

Proto-Tibetic *mbras ‘¹grain; ²rice’

Comparative reconstruction and dialect subgrouping

Joanna Bialek
Trinity College Dublin

This article proposes a new historical subgrouping of Old Tibetan dialects. It demonstrates that Tibetan dialects started to differentiate in late pre-historical times, shortly before the script introduction. The emergence of the first dialects is shown to have resulted from the military expansion of the Tibetan Empire. The reconstruction of the dialect formation supplements earlier insights by studying modern equivalents of Old Literary Tibetan *ybras*, the Proto-Tibetic form of which has been reconstructed as *mbras. In addition, comparing its cognates in other presumably closely related languages permitted the reconstruction of *mras to Proto-Bodic and, tentatively, *mrats to Proto-Trans-Himalayan.

Keywords: Tibetic, Proto-Bodic, Tibetic dialectology, comparative reconstruction, historical dialect subgrouping, rice

1. Introduction

Following extensive military conquests between the early 7th and mid-9th century, speakers of Old Tibetan not only took possession of the whole Tibetan Plateau but also established permanent and lasting settlements as far as Skardu in the west, Turfan in the north, Min Mountains in the east, and Sathang (Chinese Lijiang) in the south. This will occur even more impressive if we realise that around the mid-6th century, Pre-Tibetic was spoken only in the southern part of the Tibetan Plateau along the Gcañ-po (Brahmaputra) river, roughly between the Śaṅs valley in the west and the Rkoñ region in the east (Map 1). The subsequent centrifugal dispersal of Tibetan-speaking groups in all directions has led to the differentiation of the languages they spoke and the proliferation of Tibetan vernaculars. The modern distribution of Tibetic languages (see Map 3 in Tournadre & Suzuki 2023) reflects these early historical events.

The objective of this article is twofold; firstly, it attempts to refine the sub-grouping of Tibetan dialects presented in Bialek (2018b) by comparing modern equivalents of Old Literary Tibetan (OLT) *ybras* ‘¹grain; ²rice’ and secondly, based on modern reflexes, it reconstructs its Proto-Tibetic form. The present contribution understands the earliest sub-division of Tibetic languages as an immediate and unavoidable outcome of the military expansion and the ensuing colonisation of the conquered regions. I try to reconstruct the earliest stages in the formation of dialects from Proto-Tibetic to roughly the end of the imperial period (ca. mid-9th century).

In its central part, the article presents a case study of Literary Tibetan *ybras* ‘¹grain; ²rice’. After establishing the original meaning of *ybras* as attested in the earliest literary corpus (§ 3), the survey attempts to chronologically align various groups of Tibetan dialects based on their modern equivalents of OLT *ybras* (§ 4). The comparative reconstruction, based on data from 103 dialects, allows us to reconstruct the Proto-Tibetic form as *mbras (§ 5). To reinforce this reconstruction, I adduce data from other, presumably closely related Trans-Himalayan languages (§ 6).¹ In the final section (§ 7), I briefly discuss earlier hypotheses that sought to explain Literary Tibetan *ybras* in the light of other, usually remotely related Trans-Himalayan languages.

2. Reconstruction of historical dialects: Note on methodology

Old Literary Tibetan (OLT), the oldest documented Tibetic language, is the language of written artefacts composed roughly within the period of the Tibetan

