

himalayan linguistics

A free refereed web journal and archive devoted to the study of the
languages of the Himalayas

Himalayan Linguistics

Outlines of Chocha-ngachakha

An undocumented language of Bhutan related to Dzongkha

Nicolas Tournadre

Aix-Marseille University and Lacito (CNRS)

Karma Rigzin

Royal University of Bhutan

ABSTRACT

This paper is the first attempt to provide the outlines of the Chocha-ngachakha, a Tibetic language spoken in Eastern Bhutan. This language (particularly the Tsamang dialect described here) has preserved many archaic features that are not found in the southern Himalayas. The linguistic conservatism of Tsamang Chocha-ngachakha is not confined to phonology but extends to grammar and vocabulary. The data from Chocha-ngachakha sheds new light on the evolution of the Tibetic family.

KEYWORDS

Tibetic languages, languages of Bhutan, classification of the Tibetic family, descriptive linguistics, dialectology

This is a contribution from *Himalayan Linguistics*, Vol. 14(2): 49–87.

ISSN 1544-7502

© 2015. All rights reserved.

This Portable Document Format (PDF) file may not be altered in any way.

Tables of contents, abstracts, and submission guidelines are available at
<http://escholarship.org/uc/himalayanlinguistics>

Outlines of Chocha-ngachakha *An undocumented language of Bhutan* *related to Dzongkha*

Nicolas Tournadre
Aix-Marseille University and Lacito (CNRS)

Karma Rigzin
Royal University of Bhutan

1 Introduction

Among the eighteen Tibeto-Burman languages (henceforth TB) found in Bhutan, seven belong to the Tibetic group or “Gangjong Bhoti” (གངས་རྒྱུང་སྐད་ཀྱི་བོད་སྐད་ཀྱི་ཚོགས་པ་)¹, which was earlier called “Central Bodish” (see Tournadre, 2014). These languages include Dzongkha རྫོང་ཁ་, Chocha-ngachakha རྩ་ཅ་ཅ་ཁ་, Lakha² ལ་ཁ་, and Merak-Saktengkha³ མེ་ར་ག་ས་ག་སྐད་ཁ་, Layakha⁴ ལ་ཡ་ཁ་, Durkha⁵ དུ་ར་ཁ་ and Trashigang Kham བཏཱ་ཤིག་ག་ཁ་མཁས་སྐད་. ⁶ Lakha, Merak-Saktengkha, Layakha and Durkha are remnants of yak-herding pastoralist communities, while Dzongkha and Chocha-ngachakha were traditionally cultivator communities.

These seven languages are all derived from a form of Old Tibetan (hence OT)⁷ and are closely related to Classical Tibetan (hence CT) or *CHOS SKAD*⁸ (ཆོས་སྐད་), as it is often referred to in

¹ The term Tibetic, now widely used by the community of linguists, has replaced the earlier terms of ‘Tibetan dialects’ and ‘Central Bodish’, which presupposed the existence of a Bodish branch of TB. However, lumping together Tibetic and Tamangic languages with Tshangla and Bumthangish languages (also called ‘East Bodish’) is highly problematic. The historical comparative methodology has not yet provided common innovations that would delimit clearly the Bodish subgroup. Since the term ‘Tibetic’ *cannot* easily be rendered in written Tibetan, we propose using the term ‘Gangjong Bhoti’ (གངས་རྒྱུང་སྐད་ཀྱི་བོད་). For a recent classification of the Tibetic family, see Tournadre (2014).

² Lakha is spoken in Wangdue Phodrang district in Sāphuk gewog.

³ Merak-Saktengkha is spoken in Trashigang district, in the villages of Merak and Sakteng.

⁴ Layakha is spoken in Gasa district and in the northern regions of Thimphu (Lingzhi gewog) and Punuakha districts. Layakha is considered by van Driem (1998) to be a variety of Dzongkha, but their relationship needs further investigation.

⁵ Durkha is spoken in Bumthang district, Dur village.

⁶ See also the Dzongkha Development Commission site in Dzongkha:
http://www.dzongkha.gov.bt/indigenous_languages/index.php.

⁷ About Old Tibetan and Proto-Tibetan, see e.g. Hill (2010), Tournadre (2014), and Jacques (2014).

⁸ See the transcription convention at the end of the article.

Dzongkha. This proximity to CT makes it easy to transcribe most words using the Classical orthography. Despite their close genetic relationship, however, these languages do not allow for good mutual intelligibility.

The TB languages of Bhutan include several *non-Tibetic* languages, such as Tshangla, also referred to as Sharchopkha, “the eastern language” (the TB language with the second largest number of speakers after the national language, Dzongkha). Others are the so-called “East Bodish” group of languages (see below) and a few other languages that belong to the “Tibetosphere” (Noonan, 2012; Tournadre, 2014). The non-Tibetic “Tibetospheric” languages have been influenced greatly by CT, as well as by Dzongkha and other Tibetic languages, but they are not derived from OT, and they differ significantly from the Tibetic languages in their core vocabulary and grammar. Other TB languages of Bhutan, such as Gongduk⁹ and Lhokpu, differ entirely from Dzongkha and other Tibetic languages in every aspect of their vocabulary and grammar.

The “East Bodish” languages, which we prefer to call “Bumthangic”,¹⁰ include the following: Bumthangkha, Khengkha, Kurtoepkha, Phobjikha, Chalikha, Dzalakha, and Dakpakha (see Mazaudon, Michailovsky, Hyslop 2013). From a phylogenetic point of view, these Bumthangic languages, together with the Tamangic languages of Nepal, are the closest cousins of the Tibetic group. In fact, the Bumthangic languages are probably slightly closer to the Tibetic languages than the Tamangic languages. For example, unlike Tamangic languages, which have a negation in /a/, Bumthangic languages have a negation in /ma/ and /mi/ just as the Tibetic languages.

But let’s turn back to the Tibetic languages, or གངས་རྒྱལ་གྱི་སྐད་ “Gangjong Bhoti languages”, of Bhutan. Among these languages – Dzongkha, Chocha-ngachakha, Lakha, Merak-Saktengkha, Layakha, Durkha and Trashigang Kham – only Dzongkha has been extensively studied, while the others are still not described. The understanding of their genetic relationship, however, is important for the historical reconstruction of the Tibetic family.

As early as 1992, van Driem wrote in his Dzongkha grammar that Chocha-ngachakha may be an “older sister of Dzongkha”. He also added that “a survey of this language would shed much light on the historical development of its sister language Dzongkha.”

The description of the Chocha-ngacha language (hence CN) is indeed particularly significant because, as we will see, it has preserved many ancient features, most notably in its phonology, grammar, and vocabulary. G. Hyslop,¹¹ who has conducted fieldwork on the Tangmachu dialect (ཐག་མོ་ཅུ་ stag-mo chu) of CN spoken in Lhuentse district, suggests that CN could have an “East Bodish” substrate and that the language could have undergone an “entire relexification.” We will examine this hypothesis, which is essentially based on phonological data.

This paper is the first publication proposing a preliminary study of the CN basic phonology, grammar, and basic vocabulary. We base this paper upon the variety of language spoken in the Tokari

⁹ See van Driem 2014.

¹⁰ The Term ‘East Bodish’, proposed by Shafer (1955), is problematic because it implies a division of the Tibetic languages between Balti, Ladakhi (West Bodish), and Central Bodish, the other Tibetic languages. Bradley (1997) has used the term Bodish in a very different way. In his classification, the term Central Bodish becomes equivalent to the term ‘Tibetic’, according to the definition found in Tournadre (2014). This latter term is clearer because purely geographic terms are often ambiguous and fuzzy. Bradley’s Western Bodish corresponds to the Mazaudon’s Tamangic languages, or TGTm (see Mazaudon). Again, the later term ‘Tamangic’ is much clearer. From a geographic point of view, Bradley’s ‘Western Bodish’ languages are spoken in a region that lies nearly 1000 kms *east* of the ‘Central Bodish’ languages of Ladakh and Baltistan.

¹¹ Personal communication.

village of the Tsamang gewog.¹² Since this is a preliminary survey, some aspects of the description require confirmation and others call for further research, but even so, this paper fills a gap in the documentation of Bhutanese languages related to Dzongkha

2 Sociolinguistic and geographic presentation of Chocha-ngachakha

Chocha-ngachakha, pronounced [təhotəa ɪateə kha], is the second Tibetan language of Bhutan and has about 20,000 speakers (van Driem, 1998). This might be an overestimation, however. According to Yeshe Thinley, head of Tsamang village, there are about 7,000 speakers in the two main villages of Tsamang and Tsakaling, so the global figure might be closer to 15,000 speakers. There are also a significant number of non-native speakers, however, particularly in Mongar, Lhuentse, and Trashy Yangtse.

In terms of numbers of speakers, Dzongkha, the national language of Bhutan, has the most native speakers (roughly 160,000) and is also spoken as a second language by many Bhutanese. The number of speakers for the five other languages is more limited. According to van Driem (1998), Lakha has about 8,000 speakers; Merak-Saktengkha 3,000; Layakha 1,000; Trashigang Kham 1,000; and Durkha 300.

Chocha-ngachakha is an exonym meaning *choca*, ‘you’ (pl), and *ngaca*, ‘us’ (pl), and is a language now widely used throughout Bhutan. Local people still use loconyms such as Tsamangpe-kha ཅུམང་པའི་ཁ་ and Tsakalingpe-kha ཅ་ཀ་ཐེང་པའི་ཁ་, referring to the dialects spoken in the two main villages.¹³

For the most part, the CN speaking area is found along the lower course of the Kuri river (ཤུ་རི་ཐུ). For this reason, speakers are often called “kurmatpa” ཤུར་ཐང་པ་, which literally means [people living] in the lower course of the Kuri river, as opposed to the “kurtoetpa” ཤུར་ཐོང་པ་, who dwell in the upper valley and speak “Kurtoetpkha”, a non-Tibetic language related to Bumthangkha. Ironically, CN speakers traditionally called their language “Kurtoetpekha” ཤུར་ཐོང་པའི་ཁ་, ‘the language of the upper valley inhabitants’ (they lived higher in the valley than those from Gyelposhing or Mongar town). Some people who have the perception that Chocha-ngachakha is close to Classical Tibetan call their language “Chöke”, the ‘lithurgical language’, or ‘Dharma language’.

Chocha-ngachakha is essentially located in Mongar, Lhuentse, Trashy Yangtse, and Trashigang districts. The dialect described in this article is spoken in the Tokari hamlet of Tsamang gewog, which is located less than 50 kms from the town of Mongar. The elevation, latitude, and longitude of Mongar are 1000 m, 27, 26N and 91, 28E, respectively. This means that among the Tibetan languages, CN is spoken at the **lowest elevation** and the **lowest latitude**.¹⁴

¹² Gewog corresponds to the English rendering of Dzongkha: རྒྱུད་ལོག་ *rgad-'og*, pronounced /geo/, lit. ‘under a head/chief.’ Several orthographies are found in English: geo, geog and gewog. The latter seems to be the most frequent and has the advantage of avoiding confusion with the English formant *geo*, as in geography.

¹³ We would prefer to use the name ‘Tsamangkha’ (the name of one of the main dialects) to refer to the language as a whole, rather than the exonym Chocha-ngachakha, but since the latter name is now widely used in Bhutan, we use it as the general term to refer to this language.

¹⁴ As a comparison, we provide the latitude of some towns of the Tibetan area: Huari (Qilian, Gansu, China): 38, 09; Rebkong (Qinghai, China) 35, 51; Skardo (Pakistan): 35, 29; Leh (Jammu Kashmir, India): 34, 15; Yushu (Qinghai, China): 32, 99; Kaza (Spiti, HP, India): 32, 22; Derge (Sichuan): 31, 80; Lhasa: 29, 64; Dolpa (Nepal): 29, 05;

In Mongar, CN is spoken in Tsakaling, Tsamang, Sheri Muhung, and Saling gewogs. In Lhuentse, it is spoken in Jare, Minje, Menbi, Tsankhar, and Metsho Ungar gewogs, and in Trashi Yangtse, it is spoken in Tongshang and Gangkhapa gewogs. In Trashigang, speakers of CN are found in Bartsham, although according to local history Bartsham speakers originally came from Tsamang village. Due to the Bhutanese resettlement policy, speakers of CN can now also be found in the southern districts of Samtse and Sarpang.

CN is also used as a second language by Kurtoepkha or Chalikha speakers and generally by most people from Lhuentse and Trashi Yangtse districts. CN speakers are in contact with Khengkha speakers in the southeast, Bumthangkha speakers in the west, Dzalakha and Kurtoepkha speakers in the north, and Chali and Tshangla speakers in the east.

Along with Dzongkha, the national language, many CN speakers are able to have a conversation in Tshangla and Bumthangkha or in other closely related Bumthangish or “East Bodish” languages, such as Kurtoepkha and Khengkha. They rarely, however, speak the Chali language, which they call */phyalikha/*, nor Dzalakha.

Since English and Dzongkha are used for the medium of education in Bhutan, the Chochangacha language is not taught in schools, nor is it used by the television or radio media, which use only English, Dzongkha, Nepali, and Tshangla. No written transcription has been adopted for CN, and it remains a purely spoken language, making it vulnerable. A written language could easily be developed, however, since the spoken language is very closely related to Classical Tibetan (see the examples and appendix of this paper).

On the other hand, geography has aided the survival of the CN language and culture, as the main villages are very isolated from main roads and other towns. Until 2014, for example, some Tsamang villages could be reached only by foot. The existing mud roads make transportation quite difficult, and thus the area maintains a fairly high degree of isolation. The hamlets are scattered in the jungle, and there is no central village in either Tsamang or Tsakaling.

The area’s isolation also certainly contributed to the emergence of a dialectal variation. CN dialectal variation is, however, relatively limited. The differences between the dialect of Tsakaling and that of Tsamang, although separated by the Kuri river, are not very significant, and there is very good mutual intelligibility, at least up to Autsho (*Au-mtsho*) village in Lhuentse district. In the upper valley of Minje gewog, however, there are important phonological differences. For example, the reflexes of the bilabial of *by* turn into */sh/*. *bya-po* */byapol/* (in Tsamang) is pronounced */shapol/*, and *byas* */bya:/* ‘to do’ [past stem, in Tsamang variety] is realized as */sha:/*.¹⁵ It seems that the dialectal variation is stronger in the upper valley in Lhuentse district because of contact with Kurtoepkha speakers. The isolated dialects spoken in Trashi Yangtse and Trashigang probably present more specific features, but data is needed for the dialectal classification.

