatical differences, let’s give some more examples. Wang & bTsanlha Ngag.dbang Tshul.khrims (1992) provides the following comparison between Amdo Tibetan and bTsanlha rGyalrong in Tibetan script:

---

34. The author presents the example as Amdo however, the absence of verb auxiliary would rather lead to consider that it is a form of Literary Tibetan.
Tibetan:
(372) རྟ་ཐོག་ནས་མར་ལྷུང་
RTA THOG-NAS MAR LHUNG
horse on-ABL down fall
’S/he fell from the horse’.

rGyalrong:
(373) འབྲོ་འོ་གོ་སུ་ནུ་ཞའ་
'BRO GO-SU NU ZHA'
horse on-ABL down fall
’S/he fell from the horse’.

In the above sentences, we can easily see the great discrepancies between rGyalrong and Tibetan in the vocabulary ‘BRO’ horse’ (\(\text{m}^\text{bro}^{}/ \text{m}^\text{boro}^{}/\) depending on the rGyalrong dialects), ZHA’ to fall’, as well as in the ablative case –SU and the adverb NU ‘down’. The corresponding forms for ‘horse’ and ‘down’ are respectively cognate of RTA and MAR in all the modern languages.

Among the most striking grammatical differences between Tibetic and rGyalrongic (or more generally Qiangic languages), we have mentioned the “verb agreement” or “argument indexation.” Such systems of pronominal or argument indexations are attested in other branches of TB family such as West-Himalayish, Kiranti, etc. Here are examples of rGyalrong conjugations provided by BTSAN.IHA NGAG.DBANG TSHUL.KHRI'M (ibid.) in Tibetan script:

(374) ལྷུང་གིས་འཚིལ་
NGA-GIS GO-RTSING
1SG-ERG PREF- count+1SG
‘I am counting’

(375) དེ་འཚིལ་
NYO-GIS GO DE-RTSIN
2SG-ERG PREF- count+2SG
‘You are counting’
The verb རྩི་ ‘to count’ as other rGyalrong verbs has several forms depending on the persons (additionally they may differ also in their prefixes): རྩིང་ ‘1sg’, རྩིན་ ‘2sg’, རྩི་ ‘3sg’.

It is interesting to note that the verb རྩི་ ‘to count’ is attested in the Tibetic languages and CT. However, neither Classical or Old Tibetan, nor the modern languages have developed any agreement of this sort.

The evidential systems, which have developed egophoric forms (see Chapter 8) may not be considered as agreement or pronominal. In fact, some rGyalrongic languages which have argument indexation in their verb morphology have also developed egophoric forms and the two phenomena may coexist as shown by Jacques (2019).

10.5. Languages of the Tibetosphere

We have seen in the above sections (10.2 and 10.3) that Tibeto-Himalayan languages as well as rGyalrongic and Qiangic languages are not only closely related to Tibetic in various degrees, but have borrowed a lot of their vocabulary from CT or Tibetic languages. Furthermore, other languages belonging or not to the Sino-Tibetan stock have also been influenced in their lexicon, their prosody and even sometimes in their grammar by the neighboring Tibetic languages.

Whether in the southern Himalayas or on the Tibetan Plateau, the area over which the Tibetan cultural and linguistic influence is manifest can be called the “Tibetosphere.” This term has been coined on the model of Matisoff’s Indosphere and Sinosphere. In his article entitled ‘Megalocomparison’ that targeted Greenberg’s work, Matisoff (1990) proposed these terms to refer to “linguistic/cultural influence in Southeast Asia,” i.e. linguistic and cultural influences from Chinese on the one hand.
and from Indo-Aryan on the other hand. However, he overlooked the existence of a third "sphere," the Tibetosphere that emerged at the moment of the Tibetan Empire (seventh century A.D.). Various authors, such as Noonan (2003), DeLancey (2012) and Tournadre (2014a) have used this notion related to "areal typology" which allows explaining some lexical and grammatical convergences (such as the existence of evidential and epistemic systems, the development of light verb constructions, etc.) within this area, independently of the genetic relationship.

Let us first illustrate some Tibetospheric languages spoken in China. There are sixty non-Tibetic languages spoken within the Tibetic speaking area (see Roche & Suzuki 2018). In addition to this, at least four varieties have recently been "discovered" (10.7.6, see also Tashi Nyima & Suzuki 2019, for details). Forty-eight of these languages are located in Eastern Tibet (Sichuan, Qinghai, Gansu and Yunnan).

Here is a map showing their distribution (originally published in Roche & Suzuki 2017).

---

36. When he conceived these two areas, the author had certainly in mind a long historical period and a typological approach which lead to the idea of "areal typology."
MAP X.1. – Minority languages of the Eastern Tibetosphere

1 Western Yugur
2 Eastern Yugur
3 Mongghul
4 Kangjia
5 Salar
6 Mangghuer
7 Manegacha
8 Ngandehua
9 Heran Orat
10 Baihua
11 Northern Rmaic
12 Situ
13 Japhug
14 Tshobdun
15 Zbu
16 Stotsde
17 Khroskyabs
18 Southern Rmaic
19 Southern Rgyaltrong
20 Geshhta
21 Rtsu
22 Nyagrong Minyag
23 Choyu
24 nDrapa
25 Lhagang Choyu
26 Gochang
27 Dachua
28 Darmo Minyag
29 Shiman Minyag
30 Erhu
31 Nosu
32 Doxu
33 Namyu
34 Lizu
35 Pinni
36 Shuhang
37 Laze
38 Na
39 Naxi
40 Malimasa
41 Nung
42 Trung
43 Lu
44 Lama (Ba)
45 Southern Pinni
46 Tanglang
47 Lipo
48 Tuku

Cartography: Chandra Jayasuriya. Language data: Gerald Roche and Hiroyuki Suzuki
The presence of many non-Tibetic languages on the Tibetan plateau shows that the extraordinary linguistic diversity of this area is not limited to the languages derived from Tibetan. These non-Tibetic languages belong to Sino-Tibetan, Mongolic and Turkic language families. The situation is very complex due to the various ethnic affiliations and administrative statuses of these languages.

The situation of rGyalrongic and Qiangic languages deserve a special mention. Speakers of these languages are now Chinese citizens and all the Gyälrongwas (རྒྱལ་རོང་བ་) are ethnic Tibetans and officially recognized as "Tibetans" (བོད་རིགས, 藏族) by the Chinese government. For the speakers of Qiangic languages, the situation is more complex (see Chapter 2). Officially, they are either of the Tibetan nationality or belong to the Qiang and Pumi nationalities. Many speakers of Qiangic or rGyalrongic languages can master along with their native language an Amdo or a Kham variety as well as Classical Tibetan. They also often know Standard Mandarin Chinese or a Chinese dialect from Sichuan or Qinghai.

The Tibetans who are native speakers of rGyalrongic or Qiangic languages share with the neighboring Amdowa and Khampa the same fundamental Tibetan culture, even if from a linguistic point of view their languages are very different from the neighboring languages of Amdo and Kham (as we have seen above in sections 10.3 and 10.4). For this reason, some recognized Tibetan scholars, such as SUM.BHANDON, GRUB TSHE.RING (2011: 50-51) or Wang & BTSAN.LHA NGAG,DBANG TSHUL.KHRI.MS (1992), a native of Gyälrong himself, classify the languages of Gyälrong as "Tibetan dialects." This choice is due to political and cultural

37. Scholars of the international community just as most Chinese linguists describe these languages as being "non Tibetan languages" (Chin: fei zangyu). However, some Tibetan linguists and traditional philologues such as Duoczji (1998), BTSAN.LHA NGAG,DBANG TSHUL.KHRI.MS or Konchok Jiatso are of the opposite opinion. Whatever is the precise affiliation of these languages, they are extremely important for the reconstruction of the proto-Tibetan since they have a lot of archaic features and in any case are closely related to Tibetan.
motivations, but not to a purely linguistic analysis.\textsuperscript{38} For example, to explain some fundamental discrepancies in the lexicon of rGyalrongic and Tibetic languages, \textit{BTsan.lha Ngag.dbang Tshul.khrims} (1992: 78) uses cultural arguments. Thus, for example, the stem གནམ་\textsuperscript{39} \textit{GNAM} or རུ་(སྟད་) \textit{Nam.mkha’} ‘sky’ is found in all the Tibetic languages, but in rGyalrong, the word for ‘sky’ is བོད་ \textit{DE-MO} and obviously not related to \textit{GNAM}. The word མ་ \textsuperscript{39} \textit{SA} ‘soil, earth’, which is also a Tibetic pandialectal root is rendered as བོད་ \textit{DE-PHO} in rGyalrong. The author explains these discrepancies in the following way: the word \textit{DE-MO} contains the syllable \textit{MO} which means ‘female’, whereas the word \textit{DE-PHO} contains the syllable \textit{PHO} which means ‘male’ (in Tibetan). Even if this hypothesis were valid, it does not explain why in this rGyalrong language the words ‘sky’ and ‘earth, soil’ are not cognates with \textit{GNAM} and \textit{SA}.

In the Amdo speaking area of Qinghai, we find several Tibetospheric languages that belong to the Mongolic stock such as Mangghuer, Mongghul, Manegacha as well as Salar, a Turkic language (see Simon 2016) and Wutun, a Sinitic language (see Sandman 2016). Some of these languages have been fundamentally influenced by the surrounding Amdo dialects and by CT not only in their vocabulary but also in their grammar. For example, Salar has developed several evidential categories, which clearly reflect a Tibetic influence (Simon ibid.).

A number of languages of Yunnan and Sichuan are located at the junction between the Tibetosphere and the Sinosphere. They include languages belonging to various groups, such as Bai, Lolo-Burmese (such as Lisu and Nosu), Naic (Naxi, Na, Malimasa, Laze, Namuyi, and Shuhing) and a few other languages, whose classification is not always well established such as Drung.

\textsuperscript{38} The confusion between ethnicity, language and nationality is frequent in China and South Asia. It triggers a lot of fluctuations in the definition of nationalities. These notions are not necessarily connected. For example, Breton speaking people are considered as French citizens but from a linguistic point of view, their language is Celtic and does not belong to the same family as French. For a discussion about the ethnic names, see Chapter 3.

\textsuperscript{39} This word may be related to the Tibetan word སྒྲ་ \textit{DMU}.
Let us now turn to the Tibetospheric languages outside China. There are a lot of parallelisms with the situation we have just described. The TB languages which have been deeply influenced by Classical Tibetan and are in close contact with modern Tibetic languages are for the most part Tibeto-Himalayan languages, particularly Bodish languages. Speakers of these Tibeto-Himalayan languages are essentially Bhutanese, Nepalese or Indian citizens. Thus, a significant percentage of these speakers can also master, depending on their location, one of the major Tibetic languages such as Dzongkha (the national language of Bhutan), Lhoke (one of the official languages of Sikkim), Sherpa (a major Tibetic language of Nepal), Khunu-Spiti dialect of Himachal Pradesh in India, and so on, and in some cases Common Tibetan. And of course, a number of monks and educated people also master Classical Tibetan, which is the liturgical language of Buddhism and Bön. In Bhutan, CT is often referred to as Chöke “the Dharma language.” Additionally, the speakers of Tibetospheric languages living in the Indian Subcontinent also usually master an Indo-Aryan language such as Nepali or Hindi and sometimes English.