CN speakers live at a low altitude, about 1000 m, and their land is covered with dense forests of chir pine (ཏོག་མོ་ཤིང་ */donphöt shing/* or ཏོག་མུང་ */thangcung/*), blue pine (མུང་ཤིང་ */cangshing/*), bamboo (གཞུ་ཤིང་ */shushing/* ‘lit. bow tree’), rhododendrons (ཏོག་ཏི་མེན་ཏིག་ */arto mentok/*), as well as a large diversity of other trees and plants.

Gyalthang (Xiangrila): 27, 82; Namche Bazar (Sherpa, Nepal): 27, 80, Thimphu (Bhutan): 27, 47; Gangtok (Sikkim, India): 27, 33; Mongar (Bhutan): 27, 26.

¹⁵ In Palangphu (Mongar, Tsakaling gewog), we consulted Tshultrim (age 39) from Shongmey (Metso gewog, Lhuentse), and in Autsho village we consulted Yeshe Gyeltshen (age 28) from Jare gewog, in Lhuentse.

The Kuri river (ལུ་རི་ཆུ་) or Kurchu (ལུ་རུ་ཆུ་), has a sacred dimension and is considered to be blessed. According to folk tradition, the river flows from the mouth of a statue (hence the orthography *sku-ri* in CT) and has medicinal properties.

Concerning their religion, CN speakers are followers of the Kagyü and Nyingma schools of Vajrayana Buddhism,¹⁶ but they have preserved some local religious cults that are not characteristic of the Kagyü and Nyingma sects, such as the worship of *Guru zhal* (གུ་རུ་ཞལ་ / *Guru she:/*).¹⁷

The main *yulha* or tutelary mountain is the *Gogphel lha* གོག་པེལ་ལྷ་, the region's highest peak. Given the altitude, the region receives little snowfall, so CN speakers are traditionally cultivators, with maize (འབྲུ་ / *bra:/*), rice (འབྲུ་ / *bra:/*) and buckwheat (བྲུ་ / *braol*) as their main crops. Unlike other Tibetic speaking areas, Chocha-ngacha communities have a great diversity of fruit trees, including peach (ཁམ་བུ་ / *khambul*), orange (ཆ་ལུ་ / *tshalul*), pomegranate (ཆ་ལེ་ / *tshalem*), wild banana (ང་ལ་ / *ngala*) and domestic banana (ཆེ་ / *ce*), mango (ཨ་མ་ / *am*), and fig (ཁོ་མ་ལྷ་ / *khomdang* / *ridang*)¹⁸, as well as passion fruit (ལུ་ལོ་ / *curpol*), guava (བེ་བུ་ / *bepsu*)¹⁹, lemon (ཀ་པུ་ / *kapur*) and Asian pear (ལི་ཏོང་ / *litong*).

Chocha-ngacha communities also breed standard cows (འཇ་ / *jɛwa*) and a crossbreed of mithun and cow (བ་ / *ba*) for milk,²⁰ horses (ཏ་ / *ta*), mules (དྲེ་ / *drɛ*) to carry loads, and pigs (ཕག་པ་ / *phakpa*).²¹ The Chocha-ngacha people are also notable for their rich handicraft tradition, particularly wood carving and basket weaving.

3 Introduction to the phonology

Since the great majority of CN words are cognates with Classical Tibetan and the reflexes are quite regular, we present the phonology in relation to the language's linguistic ancestor. The phonology of CN is characterized by the preservation of ancient pronunciations reflected in Classical Tibetan.

3.1 The conservative reflexes of the syllable initials

As we will see, several features of the CN Tsamang dialect exhibit some very archaic pronunciations, which are rarely attested in most other Tibetic languages.

3.1.1 Voicing of the initial consonants

Voiced plosives, fricatives, and affricates without the preradical of CT have been preserved. Ex.: བུ་ *BU* / *bu* 'son', ཇ་ *JA* / *ja* 'tea', ཏ་གོ་ *HA GO* / *ha go* 'understand', དུག་ *DUG* / *du* 'poison', བྲང་ས་ *BRANG-SA* *brang-sa* (CT: dwelling place) / *brangsas* 'hut', ཇ་ *ZA* / *za* 'to eat', མོ་བ་ *ZOR-BA* / *zɔra* 'sickle', and རྣ་

¹⁶ A giant statue of Guru Rinpoche has recently been built above the village of Tangmachu (ཏང་མ་ཆུ་) in Lhuentse district.

¹⁷ Guru Zhal is not related to Guru Rinpoche or Padma Sambhava.

¹⁸ Ridang is a wild species of fig.

¹⁹ See CT: བེ་བུ་ལེ་ *BAL.BO'I SE'U*.

²⁰ They do not kill cows for the meat in Tsamang.

²¹ The area is too hot for sheep, and locals do not keep goats because of their low milk productivity.

DOM /dɔm/. In contrast, in nearly all the modern Tibetic languages, such as Ü-Tsang, Kham, Dzongkha, Lhoke, Ladakhi, Sherpa, etc., those initial sounds are pronounced as voiceless: 𑄧 ZA ‘to eat’ /sal/, 𑄧 BU ‘son’ /pul/, 𑄧 JA ‘tea’ /chal/, etc.

Thus, in Chocha-ngachakha the reflexes of voiced initials with preradicals are pronounced in the same way as those without preradicals and differ only in their suprasegmental features (see section 3):

Compare for ex.: 𑄧 BU /bɯ/ ‘son’ vs. 𑄧 BU /bɯ²/ ‘crawling insect’, 𑄧 DUG /dɯ²/ ‘poison’ vs. 𑄧 DU² /‘to sit’, 𑄧 SGO-BA /gɔ²a/ ‘door’ vs. 𑄧 HA GO ha gɔ /ha gɔ¹/ ‘understand’, 𑄧 BRANG-SA /brangsəl/ ‘hut’ or 𑄧 BRANG /brangtol/ ‘chest’ vs. 𑄧 SBRANG.MA /brangmal/ ‘house fly’.

The only major exception to the preservation of voicing is the initial palatal fricative ZH. In our data, this initial sound is always unvoiced even when it is preceded by a preradical : 𑄧 BZHI /shj/ ‘four’, ZHO /shɔ/ ‘milk’, ZHAL /she/ ‘mouth (H)’.

3.1.2 The pr, phr, br series

The combinations *pr*, *phr*, *br* (with or without preradical) are also well preserved in Chocha-ngachakha:

Ex.: 𑄧 BRAG /brak/ ‘cliff’, 𑄧 BRO.BA /broal/ ‘taste’, 𑄧 BRA.BO /braol/ ‘buckwheat’, 𑄧 BRANG-SA /brangsəl/ ‘hut’, 𑄧 BRANG /brangtol/ ‘chest’, 𑄧 BROG.PA /brokpa/ ‘pastoralist’, 𑄧 BRAS /brɔ²/ ‘rice (paddy)’, 𑄧 BRONG /brɔng²/ ‘wild yak’, 𑄧 SPRA /pra/ ‘monkey’,²² 𑄧 SBRUL /bri²/ ‘snake’, 𑄧 SBRANG.MA /brangmal/ ‘house fly’, 𑄧 BRAD /brat/ ‘to scratch’, 𑄧 DPRAL.BA /preal/ ‘forehead’, 𑄧 PHRENG.MA /phrengmal/ (< 𑄧 PHRENG.BA) ‘rosary’.

This sound change is totally regular, and whenever it does not occur it is likely to be a loanword. For example, the normal reflexes for ‘monkey’ and ‘snake’ are respectively /pra/ and /bri/; however, when they refer to the astrological twelve animal cycle, the forms 𑄧 tre-lo/ and 𑄧 dri-lo/ are used. Retroflex sounds are influenced by the pronunciation of Lhasa Tibetan and Dzongkha. The same is true for the words 𑄧 BRUG ‘dragon’ and 𑄧 BRI ‘to write’, which correspond respectively to /druk/ and /dri/ in Chocha-ngachakha. The expected labial sounds are not found because these two words are typically imported concepts from mainstream academic culture.

Unlike what has happened in nearly all Tibetic languages, in Chocha-ngachakha the combination *db* has preserved a labial stop: 𑄧 DBANG /bɔng/ ‘power’ (compare with Ü-Tsang /wāng/, Dzongkha /wāng/ or Amdo /ɣang/).

The combination 𑄧 DBRAL ‘to tear’ is unique in that it has lost the labial sound /re:/.

3.1.3 The py, phy, by series

The reflexes of the labial series *py*, *phy*, *by* (with or without preradical) are also very conservative.

Ex.: 𑄧 PHYE /phe/ ‘flour’, 𑄧 BYE.MA /bɛmal/ ‘sand’, 𑄧 BYA /bya/ ‘bird’, 𑄧 BYUNG /byung/ ‘come out’, 𑄧 /phyung/ ‘take out’, 𑄧 PHYUGPO /phyukpo/ ‘rich’, 𑄧 BYA (future of *byed*) /bya/ ‘to do’, 𑄧 DBYAR.KHA /byarkha/ ‘summer’, 𑄧 DPYID-KHA /pyitkha/ ‘spring’, and 𑄧 BYA.LCE /bɛchi/ ‘clitoris’ (lit. ‘hen’s tongue’).

²² Note that *SPRA* in CT means ‘ape’, while the corresponding form means ‘monkey’ in CN.

This sound change is completely regular. Whenever it does not occur, the word is likely to be a loan pronunciation, as in བྱང་ *BYANG* /*jang*/ ‘north’ and རྩོད་ཀྱི་ *PHYOGS* /*chok*/ ‘direction’.

As we have seen from the examples above, the voiced/voiceless opposition, as well as the aspirated voiceless/non-aspirated voiceless, play a fundamental role in the syllable onset of Chocha-ngachakha.

3.1.4 The reflex of *wasur*

In most modern Tibetic languages, the reflex of *wasur* does not yield any specific pronunciation, and this has even raised doubt about the idea that the subscribed *wa* has any phonetic function. However, in some languages, the /w/ does have an impact on the pronunciation. This is the case in some Balti and Ladakhi dialects (see Hill, 2006). We can add that this ancient reflex is also attested in Tsamang CN. Compare, for example, the following words in the Tsamang dialect: *RTSWA* /*tsoa*/ ‘grass’ and *RTSA* /*tso*/ ‘vein’ or *RWA* /*rua*/ ‘horn’ and *RA* /*ra*/ ‘goat’. Some southern Kham dialects have also preserved a reflex of the *wasur* (Hiroyuki Suzuki, p.c). Hill (2006: 90) showed that “the Old Tibetan word *vwa* ‘fox’ has *v-* [‘a ʌ’] and not *w-* as its initial, and that the medial - *w-* in Old Tibetan represents indeed a phonetic reality.

We may now say that this archaic feature is not only found in the northwestern and eastern regions, but also in the southernmost region in the Tsamang dialect of Chocha-ngachakha. Thus, the CN data support Hill’s statement.

3.2 Other phonological features of the initial consonants

Upon examination of other CT reflexes, we find that they are more innovative and correspond to the situation of other southern languages, such as Dzongkha and Lhasa:

- The combination *DR* yields only retroflexes.

Ex.: འཇིག་ *DRIS* /*dʒiː*/ ‘ask’ (past), འདོད་ *DRAN* /*dʒan*/ ‘remember’, འཇིག་བུ་ *DRIL-BU* /*dʒiːbu*/ ‘bell’, འདུག་ *DRUG* /*dʒuʔ*/ ‘six’.

- The combination *SR* yields /s/.

Ex.: སྐྱམ་ *SRAM* /*sam*/ ‘otter’, སྐྱམ་ *SRAB* /*sap*/ ‘horse bit’, སྐྱེག་ *SREG* /*sek*/ ‘burn, or grill’, སྐྱུང་ *SRUNG* /*sung*/ ‘to keep’, སྐྱོག་ *SROG* /*sok*/ ‘life’, and སྐྱེ་ *SRAS* /*seː*/ ‘son (H)’.

- The combinations *KY*, *KHY*, and *GY*, and the following combinations (with a preradical) *SKY*, *DKY*, *KHY*, *GY*, *SGY*, and *DGY*, yield two types of reflexes: the velar /*k*, *kʰ*, *g*/ and the palatal affricates /*c*, *ch*, *j*/.

Ex.: ཁྱི་ *KHYI* /*khi*/ ‘dog’, ཁྱིམ་ *KHYIM* /*khim*/ ‘house’, སྐྱིད་ *SKYID* /*kitongki*/ ‘pleasant, or nice’, སྐྱེས་ *SKYES* /*keː*/ ‘to be born’, སྐྱེད་པུ་ *SGYED-PU* /*getpu*/ ‘firepit made with three stones’, ཁྱེད་ *KHYED* /*chet*/ ‘you’, ཁྱེ་ *RGYAMI* /*jami*/ ‘Chinese’, བརྒྱ་ *BRGYA* /*ja*/ ‘hundred’, བརྒྱཎ་ *BRGYAD* /*jat*/ ‘eight’, ཁྱེ་གར་ *RGYA-GAR* /*jagar*/ ‘India’, སྐྱུར་མོ་ *SKYUR-MO* /*curmung*/ ‘sour’, བསྐྱུར་ *BSGYUR* /*jur*/ ‘change’.

- The combination *MY* yields a palatal nasal /ny/.

Ex.: མྱང་ *MYANG* CT: ‘to taste, experience’, /nyang/ ‘to finish’, མྱང་པ་ *MYING.PO* /nyungpo/ ‘younger brother’. This form is particularly interesting since the reflex suggest that it derives from the form *MYING.PO* found in OT,²³ and not the form མིང་པ་ *MING.PO* found in CT.

- The combinations *SL* or *BSL* yield the reflex //.

Ex.: འཇོ་མོ་གཟེངས་ /acho lo/ ‘moon’ (lit. Elder brother moon from OT: འཇོ་མོ་ *A-CO* ‘elder brother’ + སྐྱེ་བའི་ *SLA-(BA)* ‘moon’).²⁴ However, when referring to the names of the month: ‘first month’, ‘second month’, one uses the loanword ཇཱ་པ་ /dawa/ found in Lhasa Tibetan and pronounced /doa/ in CN: /doa dangpa/ ‘first month’ (of the Bhutanese calendar). Other examples include སྤོང་ *SLONG* /long/ ‘to erect’, བསྟན་པ་ *BSLABS* CT ‘study; teach’, /lap/ ‘to study’.