Just as rGyalrongic languages are typical examples of Tibetospheric languages on the Tibetan plateau, Tshangla is a good example of such a language in the southern Himalayas. Tshangla, often called ‘Sharchop’ (ཤར་ཕྱོགས་ཁ་), is the major language of Eastern Bhutan also spoken in Arunachal Pradesh (India) and in Metok County, in Tibet. It has been in intense contact with CT and modern Tibetic languages such as Dzongkha or Choča-ngača. Although Tshangla is often classified as Bodish, there is no consensus concerning its affiliation. Some of its basic vocabulary is very different from CT or modern Tibetic languages. For example, /lum/ ‘stone’ versus CT RDO, /kha/ ‘bird’ versus CT BYA, /ti/ ‘water’ versus CT CHU, /gadaŋ/ ‘hand’ versus CT LAG (па), /ʑi/ ‘blood’ versus CT KHRAG, etc. However, when one looks at a large

40. For a list of these languages see below 10.6 and for a more general list of the languages in contact with Tibetic see Chapter 3.

41. In a marginal way, they are also found in Tibet: Tshona Monpa and Basum are spoken on the Tibetan Plateau (TAR) itself.

42. Tshangla is the second language of Bhutan after Dzongkha. Sharchop which just means ‘Oriental’ in Dzongkha.
portion of the vocabulary and even the grammar, it becomes clear that the number of loanwords from CT and modern Tibetic languages is very high (see Andvik 2010). Tshangla speakers also share with Dzongkha speakers fundamental aspects of the Bhutanese culture. They use Dzongkha as a written language and have competences in CT. What has been said of Tshangla also applies to all the East Bodish languages of Bhutan.

If we now turn to Nepal, we find similar situations with Tamangic languages, such as Manangi[43] locally called Nyishang (ཧྱི་ཞང་), but also Tamang (རྟ་མང་), Gurung (རྒུང་), Thakali (ཐ་ཀ་ལི་), which have been in contact for many centuries with Tibetic languages (such as Lo-ke). They share with them many cultural and religious features. From a genetic point of view, the Tamangic languages along with East Bodish are the closest to Tibetic.

Among the Tibetospheric languages that have both close genetic connection and contact with Tibetan, one ought to mention some of the West Himalayish languages such as Kinnauri, Bunar, Tinan, Manchad and Byansi. They are spoken in India, in Himachal Pradesh and Uttarakhand. These languages have been in contact with the Tibetic Spiti-Khunu-Garzha language spoken in Himachal Pradesh and in some cases with the Purang dialects and adjacent dialects of Western Tibet. Kinnauri is in close contact with the Khunu dialect spoken along the Sutlej River in the nearby valleys, and Tinan and Bunar are respectively in contact with the Khoksar Garzha and the Töt Garzha. Manchad is in contact with the Patanam Garzha dialect. The Byangsi language and the Purang dialect of Ngari in Western Tibet are spoken in adjacent areas. Over the centuries, Tibetan had an impact on those West Himalayish languages. A number of Kinnauri people in Rekong Peo area (particularly in Morang and Puh Tehsil) can speak the Khunu dialect spoken in the upper course of the Sutlej. From a genetic point of view, West Himalayish language are not as closely related to Tibetic as the the Tamangic or East Bodish languages.

The Tibetosphere traditionally occupied a huge territory between the Indian and the Chinese spheres of influence. However cultural borders are not fixed and they

depend largely on geopolitical, economic, cultural and religious factors. For example, a West Himalayish language like Manchad has slowly been drawn into the Indosphere after the conversion of the population to Hinduism. In Uddaiipur, the main town located in the Manchad-speaking territory, the sacred Hindu-Buddhist site of ‘Triloknath’ or གི་ཁྲོག་བཞགས་པ་ Garzha phagpa reflects the meeting of the two cultural influences. This shift of cultural influence has resulted in a massive borrowing of Indo-Aryan words in Manchad. The Tibetic enclave of Garzha spoken in the Patanam valley (Myaḍ) a few km from Uddaiapur is also impacted by this cultural shift and some young people of the Patanam community are barely aware that they speak a Tibetic language and that their dialect is closely related to other dialects of Garzha and Spiti.

Finally, let’s mention an Indo-Aryan language, Brokskat ‘language of the Brokpas’, which is spoken in Ladakh, India. Brokskat is a Dardic language just as Kashmiri but as the language name itself shows, it has been heavily influenced by the neighboring Tibetic languages of Ladakh, Sham and Purik (see Tsewang Gailtsen 2018). Brokskat is spoken by less than 5,000 speakers in the Indus and Hanu valleys as well as some villages of Baltistan. Since it is not written, it is quite endangered and could disappear within two generations. Most speakers of Brokskat are trilingual, since they can also speak the Purik (or Sham) dialect, as well as Urdu.

For some languages, spoken at the margin of the Tibetosphere, the impact of the Tibetan culture and language is quite limited. Let us consider a few examples. Within Sikkim, the Lepcha language has been in contact with Lhoke and CT for many centuries. Before the development of a specific Lepcha script in the seventeenth century, the language was even written down in Tibetan script (see Plaisier 2007; van Driem 2001). Lepcha or རོང་ Rong communities who are called མོང་པ་ MON.PA in Lhoke have naturally has been influenced by Tibetan, as we have seen in 10.2, but the impact on its grammar and vocabulary has remained limited.
The same is true for another group of languages, spoken in Nepal but also in Sikkim: the Kiranti languages.\textsuperscript{44} We have already quoted in 10.2 examples from Limbu, one of the most prominent Kiranti languages together with Sunwar, Bahing, Kulung, etc. In Sikkim, Limbu speaking people who call themselves Yaktungba (ཡག་ཐུང་པ་), are referred to as $\text{TSONG}$ by the Lhopos, but according to van Driem (2001: 665), they are called $\text{TSHONG}$ by the Lepchas because they came to Sikkim as merchants. Although some Kiranti languages may be more influenced than Limbu by the neighboring Tibetic language,\textsuperscript{46} this influence remains marginal. Following the conversion to Buddhism by the Lhopos (or Dränjonpas), Limbus developed their own alphabet script at the beginning of the eighteenth century. In the case of Lepcha as well as Limbu, the choice of a script distinct from Tibetan also shows the cultural distance from the neighboring Tibetic populations.

To conclude this section, over many centuries the dominant influence of Classical Tibetan and the Tibetic languages has left a considerable imprint on many closely related languages such as Bodish, Rgyalrongic and Qiangic languages spoken in Tibet and on the Tibetan Plateau in China but also in Nepal, Bhutan and India. The impact of Classical Tibetan and the Tibetic languages is even visible on languages of the Tibetosphere which belong to other stocks such as Mongolic, Turkic and even Indo-Aryan.

However, for some languages, often located at the periphery or that are genetically remote, the imprint may be quite limited. In any case, the Tibetospheric languages are very important for the reconstruction of Proto-Tibetic as well as, more generally, Proto-TB and even Proto-ST.

\textsuperscript{44} Their genetic affiliation in the TB family is probably closer than Lepcha but unlike the latter, they have preserved some archaic features and belong to the “pronominalized languages,” i.e. languages with argument indexation in the verb. Kiranti languages are genetically a lot more distant from the Tibetic languages than Eastern Bodish or Tamangic groups.

\textsuperscript{45} “Tshong” is a loanword from Lhoke or Tibetan, which means ‘merchant’.

\textsuperscript{46} For example, the Khaling people who use to live in the territory nowadays inhabited by the Sherpas. See van Driem (2001: 710).
10.6. Methodology for the identification of a Tibetic language

As we have seen in the preceding section, some linguistic groups are not only genetically close to Tibetic, but also very influenced by this phylum. Thus, when the researcher is confronted with great similarities, s/he may wonder whether the language is indeed Tibetic or not.

When we use the methodology proposed below, the identification is immediate and usually not problematic. Some very rare cases have raised some debates, such as Baima, which a few scholars identify as "non-Tibetic." However, it is clear according to our criteria (see below the five main criteria) that Baima is Tibetic, despite the fact that some of its grammatical features exhibit discrepancies with other Tibetic languages. These peculiarities may be due to language contact or to a substratum influence. Conversely, languages such as Basum (in TAR) or Kurtö (in Bhutan) do not fulfill the criteria and can not be considered as Tibetic.

In order to identify a Tibetic language (or a "Tibetan dialect" in the classical wording) a simple test based on the word for 'seven' has been proposed by Beyer (1992: 7) and Michailovky and Mazaudon (1994: 2). In the Tibetic languages, the word corresponding to 'seven' is regularly derived from CT BDUN, while it is not the case in other languages, even when they are closely related to the Tibetic languages. The above test could be used as a first indication, but numbers can also be borrowed. Indeed, in Japhug, a rGyalrongic language, the cardinal numeral is not a reflex of BDUN but the ordinal has been borrowed from Tibetan:

βdʊnpa "seventh" < CT BDUN.PA (see Jacques 2004b).

Although the above test is valid, it gives the wrong impression that the number 'seven' would be the main difference between Tibetic languages and closely related languages (East Bodish, rGyalrongic, etc.). As we have seen in the previous sections of the present chapter, this is clearly not the case and many other lexical and grammatical features also differ. Concerning the basic vocabulary, we have seen that apart from 'seven', words such as 'red', 'to come', 'milk', 'leg', etc., are usually not cognates with their CT equivalents even in the languages that are most closely related to Tibetic.
For Natives as well as colleagues working in Human sciences (but not linguists), a simple quick test that requires only a few minutes should allow very rapidly asserting whether a language is Tibetic or not.

a) The numbers and particularly number ‘one’, ‘four’ and ‘seven’ should be directly derived from CT *GCIG*, *BZH* and *BDUN*. Any discrepancy would suggest that the language is not Tibetic. For example, if for the number ‘one’, a language has an initial /t/ instead of /č/ or /šč/ (CT < GC) or if for the number ‘seven’, the word in a given language starts with a nasal initial /n/ or /ny/ or something else instead of a plosive dental (/t/ or /č/), it is non-Tibetic.

b) The negation should always begin with an /m/ and be directly derived from CT *MA* or *M(γ)* (the vowel may change significantly /mo/, /mu/, /me/). A negation in /a/, for example, immediately signals a non-Tibetic language.

c) Auxiliary verbs directly derived from CT *YIN* and *YOD* are normally present in the Tibetic languages. Any language lacking these copulative and existential verbs is likely to be non-Tibetic.

However, the identification of a Tibetic language should be ultimately based on a number of phonological, morphosyntactic and lexical criteria (see Tournadre 2014a). We propose the following fundamental criteria:

1) The basic vocabulary is cognate to CT or OT. It means that a very significant percentage of the basic vocabulary is pandialectal (see Chapters 11 & 12). Sometimes, the modern Tibetic forms found in the various languages may be cognate with different CT roots with a similar meaning.

2) The morphological form of the lexical items is directly and regularly derived from CT or OT. The exceptions should be explained. For example, odd derivations may correspond to loanwords from another Tibetic language.

3) The morphophonological form of the lexical items must be compatible with the developments that occurred in Proto-Tibetic (see Chapter 4).

4) The grammatical words that are cognate in the various Tibetic languages

47. In several cases, we find a /ny/ initial instead, which reflects an archaic form (OT) *MYI*. 
normally include at least the negation म न(A) or म ल(M(Y)), the nominalizing suffixes प A and/or मक्खन MKHAN, and various reflexes of the classical cases such as the genitive GI/YI and the ergative/instrumental cases गि S or the dative ल A.

5) The tense-aspect-modality system is marked by a paradigm of verbal auxiliaries that include the reflexes of the copula yin YIN and the existential verb योड्द/ ‘DUG/ YONG and ग्रो ‘GRO (see Chapters 8 and 9). These auxiliaries convey epistemic and evidential meanings.

The first criteria (1) above implies that the basic vocabulary of a Tibetic language should be cognate with CT or OT. See for example a list of pandialectal words (11.1) found in all the Tibetic languages.

However, in any given language, there might be some exceptions:

a) Some of the basic words may have been borrowed from neighboring languages belonging to other families. For example, some Tibetic languages of Northern Nepal do not have the word फङ्गा PHAG(PA) ‘pig’ and use instead the word /sunggar/. This word has been borrowed from Nepali and the original Tibetic root has been lost.

b) In some rare cases, a word may reflect an archaic form not attested in CT but attested in OT. For example, some languages do not have reflexes of the word भित्त MIG ‘eye’ and भिन्न ME TOG ‘flower’ but exhibit reflexes of OT भित्त MYIG or भिन्न DMYIG and भिन्न MENTOG (see Chapter 4).

c) In some exceptional cases, words do not have reflexes of CT or OT but of a more archaic form that has to be reconstructed for Proto-Tibetic. For example, in the dialects of Spiti and Purik, the word for ‘big (rope)’ does not reflect the CT झोर्नै SBOM-PO, but a Proto-Tibetic form झोर्नै SBROM-PO (see Chapter 4).

d) Sometimes the lexical items have distinct etyma in the various languages, but they are all cognates with words attested in CT: ‘To fear’: /jiʔ/ (Tô), /`zh`wa l`an/ (Sh); /zhiks/ (La), /zhiks/ (Ba) < CT झिज्ज jigs, /`th`a/ (Kh), /l`a/ (Ho), /c`ax/ (Am) < CT झिज्ज SKRAG, /she:/ (U) < CT झेड़ ZHED, /`th`e:/ (Ts)
< CT BRED, /ˈdɔːr/ (Dz) < CT DROG (see Chapter 11 and the HCLT, Chapter 12). As another illustration, we may mention the word 'milk'. Two distinct etyma are found for the word 'milk' across the Tibetic languages. The word འོ་མ་ OMA or more rarely the variant འོ་ལྗེན་ OJEN are attested for 'milk' in nearly all the modern Tibetic languages. There are however at least two exceptions: /ʂu/ in Rongdrak (Kham), /sho/ in Choća-ngača (Bhutan). These words are clearly cognates and are probably derived from zho, the genuine word for 'milk' in Proto-Tibetic. The Proto-Tibetic form, which may be reconstructed as བོ བོ (N), is a nominal form of the verb བོ བོ to milk. The meaning of བོ ZHO 'milk' is also attested in CT (see the word 'milk' in the HCTL, Chapter 12). In most modern languages, the word བོ ZHO has acquired the meaning 'yoghurt'.

The second criteria (2) is related to the regularity of sound changes. It implies that the lexical items are not only cognate with their equivalents in CT but also regularly derived from them. Thus, for example, the initial consonantic cluster LT in any Literary Tibetan word such as བོ བོ 'to watch', བོ བོ 'to be hungry', བོ LTE 'navel', should have regular reflexes in all the Tibetic languages and dialects.

The regular reflexes of LT are: /lta/ (in Ladaks, Purik and Balti), /rta/ or /hta/ (Am), /tā/ (in Ü, Tsang and Kham), /Ta/ (in Sherpa), etc. So, if we examine the words corresponding to LTA and LTOGS in the various languages, we find the following forms:

LTA བོ 'to look at' > /lta/ (La, Pur, Ba), /tā/ or /hta/ (Am), /tā/ (Ü, Ts and Kh), /tā/ (Sherpa); LTOGS བོ བོ 'to be hungry' > /toks/ (Balti, Purik, Ladaks), /tox/ or

48. This word originally refers to 'breast' and to the act of sucking. Cf. OBDOG 'nipple', O BYED 'to kiss'.
49. The word /sho/ is reported by Tournadre & Karma Rigzin (2015). See also Ebihara et al. (2016) for a linguistic map of 'milk'.
50. It has the following inflections: /bZHOS/ (present), /BZHOS/ (past), /ZHOS/ (future), /JOS/ (imperative).
51. The cluster is t, but we mention the vowel /a/ in order to indicate tones when they are present.
52. There are a few other possible reflexes depending on the dialects.
The regular reflexes of སྲ་ SRA are: /ʂa/ (La, Am), /ʂ/ (Ts, Sh), /ʈa/ (Kh), etc.

For example, the words སྲོག་ SROG ‘life’ and སྲུང་ SRUNG ‘to keep’ have the following forms in the various Tibetic languages:

SROG སྲོག་ ‘life’ > /ʂox/ (Am), /ʂok/ (La), /stroq/ (Ba), /ʈoʔ/ (Ts, Sh), /ʂoʔ/ (Kh); SRUNG སྲུང་ ‘to guard’ > /ʂoŋ/ (Amdo), /ʂuŋ/ (La), /struŋ/ (Ba), /ʈuŋ/ (Ts, Sh), /sun/ (Kh, Ü) /ʈuŋ/ (Ü).

In some cases, for the same combination, we find two or more reflexes within a single dialect. In the above example of སྲ་ SRA, the Lhasa dialect has two reflexes: /ʈa/ or /sa/. In these cases, we can postulate that only one of reflex is a regular derivation and the other is a more recent form which may have been borrowed from another dialect or influenced by the reading pronunciation of CT.

In some Tibetic languages, reflexes can be quite remote from the original but they remain regular. E.g. ཀླད་པ་ KLAD . PA ‘brain’ > /xlatpa/ (Balti) but /läpa/ (Ü), /leta/ (Sherpa); the phonetic evolution is even more marked in Baima (Chirkova 2008; Zhang Jichuan 1997) and in many Southern Kham dialects (see Suzuki 2009) such as nJol or mBalhag. For example, in Baima, ལག་པ་ LAG . PA ‘hand’ becomes > /yঋ/; ལུག་ LUG ‘sheep’ becomes > /yü/ and ལྱིོ་མ་ MCHI . MA ‘tear’ > /dʑwɛ/.

The next criterion (3) stipulates that the morphophonological form of the lexical items is derived from a Proto-Tibetic form, which is not the case of some neighboring languages such as Kheng, Kirtö, Tamang, etc. (see Chapter 4.4). For example, all the Tibetic languages have undergone a palatalisation of dental and alveolar before y. Thus, the word ‘four’ is reconstructed as Proto- *b-la-y in Proto-TB but it has become /bʑi/ in Proto-Tibetic which is reflected in the Classical orthography བཞི་ BZHI ‘four’ and in all the subsequent Tibetic languages. This is not the case in many Bodic languages very closely related to their Tibetic neighbors such as Tamang, Kheng, Basum, Dakpa, etc. which have preserved an /l/ or /ɾ/ in their form: /pʰi/ (Tamang), /ble/ (Kheng, Bumthap), /pʰɾ/ (Dakpa), /bɾ/ (Bake), etc. Other palatalizations such *ti > /či/ and *si > /shi/ are found in CT: སྱིན་ SHING ‘wood’, *ti > /či/ ‘what’. All the Tibetic languages have reflexes derived from the palatalized form. Amdo and most Kham dialects have undergone a velarization and have the reflex /ʃ'/.
but no Tibetic language has preserved a dental /s/. However, East Bodish or other closely related languages have preserved non-palatalized forms: ‘wood, tree’: /səŋ/ (Kurtoep), /siŋ/ (Bake), /siŋ/ (Tamang), etc.

Another characteristic change, which occurred in Proto-Tibetic, is the shift from lateral to dental after m. It attested in all the Tibetic languages but not in the closely related neighboring languages. Thus Proto-TB *b/m-la ‘arrow’ has become /mda/ in Proto-Tibetic. This form is reflected in the Classical orthography ༭དའ། MDA’ and in all the modern Tibetic languages. However, East Bodish or Tamangic languages did not undergo this change. The ‘arrow’ has the following equivalents: ‘/mja/ (Kurtö, Kheng), /mra/ (Mangde), /bA/ (Dakpa), /mja/ (Tamang), etc.

A consequence of these specific evolutions is that some phonotactic combinations which are found in some closely related languages (such as Kheng, Kurtö, etc.) are simply not attested in any Tibetic language. This is the case for example of the onsets /ml/; /pl/ and /ny/.53 Let us remind that these combinations cannot be easily written down in Tibetan script.

The next criteria (4–5) are related to the grammatical words and to grammatical semantics. We have provided a small list of grammatical words or clitics (negations, grammatical cases, nominalizers and auxiliaries), which seem to be found in all the modern languages.

Most of the modern Tibetic languages have preserved to some extent a case system inherited from OT and preserved in CT (see Tournadre 2010), however, in some languages, it may be quite reduced. All the languages exhibit a form of nominal ergative marking (with only one or two exceptions), but the marking may be optional and related to pragmatic and discursive parameters or on the contrary may be compulsory and have a syntactic character.

The ancient verb tense-aspect morphology inherited from TB has gradually been replaced by a system of auxiliary verbs used with nominalized forms of the verb.

53. Such phonotactic sequences are found for instance in Kheng. See Chamberlain (2004).
Usually a set of available nominalizers include at least reflexes of \textit{PA} and/or \textit{MKHAN} (see Chapters 8 and 9).

Finally, all the Tibetic languages have developed a rich system of evidential and epistemic markers, which appear as verb auxiliaries or suffixes.

The identification of a language as "Tibetic" should thus be based on the five above criteria (and not restricted to one criterion) and on the various domains of phonology, morphology, syntax, semantics and lexicon. In order to sort out the affiliation, it may also be useful to mention the absence in Tibetan of some features that are found in neighboring non-Tibetic languages. For example, classifiers are not found in Classical Tibetan and none of the modern languages have developed any system of classifiers although a few rare classifiers do exist in a marginal way. The "pronominalized languages" include many TB languages belonging to various subgroups such as Qiangic, rGyalrongic and Bodic (Kiranti), but such a system of verb agreement or argument indexation was not attested in Old Tibetan and is of course absent of the modern languages. The morphology of Qiangic and rGyalrongic languages is often marked by prefixation (for noun, adjectives and verbs) instead of suffixation in the Tibetic languages. The verb often presents verbal directional prefixes that are not present in Tibetic languages, with a few exceptions such as Dzayêl Tibetan.

For the majority of non-Tibetic 'Bodic', 'Qiangic' or 'rGyalrongic' languages, the percentage of CT cognates together with Tibetic loanwords may reach up to 50\% of words,\textsuperscript{54} but it is still much less that the percentage of CT cognates in the Tibetic languages which is usually more than 90\%. Furthermore, as we have seen, their phonology, morphology and grammar differ in a significant way from the Classical Tibetan and the modern Tibetic languages. This has two consequences. The Tibetic languages may easily be written in a Tibetan script, preserving most of the CT spelling in a straightforward way. This has been achieved for Ü-Tsang, Amdo, Dzongkha, Lhoke, Ladaks, Sherpa and partially for Balti and Spiti and could be developed for any of the Tibetic languages. However, it is much more difficult to write down non-

\textsuperscript{54} To give a comparison, in English the percentage of words borrowed from Latin and French reaches 60\%. 
Tibetic languages such as Bodish, rGyalrongic or Qiangic languages using the Classical Tibetan orthography.\textsuperscript{55}

Another direct consequence is the fact that for a speaker of a Tibetic language, any other language of the family may be learned very rapidly, at least at a conversational level in only three months, especially if one knows Classical Tibetan, where as Tibeto-Burman languages require a much longer period to reach the same level.