- The combination *LH* yields the reflex //b/ ལྷ་ཁང་ *LHA-KHANG* /lhakang/ ‘temple’, ལྷོ་ *LHO* ‘south’.

- The preradicals *G, D, B, M, ‘, R, S, L* are no longer pronounced and do not have any impact on the pronunciation when they appear with plosives or affricates.

Ex.: གསུམ་ *GSUM* /sum/ ‘three’, བདུན་ *BDUN* /dün/ ‘seven’, བཞི་ *BZHI* /shi/ ‘four’, འབུམ་ *BUM* /bum/ ‘one hundred thousand’ or ‘remnants of threads’, མདའ་ *MDA* /da/ ‘arrow’.

- With nasals, the preradical may have an impact on the tone and trigger a high tone, as in Standard Tibetan and Dzongkha.

Ex.: མ་ /ma/ ‘injury’, མན་ /man/ ‘medicine’, མྱེང་ snying /ning/ ‘heart’, མཇན་པ་ /ngönpu/ ‘blue’, མྱེག་པ་ RMIG.PA /mekpu/ ‘hoof’, གཤམ་ཤིང་ GNYA.’SHING /nyashing/ ‘yoke’, བརྟ་ BRNYA /nya/ ‘to borrow’. In Chocha-ngachakha, however, there is some inconsistency, since the tone will be low for many words: ག་ NGA /nga/ ‘five’, རྩ་ RNA.BA /noa/ ‘ear’, མངར་མོ་ /ngarmung/ ‘sweet’, འགྲམ་པ་ nyampu/ (<MNYAM.PO) ‘together’, འཇམ་པ་ /marpu/ (< DMAR.PO) ‘red’, མུར་ RMUR /mur/ ‘to chew’, རྩང་ RNANG /nang/ ‘to choke’.

3.3 Rhyme

3.3.1 The final consonants

Among the ten final consonants found in CT (*G, NG, D, N, B, M, ‘, R, S, L*), several are well preserved in Chocha-ngachakha: *G, NG, D, N, B, M, R*.

- For the final *G*, the phoneme /k/ may have various allophones depending on the context: [k], [q], [ʔ] and [g].

Ex.: ལྷུག་ *DRUG* /duk/ ‘six’, ཇིག་ /chik/ ‘one’, སྤེག་ *SREG* /sek/ ‘to burn’, བརྟ་ *BRAG* /brak/ ‘cliff’, འདུག་ /duk/ ‘to sit’.

²³ The word *MYING.PO* is found for example in PT 1068 (OTDO, Japan). For an overview of OT phonology, see Hill, 2010.

²⁴ This formulation shows that from a paleocultural point of view, in Tibetic cultures the sun and moon are seen as relatives. The moon is perceived as male, while the sun is female. More examples of this kinship are found in other Tibetic languages: cf. Sherpa /aula/ ‘moon’ (from *a-khu sla*, lit. ‘uncle moon’). This conception is also found in neighbouring Indo-Aryan languages.

- Reflex of the nasals *N*, *M*, and *NG*

Ex.: མཎ་ *SMAN* /*man*/ ‘medicine’, འདུན་ *BDUN* /*dün*/ ‘seven’, ཁྱིམ་ *KHYIM* /*khim*/ ‘house’, ལམ་ *LAM* /*lam*/ ‘road’, དབང་ *DBANG* /*bang*/ ‘power’, སྙིང་ *SNYING* /*ning*/ ‘heart’.

- Reflex of the vibrant *R*

Ex.: མར་ *MAR* /*mar*/ ‘butter’, འཀར་ཕུ་ */karpu/* (<*DKAR.PO*) ‘white’, འཇར་ཕུ་ */serpu/* (<*SER.PO*) ‘yellow’, ངར་མ་ *NGAR.MA* /*ngarma*/ ‘strong’.

- Reflex of the fricative *S*

The *S* yields a lengthening of the preceding vowel.

Ex.: ནས་ *NAS* /*na:*/ ‘barley’, ལས་ *LAS* /*la:*/ ‘work’, སྒས་ *SNGAS* /*nga:*/ ‘pillow’, གཉིས་ *GNYIS* /*nyi:*/ ‘two’, འབྲས་ *BRAS* /*bra:*/ ‘rice (in the paddy fields)’, རུས་པ་ *RUS.PA* /*ru:pa*/ ‘bone’, མོས་ *CHOS* /*cho:*/ ‘dharma’, སྐྱེས་ *SKYES* /*ke:*/ ‘birth’, འཇེས་གཅེས་ */ce:ce:/* ‘sweetheart’.

- Reflex of the labial *B*

Ex.: ཁམ་ *KHAB* /*khap*/ ‘needle’, མོབ་ *THOB* /*thop*/ ‘to get’, ལེབ་ *LCEB* /*cepl*/ ‘to suicide’, བཀམ་ *BKAB* /*kapl*/ ‘to cover with a blanket’.

- Reflex of the dental *D*

For the final *D*, the phoneme /*d*/ may be realized as glottal stop [ʔ] (particularly in final position), [t] or [d].

Ex.: བརྟེད་ *BRJED* /*jet*/ ‘to forget’, བརྟེད་ *BRGYAD* /*jat*/ ‘eight’, ཡོད་པ་ *YOD + ?PA-YIN* /*yöt-pil*/ (or /*yet-pil*/) ‘to have’, གཤ་པོ་ *GAD.PO* /*gatpo*/ ‘old man’, གཤ་མོ་ *GAD.MO* /*gadmol*/ ‘old woman’, པའ་པ་ *PAD.PA* /*patpa*/ ‘leech’.

- Reflex of the lateral *L*

The final *L* is no longer pronounced; however, it has an impact on the vowel.

Ex.: བཟ་ *BAL* /*bɛ:*/ ‘wool’, བསྐུལ་ *BSKAL* /*ke:*/ ‘to send’, ཁམ་ *KHAL* /*khe:*/ ‘a score, or unit of 20’, མཁམ་མཐ་ *MKHAL.MA* /*khe:mal*/ ‘kidney’, རྒྱལ་པོ་ *RGYAL.PO* /*ge:po*/ ‘king’, རྩེ་པོ་ */je:po/* ‘spirit’, སྐོལ་ *SKOL* /*ke:*/ ‘to boil’, གསོལ་ཇ་ *GSOL.JA* /*se:jal*/ ‘tea (H)’, འདུལ་ *DNGUL* /*nge:*/ ‘silver’, ཡུལ་ *YUL* /*yi:*/ ‘village’, རུལ་ *RUL* /*ri:*/ ‘to rot’.

As we have seen above, the phonology of CN is clearly derived from CT. The reflexes show a great regularity and the phonotactic rules correspond to those found in other Tibetic languages, but there is one notable exception: the initial cluster /*mr*/. It is found in a few CN words, such as /*mrɔk*/ ‘to mix different foods’, /*mraka-mrɔkol*/ ‘mixed together’, /*mrek*/ ‘manure’,²⁵ and /*mre*/ ‘to scratch someone’. It is of course theoretically possible that these words are borrowed from contact languages or even reflect a non-Tibetic substrate. However, we have not found any confirmation of the “non-Tibetic substrate hypothesis” nor sources for these words in the neighbouring languages.

There is yet another explanation: while the cluster /*mr*/ does not seem to have been reported in any other modern Tibetic language, it is attested in CT for a few words, such as མཎ་ *smra* ‘to speak’, མཎ་མཎ་ *SMRE.SNGAGS* ‘lamentation’, and མཎ་མཎ་ *SMREG.PA* ‘leftovers’. Thus, the meaning of

²⁵ Note the tonal difference between /*mrɔk*/ and /*mrek*/.

/mrek/ could well be derived from མྲེག *SMREG*. In support of the CT origin of the cluster /mr/, one can also note that the expression /mr̥ka-mr̥ko/ ‘mixed together’, ‘messy’, is also found in Dzongkha but pronounced /mak-mok/ (used to describe a confused way of talking). In Dzongkha the /m/ could well correspond to the reflex of CT *SMR*. That is probably the case for /mo-sh/, a very common tag expression ‘isn’t it?’, which is probably derived from *SMROS-SHIG* ‘tell me!’ If this is correct, one could propose the reconstruction **SMRAG*/**SMROG* for ‘mixed, messy’, in Old Tibetan.

The reflex /mr/ of *SMR* would thus, if it is confirmed, be another amazingly archaic feature of Chocha-ngachakha, not found so far in any other modern Tibetic language.

3.3.2 *The vowels*

As in other Tibetic languages that have preserved many archaic features, Chocha-ngachakha has a limited inventory of vowels: /i/, /e/, /a/, /ü/, /u/ and /o/. Additionally, there is a noted opposition between short and long vowels: /i:/, /e:/, /a:/, /ü:/, /u:/ and /o:/. These long vowels are the reflexes of the Classical Tibetan final *L* or *S*.

3.3.3 *Suprasegmental features*

CN is a tonal language with two distinctive tones: high and low registers. In our transcription, only the low tone is marked by a line under the vowel; the high tone is not marked. Like other Tibetic languages, CN is a word tone language, not a syllable tone language like many Southeast Asian languages, such as Burmese, Chinese, Thai, Vietnamese, etc.

The tone reflexes are in general similar to other modern Tibetic languages such as Standard Tibetan or Dzongkha, but there are a few discrepancies, notably concerning the nasals.

The tones are usually predictable from the reflexes with CT.

For the plosives and affricates, voiced initials are normally realized with a low tone; the unvoiced are pronounced with a high tone, as can be seen from the examples in sections 3.1 and 3.2. The aspirated unvoiced are pronounced with a notably lower pitch than the non-aspirated ones; however, they are still perceived as high register.

For the nasals, the general rule is that the nasal reflex yields a low tone even when it is preceded by a preradical, unlike the corresponding reflexes in Dzongkha and Standard Tibetan, which yield a high tone. There are numerous exceptions, however, as we pointed out in 3.2.

Concerning the fricatives, /s/ is pronounced with a high tone, while /z/ is realized with a low pitch (see examples in 3.2): The /h/ is usually followed by a high tone, but the pitch is lower than the /s/: /ha lək/ ‘be surprised’, /ha gə/ ‘understand’, /hoge/ ‘salad’.

The fricative /sh/ is either high or low depending on whether it is a reflex of CT *sh* or *zh*: བཞི *bzhi* /*shi*/ ‘four’ vs. ཤི *shi* /*shi*/ ‘to die’. Note that Tokari CN has no voiced palatal fricative /zha/.

Concerning the laterals, vibrants and glides, we have the following rules:

/l/, /y/, /r/, /w/ may be realized with a high or low pitch. When they are reflexes of CT initial without preradicals, there are usually realized with a low tone and in the contrary case with a high tone. For example: ལུ *yu* ‘turquoise’, ཡུང་གྲུག་ *yungkar* ‘mustard’, རུ *re* ‘cloth’, རེ་རེ་ *rere* ‘each’, དབྱས་ *re* ‘to tear’, ཡུང་པ་ *lungpa* ‘valley’, ལུང་ *lung* ‘air’, ཡ་ *wa* ‘container’, ཡམ་ *wamo* ‘fox’, དལེ་ *wae* ‘a calling word roughly equivalent to ‘hey!’

As we have seen in 3.1.1, a contour tone – maybe correlated with a phonation quality – is distinctive in some rare monosyllabic words such as 'DUG and DUG. In this article we have indicated the distinction by the exponent numbers 1 and 2. The contour 1 corresponds to a falling tone and a more tensed phonation, while contour 2 is more lax. More research is needed to describe this opposition, but in any case it is very marginal in the phonological system of Chocha-ngachakha.

3.3.4 Phoneme inventory

To summarize, we provide a chart below of the phonemes found in CN. The letters that appear in bold correspond to our transcription, and they are followed by the I.P.A. correspondents whenever they differ.

Chart 1: Consonants

		Labial	Dento-alveolar	Retroflex	Prepalatal	Palatal	Velar	Glottal
<i>Plosive</i>	<i>vcls</i>	p	t	ʈ			k	[ʔ]
	<i>aspir.</i>	ph [p ^h]	th [t ^h]	ʈh [ʈ ^h]			kh [k ^h]	
	<i>voic.</i>	b	d	ɖ			g	
<i>Affricate</i>	<i>vcls.</i>		ts		c [tɕ]			
	<i>aspir.</i>		tsh [ts ^h]		ch [tɕ ^h]			
	<i>voic.</i>		dz		j [dz]			
<i>Fricative</i>	<i>vcls.</i>		s		sh [ɕ]			h
	<i>voic.</i>		z					
<i>Lateral</i>	<i>vcls</i>		lh [l̥]					
	<i>voic.</i>		l					
<i>Vibrant</i>	<i>vcls</i>							
	<i>voic.</i>		r					
<i>Nasal</i>		m	n		ny [ɲ]		ng [ŋ]	
<i>Semi-vowel</i>		w				y [j]		

Chart 2: *Vowels*

	Front		Central		Back	
	unrounded	Rounded	unrounded	rounded	unrounded	Rounded
High	i			ü [u]		u
Mid high	e					o
Mid low						
Low	a					

3.4 Preliminary conclusions about the phonology

We have seen that Tsamang CN exhibits a certain number of archaic phonological features. To understand the degree of archaism, it is useful to compare the characteristics of Tsamang CN to other languages of the family. We have selected seven archaic phonological characteristics that are preserved in a minority of Tibetic languages. They correspond to the (relative) preservation of Classical Tibetan preradical,²⁶ postradical, and final consonants of a syllable: 1) non-nasal preradical consonants (*g, d, b, r, l, s*); 2) nasal preradical consonants (*ʼ, m*); 3) the postradical glide /y/ after labial radicals (*p/by*); 4) the postradical vibrant /r/ after labial radicals (*p/br*); 5) the reflex of a postradical *w*; 6) the final consonant plosive dental /t/ (a reflex of *d*); and 7) the preservation of the final consonant /s/ (a reflex of *s*). As shown in chart 3 below, among the fifty or so Tibetic languages, only a handful exhibit some of these archaic features, all of which are spoken in the periphery of Tibet and none in the centre. The languages that cumulate in the most archaic phonological features are undoubtedly those spoken in the northwest regions: Balti, Purik, Ladakhi with 5 out of the 7 features, and Amdo in the northeast with 3 out of the 7. In the south, Chocha-ngacha, with 4 out of the 7, has retained most of these archaic features (see the chart 3 below).