10.7. List of the TB languages in contact with the Tibetic languages

We will list below the main Tibeto-Burman languages that are located within the Tibetosphere or in contact with the Tibetic languages and provide some information that may be useful for the linguistic classification, the substrates and the reconstruction of the proto-languages.\textsuperscript{56}

As we have seen in 10.4, the TB languages of the Tibetosphere include Tibeto-Himalayan, rGyalrongic, Qiangic and in a marginal way Tani, Mishmi, Naic and Lolo-Burmese languages.

Among the Tibeto-Himalayan languages, Bodish and Tamangic languages as well as Tshangla are located within the Tibetosphere. One should mention also Lepcha and West Himalayish, but their tibetospheric characteristics are more marginal.

10.7.1. East Bodish languages (Tibeto-Himalayan branch)

Aside from the Tibetic languages, the East Bodish languages include following languages: Basum (TAR, China), Dakpa (TAR, China; Bhutan, India), Tingpa (TAR, China), Bumthang (Bhutan), Kheng (Bhutan), Kurtö (Bhutan), ’Nyen (Bhutan), Chali (Bhutan), Dzala (Bhutan) and ’Ole (Bhutan).

We give below a small presentation of these languages.

\textsuperscript{55} As we have seen above, this is not the case for the Tibetic languages of Amdo, Bhutan, Kham, Ladakh or Baltistan, etc. Of course rGyalrongic, Qiangic, East Bodish or Tamangic may be written in Tibetan script but many words do not follow the CT orthography.

\textsuperscript{56} We restrict ourselves here to TB languages and do not list here languages of other phyla such as Mongolic, Turkic or Indo-Aryan (see Chapter 3) that are located in the Tibetosphere, because the aim is to list languages that are genetically related to Tibetan and are useful for the reconstructions of the Protolanguages.
- Basum

The Basum language is locally called བཀའ་སྐད།/bäke/. This term is derived from བྲག་གསུམ།/BRAG.GSUM SKAD/ 'the language of the Three Rocks'. Basum is spoken by about 3,000 people in two townships of Kongpo Gyamda County, around 400 km east of Lhasa, in Nyingtri prefecture; ལྗུན་Zhoka and རྟོ་མཚོ། Tshongo townships. The two villages are situated on the banks of a gorgeous lake, the Basum lake བྲག་གསུམ་མཚོ། (BRAG.GSUM MTSHO) 'lake of the Three Rocks', which is surrounded by high mountains and glaciers. The Basum valley is quite isolated and according to tradition, it is made of three doors (མགྲོན་ སྒོ།/SGO.GSUM), three lakes (མཚོ་གསུམ་ མཚོ།/MTSHO GSUM) and three rocks (བྲག་གསུམ།/BRAG GSUM). Drugla Township was once part of Basum territory. There are historical evidence supporting this. Basum people call the language spoken in Drugla འབུག་ལ།/Drugla/ 'Uncle's goat language'. This term suggests a kinship between Drugla and Basum people. There is also a tale about Drugla Khänchen, the abbot of the Drugla monastery who reportedly used Basum language to win a debate in a Lhasa monastery. Basum speaking Tibetans are well aware that their language is very specific and there are a lot of legends about it. They often claim they speak མཁའ་འགྲོ་མ་སྐད།/MCHA.GRO.MA SKAD, i.e. ‘Dākinī’s language’ (see Dākinī in the HCTL, Chapter 12). This myth is related to the fact that Basum is not intelligible for other Tibetan people who live nearby in the Kongpo area. The language might be related with Dakpa since the two languages share some lexical items as well as grammatical morphemes. Basum language is not written and Literary Tibetan is used for this purpose. A description of Ba-ke has been written by Sanchuan Wang (2020).

- Dakpa

Dakpa language ཀྲག་པ་སྐད། (DWAGS.PA SKAD) or ཇུ་གསུམ། (DAG.PA KHA) in Dzongkha, usually called ཀྲག་པ་ཁ་/DWAGS.PA KHA/ or ཀྲག་པ་ན།/DWAGS.PA NUN/ in Tshona, is spoken in Lekpo district of Tshona County (MSTHO.SNA) located in the south of the TAR by about 3,000 speakers, in Tawang རྟ་ཝང་ in Arunachal Pradesh (India) as well as in the neighboring district of Trashigang in Bhutan. The estimate figure of 1,000 speakers is mentioned by van Driem (2001: 871). In the TAR,

57. It is about 15 km long.
most people know Tibetan and some speak Chinese. Literary Tibetan is traditionally used as the written language. In Bhutan, the Dakpa area is mainly a Tshangla speaking area. A small monograph in Chinese was published in 1986 on “Tshona Mönpa.” The Dakpa language can be divided in two dialects: southern and northern, but they do not have a lot of differences and are mutually intelligible. Dakpa, ‘Ole and the Bumthang group (Bumthang, Kurtö, Kheng, Nupbikha, ‘Nyen, Dzala and Chali) are closely related and form the east Bodish group. Though Metok Mönpa, alias Tshangla, and Tshona Mönpa, alias Dakpa, are both called “Mönpa” in Tibet, they correspond to distinct languages.

• Bumthang

Bumthang བུམ་ཐང་ཁ་ (BUMTHANG KHA) is spoken in Bumthang District in Bhutan. It extends in the west to Trongsa (KRONG GSAR). Bumthang has 30,000 speakers. According to van Driem (1998: 18), “Bumthang is closely related to Kheng and Kurtöp.” There are five dialects of Bumthang: ‘Ura, Tang, Chogor, Chunmat and Nupbikha’. Van Driem (1998: 19) adds that Kheng and Kurtö could even be considered as dialects on grounds of mutual intelligibility. Michailovky and Mazaudon (1994: 545) note that:

“Bumthang languages are clearly closely related to Tibetan in addition to being heavily influenced by it […] but they are not Tibetan dialects, that is, unlike Dzongkha, they are not continuation of (roughly) the language reflected in the Tibetan writing system.” (see also van Driem 1995b)

• Kheng

Kheng མིངས་ཁ་ (KHENGs KHA) is primarily spoken in Zh‘amgang District (GZHAL SGangs) as well as in Mongar District (MONG SGAR) in Bhutan. There are 40,000 speakers of Kheng. According to van Driem (ibid.), there is considerable dialect diversity. A form of the Bön religion is practiced in the Kheng area. Chamberlain (2004) gives the following description in his thesis:

“Though there is a distinct dialect in each Kheng-speaking village, there are three major regions in Zh‘amgang district which are marked by dialect groupings. These regions are referred to as Upper Kheng (comprising the Bumthang River valley, with such villages as Shingkhar), Middle Kheng (along the Mangdi river with as villages as Zh‘amgang,
Buli and Tali) and Lower Kheng (where the two rivers meet and exit Bhutan through the Manas jungle, with such villages as Ngangla and Panbang). A fourth dialect is in the western Mongar district. (Chamberlain 2004: 7)

Lower Kheng “appears to be the most divergent of the Khengha dialects in Zh‘àmgang district” (ibid.). Khengpas practice Nyingmapa Buddhism, Bön and a form of Shamanism.

• Kurtö

Kurtö, ཀུར་སྟོདཔ་ཁ(KUR.STOD.KHA) alias Kurtoep or Kurtöp, is spoken in Lhunt’si district, དྲུང་རྗེ (LHUN.RTSE) in Bhutan in the valley of the Kuri river. There are 10,000 speakers of Kurtö according to van Driem (1998). Michailosky and Mazaudon (1994: 545) state that “Kurtoep [Kurtö], Bumthap proper [Bumthang], and by all reports Khengke, to the south of Bumthang are mutually intelligible.” See also van Driem (1998: 19). Hyslop (2011) has written a grammar of Kurtö.

• ’Nyen

’Nyen language, གོན་ཁ་(SNYEN.KHA) is spoken in Trongsa (ཐོང་གསར་KRONG.GSAR) and Wangdi phodr’a (དབང་འདུས་ཕོ་བྲང་). It has 10,000 speakers (van Driem 2001: 871). A possible etymology for the language name is ‘ancient language’ (སྔན་ཁ་SNGAN.KHA). ’Nyen has several dialects. According to van Driem (2001: 913), Phobjikha dialect “differs from other ’Nyenka dialects in its lexic.”

• Dzala

Dzala (དྲུག་ཁ་DZALA.KHA) is spoken in Bhutan, in Trashi Yangtsse (བཀྲ་ཤིས་གཡང་རྩེ་BKRA.SHS.G.YANG.RTSE) and marginally in Lhunt’si District (ལྷུན་རྩེ་LHUN.RTSE). The number of speakers is estimated at 15,000 (van Driem 1998). The language is locally called /dzala mat/. It is called མཁོ་མ་ཁ་MKHO.MA.KHA in Lhunt’si District where Kurtö is spoken (van Driem 2001: 915). According to van Driem (ibid.), Dzala is closely related to མཚོ་སྣ་མོན་པའི་སྐད་MTSHO.SNA.MON.PA.SKAD spoken on the other side of the Sino-Bhutanese border around the village of Lekpo in Tshona County མཚོ་སྣ་རྫོང་(MSTHO.SNA.RG.DZONG) located in the south of the TAR (see Dakpa language above and van Driem). Basum is also may be related to Dakpa. Interestingly, the toponym Dzala is one important village north of the Basum Lake (see above Basum language).
• Chali

Chali ལི་ཁ། (CHALI.KHA) or རི་ཁ། (PYALI.KHA) is spoken in Bhutan in Mongar District མོང་སྒར་ (MONG.SGAR), by about one thousand speakers. Chali vocabulary is heavily influenced by Choča-ngāča, a Tibetic language spoken in the same valley of the Kuri River and Tshangla, which are spoken in the nearby area (see van Driem 2001: 914).

• 'Ole

'Ole language འོ་ལེ་ཁ། (PO.ŁE KHA) is often called by van Driem (1998; 2001) “Black Mountain Mönpa.” The term 'Ole seems more appropriate since the English term as well as the Tibetan word Mönpa are exonyms and Mönpa refer to many populations of South Tibet and the Southern Himalayas. There are three dialects: a western (in Riti and Rukha), a northern (Wangling, Jiangbi, Phumz’ur) and a southern dialect (Cungseng and Berti) (van Driem 2001: 919). The language has been described by van Driem (1995a) and is quite original since it has preserved some of the Tibeto-Burman biactancial verbal agreement, a grammatical feature also found in Kiranti languages. It is spoken by 500 speakers in Tronsa ཀྲོང་སར་ and Wangdi Phodr’a སྣབ་འདུས་ཕོ་བང་ districts.

10.7.2. Tshangla (Tibeto-Himalayan branch)

Tshangla སྣང་ལ་ (TSHANGLALO) or Sharcho is spoken in southeastern Bhutan, in southwestern and northeastern Arunachal Pradesh in India and in southeastern TAR in China. In Bhutan, Tshangla is spoken in Trashigang རྒྱལ་སྨྲི་གནས, Pemagatshâ རྒྱལ་གཡང་རྩེ་, Samdru Jongkhar རྒྱལ་རྩེ་སྔད་མྱ་ལྗོངས་མཁར་, Trashi Yangste རྒྱལ་སྨྲི་གཡང་རྩེ་ and Mongar མོང་སྒར་. By the number of speakers, it is the second language of Bhutan with 138,000 speakers (van Driem 2001: 871). Tshangla, locally called Tsangla-lo in Bhutan “language of the Easterners.”