It is thus clear that Tsamang CN has preserved the most archaic features of the southern Himalayas. These conservative characteristics are not confined to the phonology. They are also attested in the lexicon and in the grammar (as we will see below).

²⁶ The 8 preradicals (*G, D, B, M, ʼ, R, L, s*) correspond in the Tibetan grammatical tradition to the *sngong-ʼjug* (*G, D, B, M, ʼ*) and *mgo-can* (*R, L, s*) consonants that precede the radical letter called *ming gzhi*. Preradicals have very specific phonetic properties. For example, they are pronounced in a light way, compared to the radical consonant. According to the Tibetan grammatical tradition, there are 4 postradicals (*R, L, Y, w*), which are called '*dogs-can* 'attached (letters).'

Chart 3: *Archaic phonological features in the Tibetic languages*

Preserved characteristics	Languages and regions
<i>Preradical</i> Non-nasal+C	Ladakhi (NW), Balti (NW) Purik (NW), Amdo (NE) Kham (SE)
Nasal+C	Amdo (NE), Hor (SE) Kham (SE), Ngari (C), Ü (C)
<i>Postradical</i> Labial + /Y/	Ladakhi (NW), Balti (NW) Purik (NW), Ngari (C), Spiti (W), Chocha (S), Drenjong (S)
Labial + /R/	Ladakhi (NW), Balti (NW) Purik (NW), Kyirong (SW), Chocha (S)
C+/W/	Chocha (S), Southern Kham (E)
<i>Final consonants</i> C+/T/	Ladakhi (NW), Balti (NW) Purik (NW), Chocha (S), Amdo (NE)
C+/S/	Ladakhi (NW), Balti (NW) Purik (NW)

4 Elements of grammar

4.1 The verbal predicate

Among the striking features of Chocha-ngachakha, one should mention the quasi-invariability of the verb. In contrast to Amdo or even Central Tibetan, the ancient verbal morphology found in Classical Tibetan for the various tenses²⁷ has not been inherited in this language, but unlike Dzongkha, it has neither developed any innovative verbal morphology. That means that there is only one verb stem independent of the tense, aspect, and modality. Thus, /*duk*/, /*tang*/, /*song*/ are the only stems used for the three tenses of ‘to sit’, ‘to do (light verb)’, and ‘to go’. However, there is a small trace of the Classical morphology. Some verbs in the past tense ending in a vowel have a lengthening (indicated in our transcription by a semicolon and in transliteration by the letter *a*) that corresponds to the reflex of the final *s* of the past in CT: ཏཱ་ /*ngu:*/ ‘to cry’, ཏཱ་ཤ་ /*ko:*/ ‘to dig’, ཏཱ་ཟ་ /*za:*/ ‘to eat’, ཏཱ་བྱ་ /*bya:*/ ‘to do’. Compare this with their CT correspondents: ཏཱ་ *ngus* བཤྱ་ *brkos* བཟས་ *bzas* བྱས་ *byas*.

²⁷ Such as ‘to eat’: *za* (present), *bza*’ (future), *zos* (past, original form) or *bzas* (past, analogical form), *zo* (imperative); ‘to plant’: *’debs* (present), *gtab* (future), *btap* (past), *thobs* (imperative).

Another exception is the existence of a suppletive form for the imperative of the verb *ong* ‘to come’: /*shok*/ ‘come!’ Apart from these exceptions, the lexical form of the verb is invariable.

4.1.1 Interrogative marker

Questions are marked by an interrogative marker /*a*/ or /*ya*/, which occurs after the final auxiliary or suffix.

- after /*pi*/: /*pi-a*/
- after /*do*/ and /*di*/: /*di-a*/
- after /*sang*/: /*sang-ya*/

1) Ex: ཇི་ཕྱེ་ལུ་ར་ཅ་དེ་ཨ་ (Dz: ཇི་ཕྱེ་ལུ་ལེ་ཅ་དེ་ག་)
chet khura za-di-a
2sg+abs bread+abs eat-asmp-q
‘Are you eating/do you eat bread?’

The suffix /*te*/ cannot be followed by the interrogative particle /*a*/: */*te-a*/. This is consistent with the fact that /*te*/ is probably the reflex of a connective *ste~te~de* found in CT and not an auxiliary verb, and thus it cannot bear an interrogative mark.

The final interrogative particle is not used with interrogative pronouns: ཅི་ *ci* ‘what’, ཀེ་ /*ke*/ ‘where’, ནྐ་ /*nam*/ ‘when’, and ཀཡི་ /*kayi*/ ‘who’, etc.

4.1.2 TAM

The CN language essentially distinguishes the following tenses and aspects: present, present progressive, simple past, perfect, progressive past, and future.²⁸ They are marked by final auxiliaries or suffixes. Here are some examples:

Ex: ཇི་ ཅི་ /*za-di*/ ‘I (you, s/he, etc.) eat’, ཇི་ ཅི་ /*za-do*/ ‘I (you, s/he, etc.) am eating’, ཇི་ ཅི་ (བཟས་ ལྟེ་) /*za-te*/ ‘I (you, s/he, etc.) ate’, ཅི་ ཅི་ ཅི་ /*za-temet*/ ‘you (he, etc.) have eaten’, ཅི་ ཅི་ ཅི་ /*za:-khandukte*/ ‘I (you, s/he, etc.) was (were) eating’, ཅི་ ཅི་ /*za-sang*/ ‘I (you, s/he, etc.) will eat’; ཇི་ ཅི་ /*la-bya-di*/ ‘I (you, s/he, etc.) work’, ཇི་ ཅི་ /*la-bya-do*/ ‘I (you, s/he, etc.) am (are/is) working’, ཇི་ ཅི་ /*la-bya-te*/ ‘I (you, s/he, etc.) worked’, ཇི་ ཅི་ /*la-bya-sang*/ ‘I (you, s/he, etc.) will work’; ཇི་ ཅི་ /*song-di*/ ‘I (you, s/he, etc.) go’, ཇི་ ཅི་ /*song-do*/ ‘I (you, s/he, etc.) am (are/is) going’, ཇི་ ཅི་ /*song-te*/ ‘I (you, s/he, etc.) went’, ཇི་ ཅི་ /*song-sang*/ ‘I (you, s/he, etc.) will go.’

4.1.3 Negation

The negation markers /*ma*/ and /*me*/ are placed before the verb and stressed. They match their CT correspondents *MA* (for present and prohibitive) and *MI* (for present and future), but they are usually realized with a long vowel noted with a semicolon /:/.

For the present, the negation is /*me*/.

²⁸ Other marginal forms are also found. See 4.2.

Ex.: ◇ མེ་མེ་ཏི་ /*mɛː-zɛ-di*/ ‘X does not eat.’ ◇ མེ་མེ་ཏོ་ /*mɛː-zɛ-do*/ ‘X is not eating.’

For the past and the imperative, the negation is /*ma*/.

Ex.: ◇ མེ་མེ་ /*maː-zal*/ ‘X did not eat’, or ‘Don’t eat!’

For the future the negative form is /*mi*/.

Ex.: ◇ མེ་མེ་ /*miː-zal*/ ‘X will not eat.’

4.2 The evidential-epistemic system

Like other Tibetic languages, CN has developed a fairly rich system of evidential and epistemic marking. Evidentiality has been defined as “the representation of source and access to information according to the speaker’s perspective and strategy” (Tournadre and LaPolla, 2014). The Tibetic systems have the following specificities: Evidential and epistemic markings are indicated by a series of verb auxiliaries/suffixes that occupy the same syntactic position. From a semantic and pragmatic point of view, Tibetic E/E systems mainly encode four types of information: evidentiality (as defined above), epistemic meaning, speaker’s commitment, and speech act (pragmatic function).

As we will see, these four functions are present in CN. In the past and present tenses, there is a distinction between “assumptive”, “sensory-inferential”, and “epistemic.” Additionally an evidentially neutral form or “factual” is also found.

4.2.1 Copulative and existential verb

One of the peculiar features of CN is the replacement of the existential verb *DUG* by /*yöt*/ (CT: *YOD*) with the same sensory-inferential function as *DUG* (in Dzongkha and Standard Tibetan). As we will see later, however, the verb /*duk*/, which means ‘to sit’ in Chocha-ngachakha, has acquired a second semantic meaning, which led to its grammaticalization (see the end of this section and 4.2.3). The verb *yöt* is sometimes realized as [*yeʔ*] instead of [*yöʔ*].

For the existential verb, the assumptive form is a compound form: /*yötpi*/ . This last form is realized as [*yetpi*] or [*yötpi*].

For the copulative verb, as expected, the classical verb *YIN* is used for the assumptive. There is, however, a long form /*yinpi*/.

Both the assumptive forms /*yöt-pi*/ (existential) and /*yinpi*/ (copulative) may have a long form /*yötpitel*/ and /*yinpitel*/.

The inferential form is /*yin-cet*/ . The origin of /*cet*/ is not clear, though. Given the fact that /*yincet*/ is pronounced /*yincot*/ in some villages of Lhuentse, the origin could be *yin-rgyu-yod* ཡིན་རྒྱུ་ཡོད་.

One interesting characteristic of CN is that it is possible to use interrogative marks only with the assumptive forms, such as /*yötpi*/ or /*yinpi*/, which become /*yötpi-al*/, /*yinpi-al*/ respectively. The interrogative forms are not used with the simple copulative verbs: */*yöt-al*/ and */*yin-al*/.

The fundamental opposition in CN for the copulative and existential verbs is between assumptive, sensory-inferential, and epistemic, as shown in the chart below:

Chart 4: *Equative and existential verbs*

E/E categories	Sensory-Inferential	Assumptive or ‘weak egophoric’	Epistemic
Equative	◇ ཡན་ཅེད་ <i>yin-cet</i>	◇ ཡན་(པེ་) <i>yin(-pi)</i>	◇ ཡན་པེ་འདྲ་ <i>yin-piong</i>
Existential/locative	ཡད་ <i>yöt</i>	◇ ཡད་པེ་ <i>yöt-pi</i>	◇ ཡད་པེ་འདྲ་ <i>yöt-piong</i>

The “sensory-inferential markers” indicate that the speaker has directly observed the event or makes an inference based on direct observation. “Assumptive” or “weak egophoric” markers specify that the speaker has a good knowledge about the information that he reports and is committed to the statement. Epistemic markers are related to hypothetical situations and express various degrees of the speaker’s certainty.

The opposition between the assumptive, sensory-inferential, and epistemic marking is exemplified by the following sentences; for a contrastive approach, we also provide a translation into Dzongkha:

With the equative verb ‘yin’

2) ◇ འ་པེ་མི་དག་སྤང་ཡན་ (Dz: མ་པེ་མི་འདི་དག་སྤང་ཡན།)

phi mi gelong yin
That man monk COP+ASMP

‘That man is a monk.’

3) ◇ འ་པེ་མི་དག་སྤང་ཅེད་ (Dz: མ་པེ་མི་འདི་དག་སྤང་ཅེད་པས།)

phi mi gelong yin-cet
That man monk COP-INFR

‘Oh, that man is a monk.’

4) ◇ འ་པེ་མི་དག་སྤང་པེ་འདྲ་ (Dz: མ་པེ་མི་དག་སྤང་པེ་འདྲ་པས།)

phi mi gelong yin-piong
That man monk COP-EPI

‘That man may be a monk.’

5) ◇ འ་འདི་ང་ཡི་བུ་ཡན་ (Dz: མ་ནི་ངའི་བུ་ཡན།)

odi nga-yi bu yin
This 1SG-GEN son COP+ASMP

‘This is my son.’

6) ◇ འ་འདི་ང་ཡི་བུ་ཡན་ཅེད་ (Dz: མ་ནི་ངའི་བུ་ཡན་པས།)

odi nga-yi bu yin-cet
This 1SG-GEN son+ABS COP-INFR

‘Oh, this is my son (looking at a picture and recognizing him).’

With the existential/locative verb ‘yöt’:

- 7) ◇ མེན་འོ་མི་པན་པ་རྒྱུད་བྱ་ཆེག་གི་ཡོད། (Dz: མེན་མ་མི་ཕུམ་ལ་འདི་བྱ་ཆེ་ཆེ་འདུག)
en ophi penpalung byachiki yöt
 Oh over there butterfly beautiful+ABS exist+SENS

‘Oh! There is a beautiful butterfly over there.’

- 8) ◇ ལྷ་མང་དེ་ཨར་རྟ་མེན་རྟག་ཡོད་པུ་ (Dz: ལྷ་མང་ལུ་ཨ་རྟ་མེ་རྟག་ཡོད།)
tsamang-nge arto mento yöt-pi
 Tsamang-DAT rhododendron+ABS exist-ASMP

‘There are rhododendron flowers in Tsamang (the speaker knows this very well).’

- 9) ◇ ལྷ་མང་དེ་ཨར་རྟ་མེན་རྟག་ཡོད། (Dz: ལྷ་མང་ལུ་ཨ་རྟ་མེ་རྟག་འདུག)
Tsamang-nge arto mento yöt
 Tsamang-DAT rhododendron+ABS exist-SENS

‘(I just discovered that) there are rhododendrons flowers in Tsamang.’

- 10) ◇ ལྷ་མང་དེ་ཨར་རྟ་མེན་རྟག་ཡོད་པུ་འོང་ (Dz: ལྷ་མང་ལུ་ཨ་རྟ་མེ་རྟག་ཡོད་པུ་འོང་)
Tsamang-nge arto mento yöt-piong
 Tsamang-DAT rhododendron+ABS exist-EPI

‘There are probably rhododendron flowers in Tsamang (inference that leads to a high probably).’

- 11) ◇ རི་དང་ཁྱིམ་པོ་ཡོད་ / ཡོད་པུ་ / ཡོད་པུ་འོང་ (Dz: རི་དང་ཁྱིམ་རྟག་རྟག་འདུག / ཡོད་ཡོད་པུ་འོང་)
ridang shimpo yöt / yöt-pi / yöt-piong
 wild fig+ABS tasty COP+SENS/-ASMP/-EPI

‘The wild figs are tasty (the speaker tasted them / the speaker knows this very well / it looks like they are tasty [speaker’s epistemic inference]).’