In the TAR, Tshangla is also spoken by about 7,000 people in the TAR in Metok County (traditionally called Pemakö), in Jingnei Townships as well as Nyingthri County in Tongjuk township (Chin: Dongjiu). See Lu Shaozun (1986: 1). In Tibet, Tshangla is referred to as Metok Mönpa རྒྱལ་རྩོ་བྲ་པའི་སྐད་ ME.TOG MON.PAI.SKAD). Most of Metok Mönpa speakers are bilingual in Tibetan and some speak Chinese.
In India, dialects of Tshangla are spoken in Kameng and Siang in Arunachal Pradesh.

According to van Driem (2001: 989):

“[...] the Tshangla of Tibet originates from eastern Bhutan and not *vice versa* and the language may represent the ancient indigenous tongue of eastern Bhutan.”

A Tshangla Grammar has been written by Andvik (2010) and Zhang (1986) has published a sketch grammar of the Cangluo Menba (Tshangla) spoken in TAR.

10.7.3. Tamangic (Tibeto-Himalayan branch)

According to Mazaudon (1978), the Tamang Branch or “Tamangic” comprise Tamang proper and Gurung (proper) as well as Thakali, Manang, Rengpungmo, Kaike, Chantyal, Kagbeni, Rohani, Ghale and Kutang Ghale. These languages are closely related but not mutually intelligible.

The Tibetic languages of Nepal are in contact with these Tamangic languages as well as some other TH languages such as Dhimal.

- Tamang

Tamang is mainly spoken in Nepal, in the Bagmati, Narayani and Janakpur zones of the Central region, south of the Langthang Lirung peak and the Ganesh peak (respectively 7,254 m and 7,415 m high) in the districts of Kábhirpálañchok, Mákvanpur, Síndhúpálañchok, Nuvakot, Dhmáning, Síndhúlí Rámechaánp, Dólañkha and Rásuwa (van Driem 2001: 963). These districts neighbor the districts of Nyalam and Kyirong in the TAR in China. The number of Tamang speakers is close to 1,000,000. Tamang has two main dialect groups, western and eastern. The eastern dialects are phonologically conservative, having preserved a syllable canon with both final consonants and initial consonant clusters. This canon is progressively simplified from east to west, ending up as CV in Gurung (for details, see e.g. Mazaudon 2017; van Driem 2001). The Tamang are mainly agriculturalists. Most are Buddhists, with an admixture of shamanic practices, particularly for medical purposes.
• Gurung

Gurung (དགུ་རོང་སྐད་) is spoken in Gandaki zone, in the districts Gorkha, Kaski, Lamjung, Parbat, Syangja and Tanahu, south of the Annapurna ridge (van Driem 2001: 958). There are about 230,000 speakers of Gurung and many more ethnic Gurung who have become Nepalophones. According to van Driem (ibid.), Gurung comprises three dialects with a low degree of intelligibility: a western group, an eastern group in Lamjung and Gorkha and a southern group in Syangja.

• Thakali

As mentioned above Thakali form with Tamang, Gurung and Manangi the Tamangic group. These four languages are thus very closely related. Thakali ཐ་ཀ་ལི་ is spoken in the area of the Kali Gandaki river in Mustang district by roughly 7,000 people (van Driem 2001: 968). There are two main dialects: the southern dialect called Tamang Thakali and the northern dialect upstream called Sekä ཁེ་སྐད་ (see van Driem 2001: 971). This northern dialect of Seké is in contact with the Tibetic language of Loka གློ་སྐད་ (GLO SKAD) and constitutes an enclave.

• Manangi

Manangi is spoken in the Gandaki zone, in Manang District. There are three dialects: Upper Manangi or Manangi proper, spoken in upper Manang locally called /njeshan/ SNYE SHANG, as well as /nar/ SNAR and /phug/ PHUG spoken in the eponym valleys.

Gyalsumdo རྒྱལ་གསུམ་མདོ་ (RGYAL GSUM MDØ), a Tibetic language of the southwest section (Hildebrandt & Perry 2011; see also Chapter 9) is also spoken in the lower Manang valley. It has been earlier misidentified by van Driem as a Tamangic dialect (2001). The estimated number of speakers is 2,600 according to 1971 census. Mazaudon (1996) has shown that the number of Tibetan loanwords in Manangi is very high. Manangi people practise Bön as well as Kagyüpa Buddhism.

• Kaike

Kaike འི་མོ་ is spoken in Dolpo དོལ་པོ་ district and thus in direct contact with the Tibetan dialect of Dolpo. The affiliation of Kaike to Tamangic has recently been
debated. The language is spoken in three villages as mentioned by Fisher (1971) and though no figure is given, the number of speakers does not exceed a few hundred.

- **Chantyal**

  Chantyal is spoken in Myagdi and Baglun districts by about 2,000 speakers (see Noonan 1999; van Driem 2001; Noonan & Hildebrandt 2017). Chantyal was identified as a language of the Tamang, Gurung Thakali group by Michailovsky (Mazaudon 1978).

- **Ghale**

  According to van Driem (2001: 984), Ghale “appears to be intermediate between Bodish and Tamangic.” Ghale is spoken in the northern part of Gorkha District. The number of speakers was estimated at 12,000 in 1975 (van Driem ibid.). The Ghale people are ethnically classed with the Gurung.

10.7.4. **West Himalayish (Tibeto-Himalayan branch)**

These languages are spoken in India in Himachal Pradesh: Kinnauri, Kanashi, Tinan, Bunam and Manchad, as well as in Uttarakhand: Rongpo (or Garwhal), Byangsi, Darma, Chaudangi and Rangkas. To this list, one has to add the extinct language of Zhangzhung which probably belongs to the West Himalayish subgroup and is considered by some scholars to be related to Kinnauri, particularly Darma (see Martin 2007; 2010; Beckwith 2012).58

All these languages (except Kinnauri) have between 1,000 and 50,000 speakers. Thus, they are quite endangered. Traditionally, some of these languages, particularly Kinnauri, have been in close contact with their Tibetic neighbors.

We provide here below some information about Kinnauri, the main language of the West Himalayish subgroup.

58. On this, van Driem writes (2001: 957): “Although the empirical basis for grouping Zhangzhung together with Bunam, Manchad, Rangkas and Kinnauri may very well be scanty, the hypothesis may none the less be correct.”
• Kinnauri

Kinnauri is spoken by about 65,000 speakers in the area of upper Sutlej Valley after Rampur and the Sangla Valley. It has several very distinct dialects. Some of the Kinnauri dialects such as Jangshung, spoken in Morang Tahsil and Shumcho and Sunam in Puh Tahsil are surrounded by Khunu Tötkat, a Tibetic dialect also spoken in the Puh Tahsil and closely related to the Spiti dialect (with mutual intelligibility).

10.7.5. Lepcha (Tibeto-Himalayan branch)

Lepcha people call themselves རོང་ Rong, but are called མོན་པ་ MON.PA by the Lhpos of Sikkim. Lepcha language is essentially spoken in Sikkim and parts of West Bengal in Darjeeling (སྦྲོ་རི་གླིང་) and Kalimpong areas (ཀ་བློན་སྦུག་). Van Driem (2001: 819) suggests that “Lepcha is the language of the aboriginal populace of Sikkim.” A few small communities of several hundred speakers are also found in Samtsi District in southwestern Bhutan and in Ilam District in Nepal (see van Driem 1998; 2001). In the 17th century, Lepcha has developed an original alphasyllabic script which is still used. According to Plaisier (2007), there are about 30,000 native speakers. In contemporary Sikkim, Lepcha is just one of the minorities. The majority of the Sikkim population is made by Nepalis who are of various ethnic groups. In a paradoxal way, Nepali and not Hindi has become the main language of this Indian state.

10.7.6. rGyalrongic languages

rGyalrongic languages are spoken in the traditional region of rGyalrong, which corresponds to the abbreviation of rGyalmo Tshawarong རྒྱལ་མོ་ཚ་བ་རོང་ (RGYAL.MO.TSHA.BA.RONG) lit. ‘the hot valley of the queen’. An alternative name ‘Tsha kho’ is also used to refer to the area. The rGyalrong region is now located in the administrative divisions of Ngawa Tibetan and Qiang Autonomous Prefecture as well as Kandze TAP of Sichuan. The rGyalrong people all belong to the Tibetan nationality.

According to Qu (1996), the Gyarongic languages are spoken in a territory of 160,000 km² by about 100,000 speakers. Sun (2000) and van Driem (2001) give a
figure of 200,000. However, given the linguistic diversity, this overall figure is not very significative. A few rGyalrongic languages such as Japhug are only spoken by a few thousand speakers (3,000 to 4,000) and thus can be considered as endangered to a certain degree.

Some scholars consider that the rGyalrongic languages constitute a subgroup of the Qiangic presentation. In this presentation, we keep the two entities separated for historical and cultural reasons. Although, rGyalrong does not refer to a linguistic entity but to an area, its existence as a traditional Tibetan province is well established.

Most of the rGyalrongic languages are not written, however, in the recent years a few scholars have written down some languages in Tibetan script or in an old rGyalrong script. Literary Tibetan is normally used as a written language.

In the first studies conducted on rGyalrongic (Lin Xiangrong, Jin Peng, Qu Aitang) in the 1950s, most authors agreed to distinguish at least three “dialects”: Eastern, Northwestern, Western. However more recent have refined this classification. Sun (2000) has listed six languages: Situ, sTodpa (Sidaba 四大坝), Japhug (Chapu 茶堡), Khroskyabs (previously referred to as Lavrung 拉坞绒), sTau (alt: rTa’u, Daofu 道孚, Ergong 尔龚) and sTodsde (Shangzhai 上寨).

With the progress of linguistic fieldwork and new data as well as anthropological research, several recent works on language classification regarding rGyalrongic propose a more detailed division of languages and renaming with clear definitions: Gates (2014), Suzuki (2012c) and Roche & Suzuki (2017). Some parts have already been reflected in Ethnologue, an online catalogue of the world’s languages.

The language previously called sTodpa (Sidaba) is now divided into Zbu (a.k.a. Showu) and Tshobdun. Other languages newly described include Geshitsa and Nyagrong Minyag.

When used by the Tibetans, the term རྒྱལ་རོང་སྐད་ RGYAL RONG SKAD may refer to any of the rGyalrongic languages, which are not mutually intelligible. However, it

60. Some scholars put sTau and sTodsde under the Horpa languages; however, the use of Horpa here is inadequate. See Sonam Lhundrop (Tunzhu) et al. (2019).
Contact languages usually refer to the languages spoken around Barkham i.e. Situ, Zbu (or Showu), Tshobdun and Japhug.

rGyalrongic languages constitute an interesting branch in regard to their relatively archaic morphology and phonology dating back in some cases to Proto-Sino-Tibetan (see Jacques 2004a). To the Tibetologists, these languages are also precious witnesses of older stages of Tibetan, as they came into contact with central Tibet at least as early as Vairocana’s exile in the rGyalrong region, in the eighth century (Jacques 2004a: 755-797).

From a linguistic and typological point, rGyalrongic languages are remarkable for the following characteristics: 1) The existence of a very complex verbal agreement (two speech participants can be indexed on the verb); 2) A system of directional prefixes encoding upstream and downstream directions; 3) Existence of stem alternations (ablaut).