The verb of existence /*duk*/

As in CT and several modern languages of eastern Tibet, the verb ‘*DUG*’ means ‘to sit’. This is also the case in Chocha-ngachakha. As expected, this verb also has the derived meaning ‘to stay’. What is original, however, is that the verb has acquired the meaning of ‘to exist’ without developing into a full copulative verb, unlike what happened in many modern central, southern and western Tibetan languages. Thus it allows us to see a very clear grammaticalization path:

‘To sit’ > ‘to stay’ > ‘to exist’ > *evidential*

The copula ‘*DUG*’ has not yet acquired the grammatical function of *sensory evidential* in CN but has clearly developed the meaning ‘to exist’, as shown in the examples below. However, as we will see later, it has developed an evidential function as an auxiliary verb.

- 12) ◇ གནའ་པ་ཟམ་པ་འདུག་ལྟེ་མན་ད་ལྟ་ཡོད་པུ་ (Dz: ཉེ་མ་ཟམ་ཡོད་མེན་ན་ ད་ལྟ་འདུག་ག་)
na:pa zampa duk-te-man data yöt-pi-a
 in the past bridge+ABS exist-PAST-TAG now exist-ASMP-Q

‘There used to be a bridge [there]. Is it still there?’

- 13) གནའ་པ་ནགས་ཚལ་ནང་དེ་རག་གཤམ་འདུག་ཟླ་ (Dz: ཉེ་མ་ནགས་ཚལ་ནང་ལྷ་སྤྱུ་དཀར་པོ་ལེ་ག་སྤྱོད་ཡི།)
na:pa naktse: nang-nge, raksha mangku duk-te
 in the past forest in-DAT langur many+ABS exist- PAST

‘Previously, there used to be many langurs (species of monkey) in the forest.’

- 14) ང་ཡི་ཕོ་རུང་ནང་དེ་ཆང་འདུག་ཅི། ཀ་ཡིས་འཐུང་ཅི། (Dz: ངའི་ཕོ་རབ་ནང་ལྷ་ཆང་སྤྱོད་ཡི། ཀ་གིས་འཐུང་ཡི།)
nga-yi phoru nang-nge chang duk-pi, kayi thung-pi
 1SG-GEN cup in-DAT chang+ABS exist- PAST+ASMP who drink- PAST+ASMP

‘There was chang in my cup. Who drank it?’

- 15) གད་པོ་གྲིས་ལ་ཆོ་ལོ་ཐེག་ནའ་པ་ཟམ་པ་འདུག་པུ་ལོ། (Dz: མ་ནས་གྲིས་སྤུའ་ཆོ་མ་ཉེ་མ་ཟམ་སྤྱོད་ཡི་ལོ།)
gatpo-k lap-tse ophe na:pa zampa duk-pi-lo
 old-ERG say over three in the past bridge+ABS exist- PAST+ASMP-HS

‘The old man said there used to be a bridge [there] a long time ago.’

Final discursive clitics /no/ and /an/

After the assumptive forms /yin(pi)/ and /yotpi/, it is common to add a sentence final particle /no/ in order to alert the hearer.

- 16) མ་ཉམ་ནང་དེ་ཙོ་ཡོད་པུ་ཙོ། (Dz: གེ་ཙ་ནང་ལྷ་ཙོ་འདུག་སྤྱུ།)
aham nang-nge nor yöt-pi-no
 maize in-DAT cow+ABS exist-ASMP-WARN

‘There is a cow in the maize field!’ (This requires a swift reaction!)

- 17) ང་ཡི་ཆང་ཡོད་པུ་ཙོ་ མ་དགོག་ཤེ། (ཞིག་) (Dz: ངའི་ཆང་ཡོད་སྤྱུས་ མ་དགོག་སྤྱུ་རེ།)
Nga-yi chang yöt-pi-no, ma trok-she!
 1sg-GEN chang+ABS exist-ASMP-WARN NEG touch-IMP

‘My chang is over there. Don’t touch it!’

- 18) ཟེ་མ་ རྒྱུས་མ་ལྷོ་ཡིན་ནོ། (Dz: རྒྱུ་མ་མིན་སྤྱུ།)
zima yin-no
 fake+ABS COP-WARN

‘This is a fake.’ (‘Be careful!’)

Another particle, /pan/, which is used for tag questions, may be used to seek confirmation or consensus from the hearer:

- 19) ཟེ་མ་ རྒྱུས་མ་ལྷོ་ཡིན་པན། (Dz: རྒྱུ་མ་མིན་སྤྱུ་པན།)
zima yin-pan
 fake+ABS COP-TAG

‘This is a fake, isn’t it?’

20) ལྷ་མང་ངེ་ཨར་རྟ་མེན་ཡོད་པན་ (Dz: ལྷ་མང་ལུ་ཨ་རྟ་མེན་ཡོད་མེན་ན།)
Tsamang-nge arto mento yöt-pan
 Tsamang-DAT rhododendron+ABS exist-TAG

‘There are rhododendron flowers in Tsamang, aren’t there?’

4.2.2 Auxiliary verbs of the present tense

Various auxiliaries are used to indicate both the tenses and evidential/epistemic meanings.

The ending /di/ is used for the assumptive present, while /do/ is used for the participatory-sensory progressive present.

Finally the form /di-ong/ is used for the epistemic present indicating that the speaker is not certain about his assertion (see Chart 5 and the examples below).

Chart 5: Present auxiliaries

	Assumptive	Participatory-Sensory	Epistemic
Present	V-di		V-di-ong
Progressive present		V-do	V-di-ong

In the Tibetic languages, the sensory meaning when referring to an outer observation normally occurs with the second or third person and not with the first person. However, in Dzongkha and in CN, some markers related to direct observation may be used with the first person. In order to distinguish the purely sensory markers from these markers, we propose to use the label “participatory-sensory.”²⁹

21) ད་ནང་མང་ལྷ་མོང་གི་ (Dz: ད་ནང་མང་ལྷ་མོང་གི་)
data nga mongar song-do
 now 1sg+abs Mongar go-pres+psens

‘Now I am going to Mongar.’

22) ཆར་པ་བཏང་དོ་ (Dz: ཆར་པ་རྒྱུ་བཏང་ལ།)
charpa tang-do
 rain LV -pres+psens

‘It is raining! (observing the rain falling).’

23) ཆར་པ་བཏང་དེ་འོང་། (Dz: ཆར་པ་རྒྱུ་བཏང་དེ་འོང་།)
charpa tang-di-ong
 rain LV- psens-epi

‘It must be raining! or ‘It must rain’ (epistemic inference).

²⁹ In Dzongkha that is the case of the suffix *yi* ‘witness past’ in Dzongkha (see Driem 1992). We propose to call *yi* a ‘participatory-sensory’ marker.

As in the case of the copulative and existential verbs, it is possible to add the particle /no/. The marker /do/ undergoes a morphophonological change: /di-no/.

- 24) ◇ ཆར་པ་བྱང་དི་ནོ། (Dz: ཆར་པ་བྱང་དེས་སྤྲུལ།)
charpa tang-di-no
rain LV- psens- warn

‘It is raining!’ (Be careful! The things outside will get wet!)

4.2.3 Auxiliary verbs of the past tense

For the past, CN distinguishes ‘simple past’, ‘progressive past’, ‘perfect’, and ‘present perfect continuous’, as well as ‘assumptive’ (or ‘weak egophoric’), ‘sensory-inferential’, and ‘epistemic’ statements, as summarized in Chart 6 below.

For the simple past, two forms are nearly interchangeable: V-te/ V-pi, but they differ somehow in a subtle way. The latter form /-pi/ is assumptive and insists on the speaker’s knowledge. It is more restricted in use than /-te/ and is more correlated with the first person subject than with the second or third person subjects, although it does occur frequently with all the three persons.

The former marker *-te* is evidentially neutral and presents the information as a fact. There is another distinction between the assumptive form /-pi/ and /-te/: It is impossible to form a question with the *-te* suffix, as we have already seen earlier: *V-*te-a*. The question form is: V-*pi-a*.

The marker *-te* is clearly derived from the CT connective *STE/TE/DE*. In Tsamang CN, this marker has two allomorphs: /*te*/ and /*de*/, depending on the final consonant. After the final dental /*d*, *n*/, and the vibrant /*r*/: /*de*/. With other consonants /*m*, *p*, *k*, *ng*/ and vowels, the allomorph is /-*te*/.

- 25) ◇ ཤང་བཀག་ཆེ། (Dz: ཤང་བཀག་ཡི།)
shing kak-te
wood+ABS chop-PAST

‘(S/he) chopped the wood.’

- 26) ◇ མེ་བསང་དེ། (Dz: མེ་བསང་ཡི།)
me sad-de
fire+ABS extinguish-PAST

‘(S/he) extinguished the fire.’

- 27) ◇ ལྷ་གང་ལྷ། (Dz: ལྷ་གང་ཡི།)
chu gang-te
Water+ABS full-PAST

‘(It) is full of water.’

28) འཕན་དེ། (Dz: འཕན་ཡི།)

phan-de

useful-PAST

‘(It) was useful.’

29) མལ་ཆ་བཀའ་བྱེ། (Dz: མལ་ཆ་བཀའ་བྱེ།)

malcha

kap-te

blanket+ABS cover-PAST

‘(It) is covered with the blanket.’

30) ཤིང་སྐམ་ལྗེ། (Dz: ཤིང་སྐམ་ཡི།)

shing

kam-te

wood ABS dry-PAST

‘The tree/wood dried.’

31) ང་སེམས་དགའ་བྱེ། (Dz: སེམས་དགའ་ཡི།)

nga

sem ga-te

1SG+ABS mind happy-PAST

‘I am happy.’ (lit. I have rejoiced)

32) ཉི་རྩ་ཐོང་དེ། (Dz: ཉི་རྩ་ཐོང་ཡི།)

tiru

bor-de

money+ABS lose-PAST

‘(S/he) lost the money.’

There are three types of sensory-inferential depending on the verb aspect. The form *V-khan-duk-te* indicates that the speaker had *sensory access* to a progressive activity taking place in the past (see examples below). Historically, this form is made of a nominalizer */khan/* (< *MKHAN*) and an auxiliary verb */duk/*, followed by the connective particle */te/*.

Chart 6: *Past auxiliaries*

	Neutral	Assumptive	Sensory-inferential	Epistemic
Simple past	<i>V-te</i>	<i>V-pi</i>		<i>V-piong</i>
Progressive past			<i>V-khandukte</i>	
Perfect			<i>V-temet</i>	
Present perfect continuous		<i>V-teyötpi</i>	<i>V-teyöt</i>	<i>V-teyötpiong</i>

33) ◇ མདང་ང་མོང་རྫར་ལེ་མོང་པི་ (Dz: ལ་ཙ་ང་མོང་རྫར་ལེ་མོང་པི་)

dang nga mongar-le song-pi/te
yesterday 1SG+ABS Mongar-DAT go-PAST+ASMP/ PAST

‘Yesterday I went to Mongar.’

34) ◇ ང་ལ་མེ་མཐོང་པི་ (Dz: མིན་བུ་མེ་མཐོང་པི་)

nga kha:yeme thong-pi/te
1sg+ABS firefly+ABS see-PAST+ASMP/ PAST

‘I have seen fireflies.’

35) ◇ མདང་ཁོང་ཆང་འབྲུང་མཁན་འདྲག་སྟེ་ (Dz: ལ་ཙ་ཁོང་ཆང་འབྲུང་དེས།)

dang khong chang tung-khandukte
yesterday 3PL chang drink- PROG+SENS+PAST

‘They were drinking chang yesterday (the speaker saw them drinking).’

36) ◇ ཆར་པ་བཏང་སྟེ་མེད་ (Dz: ཆར་པ་རྒྱབ་ནས།)

charpa tang-temet
rain LV-PERF+SENS

‘It has rained (past inferential based on the observation of the wet road).’

37) ◇ ཁོང་ཆང་འབྲུང་སྟེ་ཡོད། (Dz: ཁོང་ཆང་འབྲུང་སྟེ་ཚོད་ལྔ།)

khong chang thung-teyöt
3PL+ABS chang+ABS drink- PPCONT+SENS

‘They have been drinking chang (I saw them drinking and they are still continuing).’

Concerning the perfect, it is interesting to note that the structure V+connective+ negation is also found in some Tsang dialects, such as Phusum, Chuling and Nyemo, to indicate the inferential perfect (see Tournadre and Jiatso, 2001).

Compare the Tsamang CN and Tsang dialects mentioned above: ◇ ཆར་པ་བཏང་སྟེ་མེད་ /charpa tang-temet/ versus ཆར་པ་བཏང་ནི་མེ་འདྲག་ /charpa tang-nimindu/ ‘It has rained!’ (literally: It has rained and (now) there is no (rain)).

4.2.4 Auxiliary verbs of the future tense

The main auxiliaries of the future express the assumptive and epistemic statements. The form used for the future is /sang/, optionally followed by /yin/. It is interesting to note that /sang/ is also a nominalizer, as we will see in the next section. This is not a coincidence, since in most Tibetic languages verb tenses are often made of a nominalizer or a connective followed by an auxiliary verb.

The ending /ong/, derived from the CT verb ‘ong’, is also used for the future to express uncertainty.

Chart 7: Future auxiliaries

	Assumptive	epistemic
Future	V-sang (yin) V-mi ³⁰	V-sang-ong

38) འཇོང་མང་ལ་འགྲོ་བཤིག་ (Dz: འཇོང་མང་ལ་འགྲོ་བཤིག་)

nga mongar-le song-sang

1SG Mongar-DAT go-FUT

‘I will go to Mongar.’

39) མ་རྒྱུ་མཆེན་པ་བཟང་པ་འཇོང་བཤིག་ (Dz: མ་རྒྱུ་མཆེན་པ་འཇོང་བཤིག་)

nangpar charpa tang-sang-ong

Tomorrow rain LV-FUT-EPI

‘Tomorrow, it may rain.’

Other less grammaticalized means are available, such as the following sentence in which the verb *zon* means ‘about to’, ‘on the verge of.’ It can be used when looking at dark clouds.

40) མ་རྒྱུ་མཆེན་པ་བཟང་པ་འཇོང་བཤིག་ཡོད་ (Dz: མ་རྒྱུ་མཆེན་པ་འཇོང་བཤིག་ཅིག་འདུག་)

nangpar charpa tang-ma-zonkiyot

Tomorrow rain LV-NMLZ- ‘BE ABOUT TO’

‘It looks like it’s going to rain tomorrow.’

4.2.5 Nominalizers

The main nominalizers are /khan/, /pa/~wa/~ma/, /sa/, /thang/, ³¹ /sang/. They are all found in CT except /sang/. They correspond respectively to *MKHAN*, *PA~BA*, *SA* and *STANGS*.

Concerning the nominalizer /ma/, van Driem suggested that it might be a loan from Bumthang: “Cho-ca-nga-ca-kha has adopted the Bumthang infinitive ending *-mala*, e.g. *song-mala* ‘to go, will go’” (van Driem, 1992: 5). This does not seem accurate for several reasons. First, the ‘infinite’ /ma/ is only one of the allomorphs for this nominalizer – two other forms, /pa/ and /wa/, are encountered depending on the last phoneme of the verb (see examples below). These three allomorphs are reflexes of suffixes that are well attested in Classical Tibetan: *PA*, *BA* and *MA*. The two suffixes *PA* and *BA* function in Classical Tibetan as nominal suffixes, as well as nominalizers, and are used after verbs. The third form *MA* is only used as a nominal (and adjectival) suffix, but not as a verbal suffix. The form /ma/ used as a nominaliser in CN could just be a nasalised form of *PA* in nasal environments, as is the case in Dzongkha.

³⁰ In the case of a final /ng/, it is possible to replace /sang/ by /mi/ with the same meaning: *song-sang* or *song-mi*, *tang-sang*, *tang-mi*. More research is needed on this particular morpheme.

³¹ /thang/ is aspirated in CN just as in Dzongkha, unlike the CT form: *STANGS*.

Second, the form /la/ (or /mala/) is not used in the Tsamang dialect, as shown in the following example /song-ma ze-dol/ ‘[She] said that [she] would go.’ /to za-wa song-pi/ ‘(She) went to eat.’

In Lhuentse, the form /la/ is used, but it has been described by those we consulted as a form of /shepsa/ ཞེས་ (Honorifics). If this is the case, /la/ would be a cognate of ལགས་ LAGS. In the dialects where this form is used, /la/ may have kept a grammatical function of auxiliary that was inherited from Classical Tibetan, but further research is needed on the dialects of Lhuentse, Trashi Yangtse and Trashigang to confirm this hypothesis.

It is interesting to note that the nominalizers /ma/ and /sang/ are used to indicate future tenses, the same as their homologues /yag+(yin)/ (< CHAS+YIN) in Standard Tibetan, RGYU(+YIN) in Amdo and /ni+(ing)/ in Dzongkha.

Below are some examples of the CN nominalizers.

Marking of the Agent [A]

41) ཏི་རུ་བྱིན་མཁན་གཡི་ (Dz: ཏི་རུ་བྱིན་མི་གཞི་)
tiru bin-khan kayi
 money+ABS give-NMLZ who
 ‘Who gave [us] the money?’

42) ཀླད་བྱ་མཁན་གཡི་ (Dz: ཀླད་བྱ་བ་མི་གཞི་)
kat-bya-khan kayi
 sound do- NMLZ who
 ‘Who is shouting?’

43) ཐོ་ལབ་བརྟེན་མཁན་བུ་མོ་དང་ཐོ་ང་གིས་ཤེས་ (Dz: ཐོ་བོ་སྐབ་མི་བུ་མོ་དེ་ཚུ་ང་གིས་ཤེས་)
pholap tang-khan bumo (g)angpo nga-ki she
 talk LV- NMLZ girl PL+ABS 1SG-ERG know
 ‘I know the girls who are talking.’

Marking of the Patient [P]

44) ཁོ་གིས་འབྲས་མཁན་ཕྱག་པ་ལེགས་པོ་ཡོད། (Dz: ཁོ་གིས་ཐགས་མི་ཚེར་མ་འདི་ལེགས་ཤོམ་འདུག་)
kho-k thak-khan phrukpha lekpo yöt
 He-ERG weave- NMLZ basket+ABS nice COP+SENS
 ‘The basket he wove is nice!’

45) མོ་གིས་བཙུ་མཁན་ཆང་ངར་མ་ཡོད་པ། (Dz: མོ་གིས་བསྐྱལ་མི་ཆང་འདི་ངར་དྲགས་ཡོད་)
mo-ki cho-khan chang ngarma yöt-pi
 She-ERG prepare- NMLZ chang+ABS strong COP-ASMP
 ‘The chang she prepared is strong.’

Marking of the Manner

46) ལྟ་བུ་ཐངས་མུ་ཐོ་རོང་མ་མཐོང་། (Dz: ལྟ་འབད་ཐངས་འདྲི་བཟུམ་མ་མཐོང་)

la: *bya-thang* *utorong* *ma:-thong*
work+ABS do-NMLZ such way NEG-see

‘I have never seen anyone work this way.’

47) ག་རི་བཏང་ཐངས་འཛིགས་སིང་གི་ཡོང། (Dz: ག་རི་མུ་མ་འཁོར་བཏང་ཐངས་འཛིག་སི་སི་འདུག་)

gari *tang-thang* *jiksingki* *yöt*
car+ABS LV-NMLZ scary COP+ SENS

‘His way of driving is scary.’

Completive Clause Marker

48) སོང་མ་ཟེ་དྲ། (Dz: འགྱུ་ནི་ཟེར་སླབ་དེས་)

Song-ma *ze-do*
Go-NMLZ tell- PRES+ PSENS

‘[She] said that [she] would go.’

49) ཐུག་པ་འཐག་སང་འཇམ་རྟོང་གི་མེད། (Dz: རྩེར་ཕ་ཐགས་པ་ད་འཇམ་རྟོག་རྟོ་མེད་)

Phrukpa *thak-pa/thak-sang* *jamtongki* *met*
Basket+ABS weave-NMLZ easy COP+NEG+SENS

‘It is not easy to make baskets.’

50) ཆལ་ལྷ་བཏོག་པ་སོང་མི། (Dz: ཆལ་ལྷ་བཏོག་པ་འགྱུ་ནི།)

Tshalu *tok-pa* *song-mi*
Orange+ABS pick up-NMLZ go-FUT

‘She will go to pick oranges.’

51) རྟ་ལེན་མ་སོང་གྱེ། (Dz: རྟ་ལེན་པ་འགྱུ་ཡི།)

ta *len-ma* *song-te*
Horse+ABS fetch-NMLZ go-PAST

‘(He) went to get a horse.’

52) ལྟ་བུ་ཐངས་ལ། (Dz: ལྟ་བུ་ཐངས་ལྟོ་ཡི།)

[to *za-wa]* *song-pi*
Meal+ABS eat-NMLZ go- PAST+ASMP

‘(She) went to eat.’

Marking of the Instrument

53) ◇ འདྲི་མཁོ་བ་མེ་སང་གི་ཐེ་མེག་ལམ་གཏེ་ཡོད་པ། (Dz: མཁོ་བུ་ཐེ་མེ་མེག་ལམ་གཏེ་ཡོད་པ་མཁོ།)

odi gɔwa phe-sang-gi demilem ke yöt-pi
this door+ABS open-NMLZ-GEN key+ABS where exist

‘Where is the key to open this door?’

54) ◇ ཐུག་པ་འཐགས་སང་གི་པ་ཚར་ལེན་ས་སང་ (Dz: ཐུག་པ་འཐགས་ནི་གི་པ་ཚར་ལེན་པ་འགྱུ་ནི།)

Phrukpa tak-sang-gi patshar len-ma song-sang
basket+ABS weave- NMLZ-GEN cane+ABS take- NMLZ go- FUT

‘I am going to get the canes to make a basket.’

Marking of the Place

55) ◇ ལཱ་བྱ་ས་འོ་པེ་ཡོད་པ། (Dz: ལཱ་འབངས་ས་མ་པ་ཡོད་པ་མིན།)

la:bya-sa ophe yöt-pi
Work do- NMLZ+ABS over there exist-ASMP

‘The place where (they) work is over there.’

56) ◇ འདྲི་པེ་ཐགས་ས་ཡིན་ (Dz: མ་ན་རྒྱ་བྱ་འཐགས་ས་མིན།)

ode phe thak-sa yin
This flour+ABS grind- NMLZ COP-ASMP

‘This is the place where the flour is ground.’

57) ◇ རྩ་ཆུ་འཐུང་ས་འོ་པེ་ཡིན་ (Dz: རྩ་ཆུ་འཐུང་ས་མ་ན་མིན།)

nor chu thung-sa ode yin
Cattle+ABS water+ABS drink- NMLZ this COP-ASMP

‘Here is the place where the cattle drink water.’

4.3 Noun Phrase

4.3.1 Demonstratives

As in many southern and western Tibetic languages, the demonstrative precedes the noun. The main opposition is between proximal and distal /*odi*/ ‘this’ and /*ophi*/ ‘that’. When the demonstratives are used, the noun may be followed by the suffix /*sho*/.³² There is no definite article, unlike in Dzongkha where the /*ti*/ is used as a postponed article.

58) ◇ འདྲི་ཤིང་(ཤོ)ག་ཡི་བཙུགས་པ། (Dz: མཁོ་ཤིང་འདྲི་ག་གིས་བཙུགས་ཅི།)

odi shing (sho) kayi tsuk-pi
this tree+ABS-there who plant -ASMP

‘Who planted this tree?’

³² This seems similar to the French use of ‘là’ in “Ces enfants” vs “ces enfants-là.” According to G. Hyslop (p.c.), the form /*sho*/ with a similar function is also found in Kurtoep.

- 59) འཕྱི་མི་དཀར་སུ་གོ་སང་གི་ (Dz: མ་ཕི་མི་དཀར་སུ་གོ་སང་གི་)
Ophi mi karpu-sho ke song-pi
 That man white+ABS-there where go-ASMP
 ‘Where did the white man go?’

Chocha-ngachakha has a topicalizer /ne/ derived from the CT word *NI*, used in Dzongkha and Central Tibetan.

- 60) ལྗང་ཕྱི་ལྷགས་པོ་ཡོད། (Dz: ལྗང་ཕྱི་ལྷགས་པོ་མཉམ་འདུག།)
to-ne lekpo yöt
 Cooked rice+ABS-TOP good COP+SENS
 ‘As for the rice, it is good.’

4.3.2 Personal pronouns

The personal pronouns are: ང་ /*nga*/ ‘I’, ཁྱེད་ /*chet*/ ‘you’, ཁྱེ་ /*kho*/ ‘he’, མོ་ /*mo*/ ‘she’, ང་ཅ་ /*ngaca*/ ‘we’, ཁྱེད་ཅ་ /*chetca*/ ‘you (pl)’, ཁྱེད་ཅ་ /*khong*/ ‘they’. There is also an honorific form for ‘you’: དེའ་ /*de*/. The following dual forms are attested: ང་ཅ་པོ་ /*ngacapol*/ ‘we two’, ཁྱེད་ཅ་པོ་ /*chetcapol*/ ‘you two’, ཁྱེད་ཅ་པོ་ /*khongpol*/ ‘they two’.

4.3.3 Interrogative pronouns

The set of interrogative pronouns is clearly derived from CT. They include: གཞི་ལྡོ་ /*kayil*/ ‘who’, གཞི་ལྡོ་ /*ke*/ ‘where’ or གཞི་ལྡོ་ /*kale*/ (old form), གཞི་ལྡོ་ /*katsat*/ ‘how many’, གཞི་ལྡོ་ /*katat*/ ‘how’, གཞི་ལྡོ་ /*keki*/ or གཞི་ལྡོ་ /*kek*/ ‘from where’. The first syllable of these words is derived from CT *GANG* ‘what, who’. While in many Tibetic languages these pronouns are realized with low tones (གཞི་ལྡོ་ /*kacil*/ ‘what’, གཞི་ལྡོ་ /*ka*/ ‘who’ in Dzongkha; གཞི་ལྡོ་ /*khare*/ in Standard Tibetan; གཞི་ལྡོ་ /*ga*/ ‘who’ in Thewo Tibetan), in CN these yield a high tone. The other pronouns include the pronouns ཅི་ /*ci*/ ‘what’, ཅི་ལྡོ་ /*ci byat*/ ‘why’,³³ and གཞི་ལྡོ་ /*nam*/ ‘when’. The forms /*ci*/ and /*nam*/ are derived respectively from CT *CI* and *NAM* and are attested in many modern languages.

4.3.4 Cases

For the sake of comparison, we will first briefly present the Dzongkha cases. We find the absolutive: Ø; the genitive: གཞི་ *GI*; the ergative: གཞིས་ *GIS* /*gi*/; the dative: ལུ་ /*lu*/ (derived from CT ལ་ *LA*; the ablative ལས་ *LAS*; the associative དང་ *DANG*; the locative ན་ *NA* /*na*/; and the comparative ལ་ /*wa*/ (derived from CT ལས་ *BAS*). These seven overt cases are derived from their Classical correspondents. Concerning the locative case ན་ /*na*/, it is necessary to distinguish it from the postposition ན་ /*na*:/, which has a long vowel. This difference has been largely unnoticed. For example, van Driem (1992) noted: “The dative suffix also indicates location, like English ‘to’ or ‘at’, and may even be

³³ *Cibya:te* is a compound form made of /*ci*/ “what” and /*bya*:/ “to do” and the connective /*te*/. Its construction is analogous with standard Tibetan: *GA.RE BYAS-NAS*.

combined with the locative suffix to indicate the site of activity or situation, e.g. [...] ཡིག་ཆ་ནང་ལུ་ /*yitsha-na-lu*/ ‘at the office’.”

Our interpretation, however, is that the above example corresponds to the combination of a postposition /na:/ followed by the dative /lu/ : ནང་ལུ་ /*na:-lu*/. This hypothesis is confirmed by the fact that the postposition ནང་ /na:/ can be followed not only by the dative ལུ་ /lu/ but also by the locative case /na/, or even the ablative /lä/: ནང་ན་ /na:-na/ ‘in’ and ནང་ལས་ /na:-lä/ ‘among’.

Moreover, the forms /na:-na/ and /na:-lu/ convey slightly different meanings, respectively, the illative and the allative:

61) Ex. ང་ཁྱིམ་ནང་ན་འགྱོན་ཞིན་
nga chim na:-na jo-niing
 1SG house in-LOC go-FUT+ASMP

‘I will go into the house.’

62) ང་ཁྱིམ་ནང་(ལུ)་འགྱོན་ཞིན་
nga chim na:(lu) jo-niing
 1SG house in-DAT go-FUT+ASMP

‘I will go to the house.’