**Situ (or Eastern rGyalrong)**

Situ 四土 is the Chinese name of this language which is locally called [k̡u]. In Tibetan, the term rGyalrong is a geographic term and does not designate any precise language. Situ is the largest group in term of speakers, and used to be the lingua franca among speakers of other rGyalrong dialects, alongside Amdo. Eastern rGyalrong is divided in four main dialects: \( \text{B} \text{K} \text{R} \text{A} \text{S} \text{H} \text{I} \text{S} \text{G} \text{L} \text{I} \text{N} \text{G} \) (Lixian 理县), \( \text{B} \text{A} \text{R} \text{K} \text{H} \text{A} \text{M} \) (Maerkang 马尔康), referred to in Chinese as the speech of Situ 四土 “the four rGyal-po/Tusi 土司” of LCOG RTSE, SO-MANG, RDZONG, GAG and DAMPA, BTSANLHA 小金, on this dialect see BTSANLHA NGAG DBANG TSHUL KHRIMS 2009; Mansier 1984; Lin Xiangrong 1993; Lin You-Jing 2003 and Chu-chen (Jinchuan 金川). The best studied dialect is LCOG RTSE (Zhuokeji 卓克基; Lin Xiangrong 1993; Lin You-Jing 2003; Nagano 2018), a dictionary of which was compiled (Huang & Sun 2002). Zhang (2020) is a descriptive study of the BRAG BAR dialect. A dictionary of bTsantha rGyalrong has also been edited by a native scholar (BTSANLHA NGAG DBANG TSHUL KHRIMS et al. 2009).

61. In Tibetan, the term rGyalrong is rather a geographic term and does not designate any precise language. We propose to refer to rGyalrong proper as KURUSKAD.
• **Zbu**

Zbu (Chin: 日部 ribu) also called Showu, is spoken in Ribu township and Kangshan (Chin: 康山, Tib: KHANG.GSAR) township in Barkham County as well as Ronan and Kehe townships in Ngawa TAP. It is spoken in a couple of villages in Dzamthang and Ngawa counties. The language is called རྫོང་འབུར་སྐད་ RDZONGBUR SKAD in Tibetan. The number of speakers probably does not exceed 10,000. An alternative name Showu, based on an exonym /ɕoʁu/ used in Tshobdun has been proposed by Sun T.S. (2007). In Ngawa, Zbu is called སྨར་རོང་སྐད SMAR.RONG.SKAD. As shown by Sun (ibid.), Zbu is certainly the most archaic of the rGyarlrongic branch regarding the verb morphology. Gong (2018) is a descriptive study of Zbu.

• **Tshobdun**

ཁོ་བདུན་ TSHO-BDUN language called Caodeng 草登 in Chinese is spoken in Barkhams County in Tshobdun Township (Caodeng xiang). This language has been extensively studied by Sun (1998; 2003c). There are 3,000 people in Tshobdun. Zbu and Tshobdun languages are sometimes grouped together and called བོད་པ་ STOD.PA (si da ba), but these two languages do not allow mutual intelligibility. Sun and Shidanluo (2019) publish a collection of stories in Tshobdun.

• **Japhug**

ཇ་ཕུག་སྐད་ Japhug is spoken in Barkhams County in three townships: Longerjia ནང་བརྒྱད་ GDONG.BRGYAD (龙尔甲), Shaerzong གསར་རྫོང་ GSAR.RDZONG (沙尔宗) and Dazang རྡ་ཚང་ DA.TSHANG (大藏), called [tatsi] in Japhug) along three rivers: Japhug valley, Gajiao valley 良脚, and a small portion of the Jiaomujue valley 脚木足. Jacques (2004) has investigated this language in his PhD and in several subsequent articles, and more recently (Jacques 2021b) has published a grammar of Japhug. He has also translated in French a version of Gesar epics in Japhug (Jacques 2010).
Contact languages

Khroskyabs (Lavrung)

Khroskyabs, formerly known as Lavrungr, is a language which comprises at least three dialects: mBrongdzong (Chin: 木尔宗 Muerzong) in the eponym township, Thukje Chenmo (ཐུགས་རྗེ་ཆེན་མོ་ Thugs,rje Chen,mo) in the village of Guanyinqiao, Jorok (业隆 Yelong) in Jimu and อ. zero (俄热 Ere) township in Chuchen and Barkhams counties. Thukje Chenmo is also spoken in the northwest of Chuchen County, Sichuan. The Wobzi dialect is described by Lai (2017). Some dialects of Khroskyabs remain undescribed.

sTau (a lt: rTau, rTa'u)

sTau language (རྟའུ་སྐད་ RTA'U SKAD) refers rather to a group of closely related languages than to a single language. This group is called in Chinese 道孚 Daofu. One also finds alternative names: 尔龚 Ergong or 霍尔 Hor, which are respectively derogatory and inappropriate (because of a wrong ethnic identification). sTau is spoken by about 45,000 speakers in the following counties of Kandze TAP: rTa'u County and 茶諾 Drango. There are several dialects. The dialect spoken in rTa'u County has been described by Huang Bufang and Sun Hongkai.

Stau speakers usually also know either Tibetan Amdo or Chinese. Jacques et al. (2014) and Gates (2021) have provided a description of Stau.

Geshitsa

Geshitsa (革什扎, Gebushiza) in Tibetan is spoken in Rongdrak. It has been the subject of a monograph written by Rdo, rje (Duorjir 1998), a native speaker. It is called Geshiza (革什扎) or alternatively Gebushiza (革布什扎).

63. The term Lavrungr comes from BLABBRANG, a famous monastery in Gannan Prefecture of Gansu, and this term was misunderstood and unfortunately created as a linguistic name. The name KHRO.SKYABS corresponds to the proper name denoting one of the eighteen Gyalrong chieftains, known as 绰斯甲 Chaosijia in Chinese.

- Nyagrong Minyag

Nyagrong Minyag is a Tibeto-Burman language belonging to the rGyalrongic group of the Qiangic languages, spoken by about 1,000 Tibetans in Nyagrong [Xinlong] County, Ganzi [dKar-mdzes] Tibetan Autonomous Prefecture, Sichuan, China. It has two dialects: rGyarwaghis (spoken in a part of Jialaxi [RGYA.RWA.GHSIS] Village) and Bangsmad (spoken in a part of Bomei [BANG.SMAI] Village). Suzuki (2012c) reports that mutual intelligibility is found to some extent between sTau and Nyagrong Minyag, but this is not enough to make each speaker understood within each language.

- sTodsde

sTOD.SDE language which is called Shangzhai 上寨 in Chinese is located near the confluence of the Duke river and its tributary the Zongke river in Shili 石里, Zongke 宗科 (RDZONG.KHOG) and Puxi 蒲西 (PUXI) townships in Dzamthang County བོད་རྩོང་ (Sun 2000: 164). According to Sun (ibid.: 166), “Shangzhai and Horpa [Tau] stand in a dialectal relationship to each other.” The number of speakers is unknown.

10.7.7. Qiangic languages

Qiangic languages are spoken in Sichuan, Yunnan, Gansu in South-western China, mostly in Tibetan Autonomous prefectures.

They include Rmaic (also called Qiang), Darmdo Minyag, Shimian Minyag, 64 Prinmi (Pumi), nDrapa (Zhaba), Choyu (Queyu), Lhagang Choyu, 65 nGochang (Guiqiong), Ersu, Lüzu, Shuhing (Shixing), Doxu, Lamo, Larong sMar and Drayap

64. Darmdo Minyag and Shimian Minyag are often considered as two dialects of the Minyag languages. Dawa Drolma & Suzuki (2016) propose to divide it into two languages. See also Roche & Suzuki (2017, 2018).


According to most authors (Sun H. 1983; Jacques 2004a; Sun T.S. 2000), Qiangic also comprises rGyalrongic languages: Situ, sTau, Geshitsa, Nyagrong Minyag, Stodsde (Shangzhai), Japhug, Zbu, Khroskyabs (Lavrung) and Tshobdun (see above).

Other than those, the existence of at least four languages spoken in Chamdo and Nyingthi Municipalities (TAR) has recently been reported: Lamo, Larong sMar, Drag-yab sMar, and gSerkhu. They are provisionally considered as members of Qiangic languages (Suzuki & Tashi Nyima 2016; Suzuki et al., 2018; Tashi Nyima & Suzuki 2019).

Most of these languages have recently been listed by many linguists such as Sun H. (1983), Dai et al. (1991), Huang and Dai (1992), Qu (1996), Lin (1993), Duoerji (1998), Sun T.S. (2000), Ikeda (2002) and Jacques (2004a). Qiangic languages are concentrated in a few areas, at the border of the Tibetic linguistic area, mainly in the Ngawa Tibetan and Qiang Autonomous Prefecture as well as Kandze TAP of Sichuan and the Liangshan Yi Autonomous Prefecture of Sichuan; hence, this region is often called “The Ethnic Corridor of West Sichuan” (Sun H. 1983).

None of the modern Qiangic languages has a written tradition (see Jacques 2014). One should however mention that the ancient Tangut script, a logographic script invented in the eleventh century was used to write down the Xixia or Tangut language. According to Jacques (2014a), there are enough elements to consider that this language belonged to the Qiangic branch of the ST macrofamily, and more specifically to the “macro-rGyalrongic branch.” (2014a). The script which has nearly 6,000 characters and is inspired by Chinese, is considered by Sofronov (1968) as the most complex script in the world. The reason for this complexity is that the script is not pictographic and has only a limited number of phonetic elements.

---

66. Some scholars still doubt that the group “Qiangic” forms a genetic group as in Nishida 1998 and it is sometimes called “Ethnic Corridor of West Sichuan languages.”

67. Maya glyphs may also be good candidates for this world record.
- **Rmaic (Qiang)**

Qiang ཀྲིང་སྐད་ CH’ANG SKAD (羌语 Qiangyu), locally called རྨ་སྐད་ RMA SKAD is spoken by about 100,000 speakers in Lunggu སྦུང་ཆུ (Chin: Wenchuan), Zungchu བཀྲ་ཤིས་གླིང་, Trashiling བཀྲ་ཤིས་གླིང་, Mao མའོ་, Rongdrak རོང་བྲག་, Throchu ཁྲོ་ཆུ་ (see the Appendix 3 for the Chinese equivalents) and in some parts of Beichuan Qiang Autonomous County (北川) in Sichuan province (see e.g. Sun Hongkai et al. 1991). One can distinguish a northern and a southern dialect. The northern dialect is mainly used in the southern parts of Maowun, Trashiling and Zungchu. The northern dialect is used in the northern part of Mao County and in most of Throchu. The Qiang living in the Qiang communities can speak both Qiang and Chinese while the people scattered in various communities use only Chinese. A good number of Tibetans of Throchu County are native speakers of some Rmaic languages. The Qiang language is not written and people use Chinese or Tibetan. LaPolla and Huang (2003) has written a grammar of Qiang.

- **Prinmi (Pumi)**

Prinmi (普米语), alternatively called Pumi, is locally called [phzôma] རྨ་ཆི, which means "white people" is spoken by about 47,000 speakers in Pumi and Bai Autonomous County of Nujiang Lanping Prefecture, Naxi Autonomous Lijiang County and Yongsheng County in Yunnan Province as well as in Muli རུ་ལི་, Gyäzil ཐྭོ་ནོ་, and Yanyuan 盐源 Counties in Sichuan Province (see e.g. Sun Hongkai et al. 1991: 192; Hu & Hu 2014; Daudey 2014, see also Appendix 3). The Prinmi language is generally not written and people use Chinese as their literary language. Apart their own language, the Prinmi people also use Chinese and other neighboring languages. Some Naxi and Bai can also speak Prinmi. The Prinmi language has two main dialects: Northern Prinmi and Southern Prinmi. Some scholars say that they do not allow good intelligibility, however, many native speakers feel that there is good

---

68. According to 1990 census, there are only 29,675 Pumi ethnic people. Other minorities also use Pumi as a vernacular language.

69. However, several ritual manuscripts have recently been found, in which the Prinmi language is written in the Tibetan script. See Hu & Hu eds. (2014).
The Prinmi language may be referred to as གུ་མི་སྐད་ PHU SKAD in Literary Tibetan but there is no known traditional Tibetan name for this language. Prinmi speakers are considered as Tibetans BODBIRIGS in Sichuan but otherwise are considered as a separate Pumi nationality གུ་མི་སྐད་ PHU SKAD also spelled གུ་མིའིསྐད་ PHU SMIRIGS. Concerning their religion (see Wellens 2010): Hangui, possibly corresponds to SNGAGSMGON in Literary Tibetan. Their rituals are very similar to the Bön religion, however, local features are also found. Ritual masters generally use sutras common to Bön written in Literary Tibetan, however, several sutras are written in Prinmi by using the Tibetan script.

* Darmdo Minyag and Shimian Minyag

The Minyag language is referred to as རྒྱལ་ཐང་སྐད་ MINYAG SKAD in Tibetan and is called 木雅 Muya in Chinese (see e.g. Sun Hongkai et al. 1991). Minyag is spoken by around 10,000 speakers in Dartsendo ཀྲུ་ཆེན་, Nyagrong ཐག་རོང་ and Gyäzil Counties རྒྱལ་ཐིལ་ (see Appendix 3). It is also spoken in Shimian County located in Ya’an Prefecture, the next county bordering Dartsendo in the south. All the three counties are situated nearby the Minyak Gongkar range སྐུ་མོ་གནམ་དཀར་ which is 7,500m high. Now the languages of these two areas are not in contact and divided into two different languages on the basis of their mutual unintelligibility: Darmdo Minyag and Shimian Minyag (Dawa Drolma & Suzuki 2016).

In Dartsendo, the majority of the toponyms are Minyag. The language is mainly used in the villages and at home. The people living on the eastern slopes of the Gongkar range also speak Chinese while those who dwell on the southwestern slopes also speak Tibetan.

From a historical perspective it is important to note that Minyak has played an important part in the Tibetan history. The 11th Dalai Lama མཁྲིས་གྲུབ་རྒྱ་མཚོ་ (MKHAS,GRUB RGYA MTSHO) was of Minyak origin. The 93rd and the 95th Gandan Thriras གནད་མཆོག་ ཐླི་པ (GKRA LHA’IS, DUS) respectively Yeshes Wangden and Trashi Tongdjii (BKRA LAS HIS STONG DUS) were also of Minyak origin. The five learned of Kagyu at the time of

70. Note that we use the spelling Minyag to refer to the Qiangic language whereas we use Minyak for the Tibetan Kham variety spoken in the area as well as for the toponyms.
Tsongkhapa were also from Minyak region. There were giving Dharma teachings in Miniag language.

- **Choyu (Queyu)**

  Queyu has about 7,000 speakers at the border of the three following counties: Nyagchu, Lithang and Nyagrong in Sichuan (see Appendix 3). The language is used within the family or inside the valley, but outside of these contexts, everybody speaks Kham Tibetan, Chinese or another neighboring language (see Dai et al. 1991; Huang & Dai 1992).

  Queyu 靖域语 is the Chinese name to refer to this language. It may be referred to as Cho yul skad in Literary Tibetan but there is no known traditional Tibetan name for this language. The language called Gawa in Lithang County in the villages of Ga and Gyarawa possibly refers to Queyu.

- **Lhagang Choyu**

  Lhagang Choyu is a newly recognized moribund language spoken only in one hamlet named Tage of Tagong Town, Kangding (Dar Mdo) Municipality, Ganzi (Dkar Mdzes) Tibetan Autonomous Prefecture, Sichuan Province, China. There are less than a hundred competent speakers, most of whom, unfortunately, no longer use the language in daily life, speaking Kham Tibetan instead. See Suzuki & Sonam Wangmo (2016b; 2018a; 2019b).

- **nDrapa**

  The nDrapa language, called Zhaba in Chinese (Zhaba Zhabayu), is used by 7,700 Tibetan speakers (1986 census) in Tau County (Zhaba district) and in Nyagchukha Counties (Zhamai district) in Sichuan. These two varieties allow good intelligibility. Their language does not allow any intercomprehension with Tibetan or Minyag. Outside the family and the village, people use the Chinese language to communicate (see Huang & Dai 1992: 643; Sun 1991). The language may be referred to as DRA.PASKAD in Literary Tibetan but there is no known traditional Tibetan orthography for this language.
• **nGochang (Guiqiong)**

Guiqiong (貴琼語) locally pronounced 'GO.CHANG [gu’ tɕʰ] or [gwi’ tɕʰ] is used by 7,000 speakers (1986 census) in Dartsendo County (Yutong district) in Sichuan at the border with Chagzam County (Dai et al. 1991; Sun 1991). There are possibly two autonyms with a slight difference of pronunciation, i.e., nGochang and Guichang, both of which lack a complete correspondence in Literary Tibetan form but its etymology is related to a toponym 'GO.THANG, corresponding to an old Chinese toponym 鱼通 Yutong. The Chinese name probably reflects the latter. Outside the family and the valley, people use Tibetan or Chinese to communicate. Guiqiong language may be referred to as 'GO.THANG.SKAD in Tibetan.

• **Ersu**

Ersu (尔苏语 Ersuyu) has about 20,000 speakers in Ganluo 甘洛, Yuexi 越西, Mianning 冕宁, Muli 普雄, Shimian 石棉, Hanyuan 汉源 as well as Gyäzil བརྒྱལ་ཟིལ་ Counties (Sun 1991: 231, see also Huang & Dai 1992). Ersu language may be referred to as 'ER.SUSKAD in Literary Tibetan but there is no known traditional Tibetan name for this language.

• **Lüzu**

Lüzu (吕苏语 Lüsuyu) locally pronounced [’u ʊ] is spoken in Gyäzil བརྒྱལ་ཟིལ་, Mianning 冕宁, Puxiong 普雄, Ganluo 甘洛 and Muli 普雄 Counties in Sichuan (Huang & Dai 1992: 647-648). Lüzu is sometimes spelled Lyusu in the English literature (LaPolla 1995) or even Lisu which is a source of confusion with the Lolo-Burmese language also called Lisu; however, there are no varieties that use a voiceless [s]-sound as Lüzu, but always a voiced [z] as Lüzu. It is considered as western dialect of Ersu by Sun H. (1983), but according to Huang & Dai (1992 ibid.) the linguistic classification of the Lüzu needs further research (see also Chirkova 2017b). The number of Lüzu speakers is not known, but it probably does not exceed a couple of thousand.

71. These counties are located in Lianshan Yi Autonomous Prefecture, Kandze TAP, and Ya’an Prefecture of Sichuan.
The meaning of Lüzu is “White people.” In the past, the Han Chinese used to call them the Small Xifan “Small Western Barbarians” or Xijiao (“Western religion”). There are now officially considered as Tibetan.

Lüzu language may be referred to as LUSZU SKAD in Literary Tibetan but there is no known traditional Tibetan name for this language.

- **Doxu**

Doxu, often regarded as a dialect of Ersu, is a moribund language. This language is well-known due to the study of Nishida (1973), who predicted the possibility of existence of a language called Doxu in the present society based on a Chinese-Tibetan vocabulary edited in the eighteenth century. See also Chirkova (2014).

- **Lamo**

Lamo is a language spoken in Dongba and Zhonglinka Townships, Dzogang County, Chamdo Municipality, TAR. Suzuki & Tashi Nyima (2016) first reported this language, named mBo skad based on an exonym, and pointed out that it might be a member of Qiangic. Tashi Nyima & Suzuki (2019) is a brief introduction to Lamo as well as for the following three languages LarongsMar, Drag-yab sMar, and gSerku. Lamo has approximately 4,000 speakers. It is divided into two dialects: Lamo and Lamei (see Suzuki et al. 2021 for Lamei). Lamo is allegedly closely related to Minyag; however it awaits confirmation.

- **Larong sMar**

Larong sMar is a language spoken on the riverside of Zlachu in Dzogang and Markham Counties, Chamdo Municipality, TAR. There are approximately 15,000 speakers. See Tashi Nyima & Suzuki (2019) for details.

- **Drag-yab sMar**

Drag-yab sMar is a language spoken in Drayap County, in Rongdrup, Khuda, Shamdu, Tsalzang, and Pälri townships, Chamdo Municipality, TAR. There are approximately 20,000 speakers. See Tashi Nyima & Suzuki (2019) for details.
• **gSerkhu**

gSerkhu is a language spoken in the Adrag Valley in Shangchayu Town, Dzayül County, Nyingthri Municipality, TAR. The ancestor of the gSerkhu speakers are said to have been from Dongba, the Lamo-speaking area, and Lamo and gSerkhu are mutually close to each other. There are approximately 400 speakers. See Tashi Nyima & Suzuki (2019).

10.7.8. **Nungish**

• **Trung (T'rong)**

Trung (独龙语 Dulongyu, སྲུང་སྐད་ DRUNG SKAD) alternatively called Dulong, is a Qiangic language spoken by the Trung minority (独龙族 Dulongzu) in 贡山 Gongsheng Trung and Nu Autonomous County in Yunnan and in Myanmar (Huang & Dai 1992: 649), where it is called Rawang. Trung is reportedly also spoken in Balung County འབའ་ལུང་ in Yunnan and in Dzayül County རྫ་ཡུལ་ in TAR. The number of speakers is approximately 6,000. Young people speak mainly Lisu or Tibetan.

• **Nung**

Nung has been called Anu and regarded as a dialect of Trung by Sun (1982). It is spoken in Gongshan County, along Salween (Nu Jiang) River. However, Qin & Suzuki (2016) claim that the autonym is not ‘Anu’ but ‘Nung’ and the Chinese character 怒 is in that case pronounced Nung (and not Nu). In addition, they distinguish the Nung variety from Trung based on the identical observation of speakers of these languages.

• **Zërwang and Daru**

Zërwang and Daru are two other Nungish languages spoken in Burma (Bradley 1994). They are closely related to Trung and Anung. All these four languages are spoken in the upper of Irrawadyi on the both sides of the Sino-Burmese border. The

---

72. There is not a traditional term in Tibetan to designate the Trung language.

73. See e.g. the *Atlas of distribution of National Minorities in China* (He Shiyuan 2002) and Bradley 2007.
number of speakers is 15,000 (Zorwang) and 35,000 (Daru). These two languages are in contact with the Kham Tibetan spoken in Burma (ibid).

10.7.9. Tani and Mishmi languages

Tani languages are also known as Abor-Miri-Dafla (Matisoff 2003). There is still some uncertainty about the classification of Tani languages (Sun 2003; Post and Sun 2017). Tani and Mishmi languages are spoken in Arunachal Pradesh but marginally in Assam, but they are not directly in contact with Tibetic languages.

Tani languages that are located in the Tibetosphere include Bokar and Sulong, whereas Mishmi languages Idu, Kaman, Taraon and Zaiwa are outside the Tibetosphere. These languages are essentially spoken in the southeastern area of the TAR and on the other side of the Sino-Indian border in Arunachal Pradesh (India).