The postposition /na:/ ‘in’ is increasingly used alone without being followed by case and thus, in the process of grammaticalization, it might finally become a case marker. However, we still describe it as a postposition because of its optional syntactic combination with the case markers /lu/ and /na/. In Dzongkha, two other postpositions, ཁ་ *kha* ‘on’ and འགྲུ་ *gu* ‘on’, are also on the verge of becoming cases. They are not found in Classical Tibetan and are derived from relator nouns *KHA* ‘mouth’, ‘surface’, and *GU* < *GO* < *MGO* ‘head’. They are also optionally followed by case markers such as /lu/ or /r/ in formal Dzongkha but are more often used directly after the noun.

If we consider ཁ་ /*kha*/, ནང་ /*na*/ and འགྲུ་ /*gu*/ as postpositions, Dzongkha has seven cases.

As we will see, the CN case system presents many functional similarities with Dzongkha, but it also exhibits formal discrepancies. CN has six overt cases: absolutive (Ø), ergative /gi~k(i)/, genitive /gi~yi/, dative /le/ (and allomorphs), associative /dang/, ablative /leki/, comparative /wata/. It is easy to see that the first four overt cases correspond in general to the reflexes of their CT correspondents: གིས་/ གྱིས་ *GIS~KYIS* (ergative), གི་/ གི་ *GI~YI* (genitive), ལ་ *LA* (dative), and ངང་ *DANG* (associative). The ablative and the comparative look more problematic for the reconstruction, but we will come back to this issue later.

In some Tibetic languages of southern Kham and to a certain extent in Dzongkha, the ergative marking the grammatical Agent is often optional and essentially plays a pragmatic role. This is also the case in CN: In many examples of this paper, the ergative is simply dropped.

The ergative

The ergative is marked by various allomorphs: /gi/, /ki/ or the short form /-k/. After consonants, the form is /gi/ but after vowels, the form /-k(i)/ is used.

63) ཁོ་གི་ཁོ་གིས་སློབ་བཏང་གྱི། (Dz: ཁོ་གིས་སློབ་བཏང་གྱི།)

kho-k/kho-ki sun tang-te
3SG-ERG story+ABS LV-PAST

‘He told a story.’

64) ཚེ་རིང་གིས་པ་ཚར་འཐེན་གྱི། (Dz: ཚེ་རིང་གིས་པ་ཚར་འཐེན་གྱི།)

tshering-gi patshar then-te
Tsering-ERG cane pull-PAST

‘Tshering collected some canes (lit. pulled the canes).’

65) རོ་གྱིས་ཨ་ཁྱ་ཟེ་གྱི། (Dz: རོ་གྱིས་ཨ་ཁྱ་ཟེ་གྱི།)

nor-gi aham za:-te
cow-ERG maize+ABS eat-PAST

‘The cow has eaten the maize.’

66) སྐྱ་གྱིས་ཉ་ཐ་དི། (Dz: སྐྱ་གྱིས་ཉ་ཐ་ཨིན།)

sam-gi nya za-di
otter-ERG fish+ABS eat-PRES

‘Otters eat fish.’

The genitive

The genitive is identical to the ergative /gi/ after consonants and differs only when it follows a vowel, as shown in the chart below:

Chart 8: allomorphs of the genitive and the ergative cases

	Genitive	Ergative
After consonants	/gi/	/gi/
After vowels	/yi/ or /gi/	/k(i)/

67) འ་འདི་ང་གི་བྱ་ཡིན། (Dz: འ་འདི་ང་གི་བྱ་ཡིན།)

odi nga-yi bu yin
this 1SG-GEN son+ABS COP-ASMP

‘This is my son.’

The dative

The dative is formed with the suffix /le/, which has several allomorphs: /le/, /ge/, /nge/, /e/.³⁴ The form /le/ is found in Kham Tibetan for the dative,³⁵ and it is cognate with the CT *LA* form as well as the Dzongkha form /lu/, Lhoke /lo/ and has the same meaning. The CN /le/ (and its allomorphs) is used to indicate the function of dative as well as locative and allative.

Chart 9: The allomorphs of the dative marker

Phonological context	Dative
After /t,n,p,r/	/le/
After /k/	/ge/
After /ng/	/nge/
After /m/ vowel	/e/

68) རྩ་ལ་རྩ་བྱིན་ལྟེ་ (Dz: རྩ་ལ་རྩ་བྱིན་ལྟེ་)
nor-le tsowa bin-te
 Cow-DAT grass+ ABS give-PAST

‘(S/he) gave the cow some grass.’

69) མི་བརྒྱད་ལ་དྲི་ལྟེ་ (Dz: མི་བརྒྱད་ལ་དྲི་ལྟེ་)
Mi jat-le dri:-te
 Person eight-DAT ask(P)-PAST

‘He asked eight people.’

70) རྩ་མ་བྱིན་ལྟེ་ (Dz: རྩ་མ་བྱིན་ལྟེ་)
kho-e bin-de
 3SG-DAT offer-PAST

‘(S/he) gave (it) to him’

71) རྩ་མ་ཚེ་རིང་ལ་བྱིན་ལྟེ་ (Dz: རྩ་མ་ཚེ་རིང་ལ་བྱིན་ལྟེ་)
Kho-k tshering-nge trot-de
 3SG-ERG Tsering-DAT give-PAST

‘He gave (it) to Tshering.’

³⁴ It triggers a pharantalization of the preceding vowel.

³⁵ Despite the fact that it could theoretically be derived from the ablative *LAS*, it is much more plausible that it comes from the dative, since most Kham and Southern Tibetic languages such as Dzongkha and Lhoke exhibit a great variation in the vowel quality: /la/, /le/, /lo/, /lu/, etc.

72) ཁམས་མེ་སོང་ནེ་ (Dz: ཁམས་ལུ་གཏོ་ཡི།)

kham-e song-te

Kham-DAT go-PAST

‘He went to Kham.’

73) ཁོ་མེ་བུ་སྐུ་གསུམ་ཡོད་ཅི། (Dz: ཁོ་ལུ་མེ་བུ་གསུམ་ཡོད་ཅི་ཅེ།)

kho-e busa sum yotpi

3SG-DAT child three+ABS exist-ASMP

‘He has three children.’

74) རྣམས་ཚལ་ནང་ལུ་ (Dz: རྣམས་ཚལ་ནང་ལུ་)

naktse: nang-nge

forest in-DAT

‘In the forest’

It is possible to have a form /tate/ postponed to the dative. This conveys the meaning of the preposition ‘for’ or ‘to’ in English.

75) ཁོ་མེ་ཏཱ་བྱིན་དེ་ (Dz: ཁོ་ལུ་བྱིན་ཡི།)

kho-e-tate bin-de

3SG-DAT-for offer-PAST

‘(S/he) gave it to him.’

The associative

The associative is marked by /dang/. It is often optional and has marginal status in the case system.

The ablative

As we suggested earlier, the ablative and the comparative constitute the most exotic cases of CN. At first, the long form of the ablative /leki/ does not seem to reflect any form of CT. It is not directly related to the Classical ablative *LAS* nor its reflex /lä/ in Dzongkha because nothing could explain the /ki/ sequence. When we examine this more closely, it becomes obvious that the ablative is a compound case made of the dative and the ergative. This hypothesis is confirmed by the fact that /leki/ has several allomorphs that behave in exactly the same way as the ablative /le/ and the ergative /ki/, as appears from the chart 10 below. (Compare with the chart of the dative.)

Chart 10: The allomorphs of the ablative

Phonological context	Ablative
After /t,n,p,r/	/le-ki/ or /lek/
After /k/	/ge-ki/ or /gek/
After /ng/	/nge-ki/ or /ngek/
After /m/ vowel	/e-ki/ or /ek/

Examples:

76) ལྷ་གར་ལེ་གི་ལོག་འོང་གི་ (Dz: ལྷ་གར་ལས་འོང་གི།)

gyagar-leki ong-pi
India-ABL come-PAST

‘(We) came from India.’

77) ལུང་ཐང་དེ་གི་(འོག་ལོང་གི་) (Dz: ལུང་ཐང་ལས་འོང་གི།)

bumthang-ngek ong-pi
Bumthang-ABL come-PAST

‘(She) came from Bumthang.’

78) ལམ་ཁ་མེ་གི་(འོག་ལོང་གི་) (Dz: ལམ་ཁ་ལས་འོང་གི།)

kham-ek ong-pi
Kham-ABL come-PAST

‘(I) came from Kham.’

The fact that the ablative is a compound case is not very surprising. The CT forms of the ablative *LAS* and the elative *NAS*, as suggested by some scholars, are also compound forms, respectively made of the dative+ergative (*LA+S*), and the locative+ergative (*NA+S*). If our hypothesis is correct, this would mean that CN has preserved an archaic form **LA-KYIS* that is not attested in Classical Tibetan but built on the same model as *LA-S*.

The comparative

The comparative case is also interesting from the point of view of historical linguistics. It appears clearly that the form /wata/ is cognate with the Dzongkha form /wa/, which is itself derived from Old Tibetan *BAS* (see Tournadre 2010). Thus the /wata/ form must also be reconstructed as a compound form made of /wa+/ta/. One hypothesis could be that the form /ta/ could come from the verb *LTA* ‘to watch’, ‘to consider’, which is also pronounced /ta/ in Chocha-ngachakha. This hypothesis could be supported by the fact that *LTA* has been grammaticalized in other Tibetic languages, notably some Amdo dialects, to form the comparative construction. Of course other hypotheses should be examined, such as the borrowing of the form from neighbouring languages such as ‘East Bodish.’³⁶

Here are two examples of the comparative construction:

79) ང་མ་ཏ་ལྷོང་ནས་ (Dz: ང་བས་ལྷོང་ནས།)

nga-wata chet ge
1SG-COMP 2SG+ABS old

‘You are older than me.’

³⁶ The marker /wa/ is also found in Kurtoep (Hyslop, p.c.).

80) གྲོང་གསར་ལ་ཉ་མོང་སྤར་དྲོ་དོ་ (Dz: གྲོང་གསར་བས་མོང་སྤར་དྲོ་མ་ལ།)
trongsar-wata mongar dro-do
 Trongsar-COMP Mongar+ABS warm-PRES
 ‘Mongar is warmer than Trongsar.’

5 Lexical differences between Chocha-ngachakha and Dzongkha

The CN vocabulary is nearly entirely derived from Tibetic forms related to Old Tibetan. When comparing Chocha-ngachakha and Dzongkha, we also find many direct correspondences in the lexicon of the two languages.

Chart 11: *Similarities between Chocha-ngachakha and Dzongkha*

Chocha-ngachakha	Dzongkha	Meaning
ལ་བྱའ་ /la: bya:/ (< CT ལས་བྱས་)	ལ་འབད་ /la: be/ ³⁷	‘to work’
བྱ་ཆེ་གི་ /bya chiki/	བྱ་ཆེ་ཆེ་ /pcha chichi/	‘nice’
ལེགས་ལུ་ /lekpu/	ལེགས་ཤོམ་ /leshom/	‘good’
བྱིན་ /bin/	བྱིན་ / ^p ch _{in} /	‘to give’
འ་ཕི་ /ophi/	ཨ་ཕི་ /aphi/	‘that’
སྤོ་བ་བྲེས་ /towa tre:/	སྤོ་མ་བྲེས་ /tou ke:/	‘to be hungry’

In nearly all the above examples, CN has a more conservative form than the Dzongkha equivalents. However, we also find a number of differences even for very common words. The following chart provides a list of examples.

³⁷ The CN form suggests that the Dzongkha form /be/ could also derived from the CT verb ‘to do’: *BYED* (pres) or *BYAS* (past).

Chart 12: Differences between Chocha-ngachakha and Dzongkha

Chocha-ngachakha	Dzongkha	Meaning
ཞོ / <i>sho</i> / ³⁸ < <i>zho</i> ‘yoghurt’	མཚ་ / <i>om</i> / < ‘ <i>o.ma</i> ’ ‘milk’	‘milk’
ཞིམ་ལ་ / <i>shimola</i> / < <i>zhim.bu</i> ‘cat’	བྱིལ་ / <i>pcili</i> / < <i>byi.la</i> ‘cat’	‘cat’
འཇིགས་ / <i>jik</i> / < ‘ <i>jigs</i> ’ ‘to fear’	འདྲོག་ / <i>dro</i> / < ‘ <i>drog</i> ’ ‘to fear’ (for horses)	‘fear’
བརྩི་ / <i>bardzi</i> /	ནོར་ཇིང་ / <i>nodzip</i> /	‘herdsman’
མོག་པ་ / <i>mokpa</i> < <i>rmog</i> ‘helmet’	མཚ་ / <i>sham</i> / < <i>zhwa.mo</i> ‘hat’	‘hat’
གཤེན་པ་ / <i>sherpa</i> /	ཆུ་ / <i>chu</i> /	‘urine’
ཆུག་པ་ / <i>cakpa</i> /	ཨ་ལ་ / <i>awa</i> /	‘feces, shit’
ཆུ་ / <i>chachu</i> /	ཆུ་ཆོག་པ་ / <i>cagop</i> /	‘garlic’
བྱེད་ཆུང་ / <i>pretmung</i> /	མཚུང་མོ་ / <i>dzum</i> /	‘finger’
ཕོ་ལ་ / <i>pholap</i> /	ལོ་ / <i>lo</i> /	‘conversation’
བུ་ / <i>busa</i> / < CT < <i>bu+tsha</i>	ཨ་ལུ་ / <i>alu</i> /	‘child’
ཆར་པ་བྱང་ / <i>charpa tang</i> /	ཆར་པ་ཆུབ་ / <i>cha:p chap</i> /	‘to rain’
བོར་ / <i>bor</i> /	བྱང་ / <i>pcang</i> /	‘to lose’

Concerning the lexicon, the word /*sho*/ for ‘milk’ is particularly striking. The word **gyu*³⁹ found in neighbouring East Bodish and (ཏཱ)-lu⁴⁰ in Rgyalrong, both meaning ‘milk’,⁴¹ are obviously cognates with /*sho*/.

Thus /*sho*/ ‘milk’ in CN might well correspond to the Proto-Tibetan meaning, since the word in CT had already undergone a semantic change: The word now means ‘yoghurt’ in the modern Tibetic languages, and the word ‘*o.ma*, originally related to ‘breast’, has replaced it for the meaning of ‘milk’. As noted independently by Jacques (2014: 29-30), the meaning ‘curd’ for *zho* is an innovation in the modern Tibetic languages: “Le tibétain a [...] développé un verbe dénominal *ndzo*, *bzos* “traire” tiré du nom *zo* < **ljo* dont le sens original est certainement lait.” The CN word /*sho*/ is thus directly related to the CT verb ‘to milk’ (pres: *jo*, past *bzhos*, fut *bzho*, imp: *jos*) and has preserved the original meaning. If this explanation is correct, to our knowledge CN would be the only language out of the 50 or so Tibetic languages, with the word *zho* retaining the original meaning ‘milk’.

These lexical discrepancies may suggest that CN could have evolved separately from Dzongkha from a distinct branch of the Tibetic family at an early stage.

CN does not have a great proportion of loanwords. The main sources for borrowing are English, Hindi and Dzongkha. Here are some frequent examples:

Hindi: /*pura*/ ‘all’, /*gari*/ ‘car’.

English: /*phon*/ ‘phone’, /*miskoll*/ ‘miscall’, /*sikru drayber*/ ‘screwdriver’, etc.

³⁸ *Zho* designates milk as well as yoghurt; in order to distinguish them one uses an adjective: /*sho trengma*/ ‘fresh milk’, vs. /*zho ripa*/ ‘yoghurt’ (lit. rotten milk).

³⁹ a reconstruction proposed for East Bodish by Hyslop (p.c.)

⁴⁰ About sound changes **lj* > *zh*, see e.g. Hill (2013).

⁴¹ Jacques (2012: 214)

Dzongkha: /*kapol*/ < *rgadpo* /*gap*/ ‘headman’, /*jon*/ ‘to come, go’ (H). The normal reflexes of these two words in CN should be respectively /*gatpol*/ ‘elder’, /*byon*/ ‘to come, go’.

Likely there are also loanwords from neighbouring languages, such as Tshangla or Bumthangkha, but none is attested in our corpus.

6 Conclusions

The Tsamang dialect of CN exhibits many conservative pronunciations that are rarely attested in modern Tibetic languages. Among the most archaic features, we can mention the preservation of the Old Tibetan voiced initial consonants, as well as the preservation of combinations of an initial labial followed by a vibrant or a glide (P/B+R or P/B+Y). A reflex of the wasur *W* is attested in CN, while it is extremely rare in the Tibetic languages. Another unique feature is the existence of the sequence /*mr*/ in some words. This could well be a reflex of Old Tibetan *MR*. If this is the case, it would indicate that CN has preserved some of the most ancient phonological features of the Tibetic languages.

These archaic features have essentially been preserved in the Balti and Purik languages spoken in the northwestern Tibetic areas by Muslim populations of Ladakh and Baltistan. In these regions, the linguistic conservatism might be due to the peripheral location, as well as to the Shiah Muslim religion, both of which have contributed to the relative isolation of the language from their Ladakhi neighbours’. In the case of CN, the linguistic conservatism cannot be explained by a religious factor, since CN people mainly practice a form of Nyingma and Kagyü Buddhism dominant in Bhutan. It may again partly be explained by the peripheral geographic location of this language in eastern Bhutan and the fact that CN people live in a rather ‘non-Tibetic’ jungle environment at a low elevation and are located at the southernmost region of the Tibetic speaking area. Moreover, CN people are isolated from the other Tibetic speaking groups of Bhutan (Dzongkha in the west and Merak-Sakteng in the east).

We have also shown that the linguistic conservatism of Tsamang CN is not confined to phonology but extends to grammar and vocabulary. All these archaic characteristics reflect Old Tibetan forms and structures, so there is no reason to believe that CN has a Bumthangish or ‘East Bodish’ substrate, and we have not thus far found evidence reflecting the phonology or grammar of those languages.

In the field of grammar, Tsamang CN exhibits some very original features, such as the use of the ancient verb *YOD* to convey a sensory-inferential meaning.

The case system also has original features, some of which might be derived from the Old Tibetan case system. For example, the ablative in Tsamang CN appears to be a compound case made of the dative and the ergative, reminiscent of the Old Tibetan elative and ablative cases, which can be analysed as a combination of dative and ergative. In the field of lexicon, some words could reflect a very ancient meaning. For example, the word /*sho*/ for ‘milk’, which in virtually all other Tibetic languages means ‘yoghurt’, probably reflects the original Proto-Tibetic meaning. If it is the case, CN is so far the only Tibetic where this meaning has been preserved.

Therefore, together with Balti, Purik, Ladakhi, Amdo, and some archaic Kham dialects, CN belongs to the small group of languages with the most archaic features of the Tibetic area.

More data is needed to describe the Tsamang dialect, and of course other dialects of CN which have not been documented. It is particularly interesting to note that a variation from labial initial sounds to fricative sounds is attested between the closely related dialects of CN.

The significant differences between CN and Dzongkha suggest that CN is not a more conservative form of Dzongkha, but could have evolved separately from a form of Old Tibetan. The migration patterns of the CN people are of course a major question which has yet to receive an answer⁴². Did the group that now speaks CN follow the same migration patterns as those who have settled in Hâ and the Ngalong area? Do they originally come from Lhobrak? Have they crossed the Himalaya following the valley of the Kuri river? Are the CN people descendants of earlier migrations related to the Tibetan empire expansion when it reached the Gulf of Bengal?

Many other questions are waiting for answers, but it is already clear that the data from Chocha-ngachakha sheds new light on the evolution of the Tibetic family and that this language of eastern Bhutan has preserved the most archaic features among Tibetic languages spoken in the southern Himalaya.

ACKNOWLEDGMENTS

First, I want to thank the Institute of Language and Culture Studies (ILCS in Taktse, Trongsa, a college of the Royal University of Bhutan) and Lopen Lungtaen Gyatso, ILCS Director, for organizing my visit to Bhutan and for giving me opportunities to hold workshops and talks at ILCS and in Thimphu. I would also like to express my gratitude to the Lacito (CNRS) for once again supporting my research and providing financial help.

I would also like to thank Karma Rigzin, co-author of this article, who is a lecturer at ILCS. Karma Rigzin is a native of CN, born in Tokari hamlet in Tsamang gewog. Having received a degree of Acharya from the Central University of Tibetan Studies in Sarnath (Varanasi, India), Karma also has a good knowledge of Sanskrit.

The collaboration of two linguists, among whom one is native, allowed maximal efficiency in the analysis of the CN language. Despite the isolation of his village in the dense forests of Tsamang and the rough environment, we were able to carry out our fieldwork in excellent conditions, thanks to Karma's family members: father Dorji, brother-in-law Thinley, and sister Sithar. May they all receive our warm gratitude for their hospitality and help with our fieldwork.

Many thanks to Gwendolyn Hyslop for her comments on this article and for sharing her views of the relationship between "East Bodish" and Chocha-ngachakha.

Finally we would like to warmly thank Jude Polsky, ILCS English lecturer, for correcting and editing this paper.

ABBREVIATIONS

ABL	ablative	FUT	future
ABS	absolutive	IMP	imperative
ASMP	assumptive	INFR	inferential
COMP	comparative	NMLZ	nominalizer

⁴² Genetic studies, and particularly the tracing of the Y chromosome haplogroup, might provide helpful information.

DAT	dative	OT	Old Tibetan
ERG	ergative	PPCONT	present perfect continuous
TB	Tibeto-Burman	SENS	sensory
COP	copula	PERF	perfect
CN	Chocha-ngachakha	PSENS	participatory-sensory
CT	Classical Tibetan	TOP	topic
E/E	Evidential/ Epistemic	Q	question marker
EPI	Epistemic	WARN	warning particle

CONVENTIONS

The Tibetan script is used to note Classical Tibetan, Dzongkha and Chocha-ngachakha. The sign ◊ indicates Chocha-Ngacha sentences or words (whenever the spelling does not match the classical Tibetan orthography or when the words in the two languages have different meanings).

For the transcription of CN in Tibetan script, we use an additional diacritic sign (called tsalak), also used in Lhoke (Sikkimese) with labials followed by a glide ʂ /by/. The diacritic sign is not used (ʂ) when the sound is realized as an affricate /j/ (see the chart for correspondances with IPA).

The reason to note CN with Tibetan script is that it allows a better preservation of the language and also gives an easier access of information to native speakers. The co-author of this article is a native speaker of CN and also a master of Classical Tibetan. Most CN examples have been translated in Dzongkha, the national language of Bhutan. These translations are preceded by the abbreviation “Dz” and they are intended for the Bhutanese readers for the sake of comparison between Dzongkha and CN.

When used for CT, the Tibetan script is always followed by the Wylie transliteration in small capital and italics which renders the orthography. In the case of Dzongkha and CN, the script is followed by a Romanization which renders the pronunciation.

REFERENCES

- Bradley, David. 1997. “Tibeto-Burman languages and classification”. In: David Bradley (ed.), *Papers in Southeast Asian linguistics* No. 14: *Tibeto-Burman languages of the Himalayas*, 1-72. Canberra: Pacific Linguistics.
- Bradley, David. 2013. “The ancestry of Tibetan”. In: Gray Tuttle, Kunsang Gya, Karma Dare and Johnathan Wilber (eds.), *The Third International Conference on Tibetan Language, Volume 1: Proceedings of the Panels on Domains of Use and Linguistic Interactions*, 363-397. New York: Trace Foundation
- van Driem, George, 1992. *The grammar of Dzongkha*. Thimphu: RGoB, Dzongkha Development Commission (DDC).
- van Driem, George; and Tshering of Gaselô, Karma. 1998. *Dzongkha: Languages of the Greater Himalayan region*. Leiden: CNWS Publications.
- van Driem, George. 2007. “Dzala and Dakpa form a coherent subgroup within East Bodish, and some related thoughts”. In: Roland Bielmeier and Felix Haller (eds). *Linguistics of the Himalayas*

- and beyond*, 71–85. De Gruyter Mouton. [Series Trends in Linguistics, Studies and Monographs, 196]
- Hill, Nathan, W. 2006. “Tibetan vwa ‘fox’ and the sound change Tibeto-Burman *wa -> Old Tibetan o”. *Linguistics of the Tibeto-Burman Area* 29.2: 75-90.
- Hill, Nathan W. 2010. “An overview of Old Tibetan synchronic phonology”. *Transactions of the Philological Society* 108.2: 110-125.
- Hill, Nathan, W. 2013. “Relative order of Tibetan sound changes affecting laterals”. *Language and linguistics* 14.1: 193-209.
- Hyslop, Gwendolyn. 2013, “On the internal phylogeny of East Bodish”. In: Gwendolyn Hyslop, Stephen Morey and Mark W Post (eds.), *North East Indian Linguistics*, Volume 5, 91-112. New Delhi: Cambridge University Press India.
- Jacques, Guillaume, 2012. “An internal reconstruction of Tibetan stem alternations”. *Transactions of the Philological Society* 110.2: 212–224.
- Jacques, Guillaume, 2014. *Esquisse de phonologie et de morphologie historique du tangoute*. Leiden: Brill.
- Jackson, Sun T. S (ed). 2014. *Phonological Profiles of Little-Studied Tibetic Varieties*. Taipei: Academia Sinica. [Language and Linguistics Monograph Series]
- Shafer, Robert 1955. “Classification of the Sino-Tibetan languages.” *Word* 11: 94-111.
- Tournadre, Nicolas, 2014. The Tibetic languages and their classification. In: Nathan W. Hill and Thomas Owen-Smith (eds.), *Trans-Himalayan linguistics: Historical and descriptive linguistics of the Himalayan area*, 105-129. Berlin: De Gruyter Mouton.
- Tournadre, Nicolas; and LaPolla, Randy J. 2014. “Towards a new approach to evidentiality”. *Linguistics of the Tibeto-Burman Area* 37.2, 240–263.
- Tournadre, Nicolas; and Konchok Jiatso. 2001. “Final auxiliary verbs in literary Tibetan and in the dialects”. *Linguistics of the Tibeto-Burman Area* 24.1: 49-110.

Nicolas Tournadre
nicolas.tournadre@univ-amu.fr

Karma Rigzin
karma.ilcs@rub.edu.bt

APPENDIX: Lullaby of ‘Acho Lala’ (lit.: ‘Elder Brother Moon’)

The following text is a lullaby that is well known in the Tsumang and Tsakaling areas. We propose here a transcription in Tibetan script of the Chocha-ngachakha. We have kept the classical orthography whenever it matched the Tsumang pronunciation.

ཨ་ཅོ་ལ་ལ་རྩོ་རྩོ་རོ་རྩོ་དྲུ་པོ་ཅུང་བར་རྒྱུང་ག་ལེ་རོག་གེ་བཞག་ལྟེ་རོག་ག་ལེ། མེ་གིས་ཚོག་ལྟེ་མེ་ག་ལེ། རྒྱ་གིས་བསང་དེ། རྒྱ་ག་ལེ།
 རྒྱང་གིས་འབྱུང་ལྟེ། རྒྱང་ག་ལེ། རྒྱག་གིས་བསང་དེ། རྒྱག་ག་ལེ། རྒྱང་གིས་བསང་དེ། རྒྱང་ག་ལེ། རྒྱིན་གྱིས་བསང་དེ། རྒྱིན་ག་ལེ། སང་དགན་མཁྱི་བར་ལེ་མཛུག་མ་གིར། ལྷ་མ་ལ་རྒྱ་འབྱུང་ཏིབ་
 གེ། སྤང་ལ་རྒྱ་བྱང་དེ།

Acho la:la toto ro:ro chö-gi phurung bangchung kale rog-e shagte rok kale me-ki tshik-te, me kale chu-ki sad-de, chu kale, lang-gi thung-te, lang kale, tak-gi sad-de, tak kale, gung-gi sad-de, gung kale, trin-gi sad-de, trin kale, sa dang nam-gi bar-le jugma kir, umla chu thung tipki pangla tsa za pratki.

Translation:

“Elder Brother Moon, it is time for food. Where is your plate and cup? It’s on the shelf. Where is the shelf? It has been burned by the fire. Where is the fire? It’s been extinguished (killed) by the water. Where is the water? The ox has drunk it. Where is the ox? The tiger has killed it. Where is the tiger? The leopard has killed it. Where is the leopard? The cloud has killed it. Where is the cloud? Between the sky and the earth it disappeared (lit.: wagging the tail). Water in the small pool has been completely drunk. All the grass in the pasture has been eaten.”