• Bokar

The Bokar (博嘎尔珞巴 Bogaer Luoba) is a Tani language of the Gallon group. Bokar is spoken in Mänling County (Chin: 米林 Milin) and on the other side of the de facto border with India in the state of Arunachal Pradesh. Bokar speakers belong to the KLO.PA Lopa minority called 珞巴 Luoba in Chinese and are recognized by the Chinese government as a separate ethnic group. Lopas are found in Mänling, Metok, Dzayül and Lhuntse counties of TAR (Huang & Dai 1992), and on the other side of the defacto border with India in the state of Arunachal Pradesh. Altogether, there might be around 200,000 Lopas but it is very difficult to give a precise estimation of their number. On the northern side of the border, in the southern part of Mänling County, the 1990 census reported only 2,312 speakers. From the linguistic point of view, the Lopa minority corresponds in reality to various linguistic communities speaking distinct languages such as Bokar, Idu, Sulong and Damu. In Metok County, some Lopa people use only Tibetan Kham dialect whereas others can still speak Lopa, apart from Kham Tibetan and Monpa. According to Huang & Dai (1992: 653), Bokar and Bangni are rather close and allow communication. This author considers the above three varieties as dialects.
• Puroik (Sulong)

Puroik locally called [su\(^{33}\)loŋ\(^{33}\)] (སུ་ལོང་) is spoken in Dzayül County དཔོ་གཅིག in TAR by 2,000 speakers and in Tawang དཔལ་ནག་ area (Arunachal Pradesh, India). Sulong is the name by which they are known by their neighbors but their autonym is actually [poh\(^{33}\)yut\(^{33}\)]. Puroik speakers are also considered as part of the བོད་ Lopa minority.

• Idu

Idu (义都珞巴 Yidu luoba), alternatively called Chulikata or Midu is spoken in Dzayül County དཔོ་གཅིག in TAR by 7,000 speakers. Idu is classified as a Mishmi language. According to Huang & Dai (1992), Idu and Taraon are relatively close and have 40% of common vocabulary. As mentioned earlier Idu speakers are considered as part of the བོད་ Lopa minority.

• Taraon

Taraon and Kaman (see below) are spoken by the Deng minority. Deng languages are called Tingpa རིང་པའི་སྐད་ (RTING PA’I SKAD) in Tibetan. Taraon which is locally pronounced /ta\(^{13}\) lan\(^{15}\)/ is called 达让 僜 Darang Deng in Chinese. Alternative names for Taraon are Digaru or Methung. Taraon is spoken in Dzayül County དཔོ་གཅིག (TAR) as well as on the other side of the border with India (Huang & Dai 1992: 652). It has around ten thousand speakers, but mainly on the Indian side. Within TAR, Taraon is spoken by about 700 speakers.

• Kaman

Kaman, the other Deng language is locally called /ku\(^{11}\) man\(^{15}\)/. In Chinese it is referred to as 格曼 僜 Geman Deng (Huang & Dai 1992: 653). This language is spoken in Dzayül County དཔོ་གཅིག in TAR by about 200 speakers, according to the 1976 census.

10.7.10. Naic languages

The Naic branch includes Naxi as well a few related languages: Na (or Moso), Malimasa, Laze, Shuhing and Namuyi. There are some debates about the classification of these languages.
Naxi has been previously classified as Lolo-Burmese (see Thurgood & LaPolla 2003) have questioned this classification. Shihing and Namuyi were considered as Qiangic languages. Recently Michaud and Jacques (2011) have shown the close affiliation of these languages and proposed to create a separate Naic branch.

**Naxi**

Naxi (纳西) is the Chinese name of a language (or a group of closely related languages) which is called locally /nɑ˩hɪ˧/ (in Lijiangba dialects) and གྱིང་སྐད་ JANG SKAD in Tibetan.

Naxi is spoken by about 290,000 speakers mainly within Yunnan Province in Gucheng district, Yulong Naxi Autonomous County, Yongsheng County (all located in the area of Lijiang Municipality) and in Dechen TAP in the following Counties: Shangri-La, traditionally called Gyälthang རྒྱལ་ཐང་ (Chin: 香格里拉 Xianggelila) and Balung སྐུ་མོ་ (Chin: 维西 Weixi). Smaller populations of Naxi speakers are found in Dechen སརྒྱུར་ཙུམ་ (Dechen TAP), Gongshan, and Ninglang Counties and even as far as in Muli Tibetan Autonomous County (in Eya Township) or Markham County སྨར་ཁམས་ (Chin: 芒康 Mangkang; in Tshwakhalo Township) in the TAR.

Naxi people have their own religion called /to-mba/ and referred to as Dongba (东巴) in Chinese (see Michaud 2011). This term which is derived from the CT word སྟོན་པ་ STON.PA ‘guide, master’ is also used to designate Naxi Bön priests or Shamans. According to Michaud (2011):

“The Dongba religion and Bön share many rituals and deities as well as a common founder […] Dongpa Shilo is none other than sTonpa Shenrab Mibo, the founder of the Bön religion.”

However, the Naxi form of Bön is quite different from the modern Yungdrung Bön (ཡུང་དྲུང་བོན་) found in Tibet. The word /to-mba/ also designates a unique pictographic script used by the priests during their recitation. A Naxi syllabary called

---

74. The Naxi are sometimes referred to as “Western Naxi.” See He and Jiang (1985), Michaud (2011). The “Eastern Naxi” refer to other closely related languages such as Na.
Geba has also been developed. There are many syncretic elements in the Naxi religious practices. Many Naxi are also followers of the Kagyu school of Tibetan Buddhism.75

If we look at the Naxi religion and scripts, we understand that the Naxi people live at the crossroad between the Tibetosphere and the Sinosphere: the religion has been influenced by the Tibetan Buddhism and Bön but the pictographic scripts rather point toward a Chinese influence.

- Na or Moso

The term Moso (摩梭 mosuo) corresponding to the indigenous pronunciation /moso/ is used to designate a small ethnic group also called Na who is settled in the Yunnan and Sichuan Provinces, close to the border with Tibet. From the ethnic point of view, they are also classified as Naxi. Consisting of a population of approximately 40,000, many of them live in the Yongning region, around Lugu Lake, in Labai and Muli མུ་ལི, in Yanyuan. The Moso are famous for the absence of “marriage” and their matriarchal society. They have developed their own religion called Daba. Michaud (2017) describes a full picture of the tone of Yongning Na.

- Malimasa

Malimasa (玛丽玛萨) is a Naic language spoken by around 1,000 inhabitants in Kenuo, Haini, and Chuanda administrative villages of Tacheng Town and Balung རྲུལ་ (Chin: 维西 Weixi). It has two dialectal varieties, one of which is spoken only in Ruke Hamlet, while the other is more widespread. Speakers do not belong to a single officially recognized nationality: some are Naxi, others are Lisu, and some are Tibetan. Local folklore claims that their ancestors were Moso (Na-speaking people) from Muli (Sichuan) and that the autonym Malimasa originates from Muli-Moso. However, this oral tradition is not shared by all Malimasa. The Lisu Malimasa-speakers especially, do not have such an oral tradition. Short linguistic descriptions of Malimasa are available (Li 2013; Suzuki 2015c).

75. Bradley (2012) discusses a Tibetan-reading pronunciation in the Dongba script sutras. The contents include Bon and even Buddhism.
• Shuhing (Shihing Shixing)

Shuhing (史兴语) /ʂu55 hĩ55/ is spoken by about 1,800 speakers in སྦུ། རི། Muli County (Shuiluo Township), Liangshan Yi Autonomous Prefecture, Sichuan (Huang & RIG'DZIN DBANG'MO 1991: 240; Huang & Dai 1992: 646). Outside the family and the valley, people use Chinese, Tibetan or Prinmi to communicate. Shuhing language is referred to as /ɕu33 mɯ55/ by the Kham-speaking Tibetans of Muli. Thus the Prinmi speakers can be rendered as ལུ་མུ་སྐད་SHUMUSKAD in Literary Tibetan but there is no known traditional Tibetan name for this language. Shihing speakers are officially considered as Tibetans by the Chinese administration.

• Namuyi

Namuyi (纳木义语 or 纳木依语 Namuyiyu) has about 5,000 speakers in བརྒྱད་ཟིལ་Gyäzil, རི་ྗིི། Muli, རི། Mianning, རི། Xichang, and རི། Yanyuan counties in Sichuan (Huang & Dai 1992; Dai Qingxia et al. 1991: 236). Namuyi has many alternative names: Namyi, Namuji, Namʑi, Namuzi. Traditionally local Han Chinese used to call Namuyi “Western Barbarian” (Xifan). Namuyi are officially considered as Tibetan by the Chinese administration. According to Libu Lakhi (2017), Namuyi denotes Namuyi people, and their language should be called Namuyi Khatho. The Namuyi language may be referred to as གནམ་མུ་ཞི་སྐད་GNAMMUZHISKAD in Literary Tibetan but there is no known traditional Tibetan name for this language. A textbook and grammar of Namuyi has been published (Libu Lakhi 2009; 2017).

10.7.11. Lolo-Burmese languages

Several Loloish languages which belong to the Lolo-Burmese branch of TB are also spoken in contact with Kham languages at the Southeastern periphery of the Tibetosphere. They include Lisu and Nosu (or Nuosu). The former is mainly spoken by Lisu people, and the latter is principally spoken by the Yi Nationality གུ་ནིིུ། (Chin: 彝族 yizu), officially recognized by the Chinese government. During and after the Tang Dynasty, the Yi had already developed a logographic script referred to as Classical Yi. The first attested document dates back from the fifteenth century. It had more than 8,000 characters but was not entirely standardized. This script is no
longer in use. At present one standardized Yi script system (规范彝文 Guifan Yiwen) has been established and is mainly used in Sichuan.

- **Lisu**

Among the Lolo-Burmese languages spoken in Yunnan, Lisu (傈僳语 Lisuyu) is in contact with southern Tibetic languages. It is spoken by about 500,000 speakers in Nujiang Lisu Autonomous Prefecture (怒江傈僳族自治州); Dehong Dai and Jingpo Autonomous Prefecture (德宏傣族景颇族自治州); Dechen TAP (德钦藏语 TAP) Dali Bai Autonomous Prefecture (大理白族自治州 ) and Lijiang Municipality. Lisu corresponds to one nationality officially recognized by the Chinese government. In Tibetan, there is apparently no traditional term to call this language and people but it may be referred to as བི་སུའུ་རིགས LI SU RIGS and བི་སུའུ་སྐད་ LI SU SKAD. Lisu has had three script systems used in the Tibetosphere: Fraser script, New romanized script, and 汪忍波 Wang Renbo’s bamboo stick script (see Mu & Sun 2012).

- **Nosu**

Nosu, called 诺苏 Nuosu in Chinese (pronounced /nosu/ in Southwestern Mandarin), is also known as Northern Yi is considered as the prestigious language of the Yi people and has been the basis of the standard Yi language (Chin: 蒙语). Nosu is spoken in Liangshan Yi Autonomous Prefecture and its adjacent area of Kandze Tibetan Autonomous Prefecture in Sichuan as well as Yunnan, in Dechen Autonomous Prefecture and Lijiang. It contacts with Kham Tibetans.

The language is taught in schools, both in its oral and written forms. At the end of the twentieth century, Nosu have developed a complex syllabary (with 756 signs) derived from the Classical logographic script. This modern script is officially recognized and used to write written Yi. It is based on the Liangshan dialect of Yunnan. In Tibetan, there is apparently no traditional term for Nosu in Tibetan language, but it is generally referred to as དབྱིས་སྐད་ DBYIS SKAD.