1. I use the term “Trans-Himalayan” to refer to the language family formerly known as Sino-Tibetan or Tibeto-Burman. Firstly, “Trans-Himalayan” has geographic rather than ethnic, political, or linguistic connotations. Secondly, the term allows us to avoid the pitfalls of identifying ancient communities with historical or modern ones. Linguistic elements such as Sino-, Tibeto-, or -Burman refer to historically attested identities but might not be well applicable to pre-historic groups. Thirdly, it is not the function of such labels to capture the internal subdivision of a language family. And lastly, terms such as Sino-Tibetan or Tibeto-Burman are exclusive rather than inclusive; they pass over in silence hundreds of other languages included in the language family, making the impression that the “chosen” languages (Sinitic, Tibetic, or Burmish) are somehow better, or at least historically more important. This is not the case. In the present paper, I also use the widely accepted but linguistically poorly defined term “Bodic”. I do so for heuristic reasons to have a more encompassing name for an order higher than “Tibetic”. Since I plan to devote a separate publication to the linguistic definition of “Bodic” (see Bialek Forthcoming b), I leave this problem aside here with the sole remark that Tibetic, Tshanglic, East Bodish, Kaike-Ghale-Tamangic, and West Himalayish groups are hypothesised as potential candidates to be included in the Bodic branch of the Trans-Himalayan language family.

Empire (ca. 600–850s; Bialek 2021b) after the script invention in the 630s. The language should be distinguished, on the one hand, from Middle Literary Tibetan (popularly referred to as Classical Tibetan) and, on the other hand, from Old Tibetan. The latter was a spoken language dated approximately to 630s–800 (Bialek 2018b). From the second half of the 7th century onwards, spoken and written languages developed parallel, although the former continued to influence the latter. Building upon the proposal of Bialek (2018b), I distinguish between four main phases of Old Tibetan: Ancient, Early, Middle, and Late, each characterised by a set of distinctive linguistic traits — shared innovations. I adduce a brief discussion of Early Middle Tibetan dialects for these (i) seem to have formed shortly after the fall of the Tibetan Empire, (ii) were informed by the latest stage of Old Tibetan, and (iii) are still partly reconstructible from the available data.

Most recently, Tournadre & Suzuki (2023: 1022ff.) provided a list of 308 thus far identified modern Tibetan dialects. Their classification is complicated for mutual interactions of neighbouring dialects, multi-layered influences from prestigious dialects, and intense contacts with other non-Tibetic languages (especially in border areas) constitute serious obstacles to a clear-cut subgrouping. Moreover, the internal classification of modern Tibetan languages remains a complex problem due to the deep history of language documentation that goes back to the 7th century. Synchronically, modern Tibetan dialects are classed into dialect groups or, less securely, languages. Scholars have distinguished between five and eight such dialect groups depending on the methodology chosen.²

Languages do not move across space nor change over time by themselves; they need a speech community for both these processes to take place. Comparative reconstruction, the linguistic method resorted to in this article, is therefore *supplemented* here by our current knowledge of socio-historical facts from the earliest period of Tibetan history. Studying the history of Tibetic languages, we are lucky not only in having a rich linguistic corpus that spans approximately 1,300 years but also in having abundant knowledge of the social, political, and military history of the region and its peoples, which can help us illuminate the broader circumstances of language change. To understand the approach taken in this article, the reader should realise that at the end of the 6th century, Tibetan-speaking area restricted geographically to the south-eastern valley system of the Tibetan Plateau (see Map 1 and Bialek 2021b). In this region, the Tibetan alphabet was invented, and the orthography settled to be soon carried to all corners of the Tibetan Empire by Tibetan soldiers and administrators. In its starting point,

2. See Uray (1954: 312f.), Róna-Tas (1966: 21f.; 1985: 153ff.), Bielmeyer et al. (2018: 38ff.), and most recently, Tournadre & Suzuki (2023: 469ff.) for approaches that combine linguistic, historical, and geographical criteria. A critique of these rather heuristic groupings can be found in Sun (2003: 794ff.) and Hill (2019: 5).

the Tibetan writing system had one source, which was coreferential with the central chancelleries of the Empire. Although no documents witnessing the earliest spelling rules have been preserved, their general conservatism, which continues to this day, warrants the assumption that the orthography of the 630s did not considerably differ from that of the mid-8th century. More importantly, the cradle of the Tibetan polity and, in general, civilisation was simultaneously the centre of linguistic innovation from which consecutive waves of Tibetan speakers diffused over the plateau, each time bringing novel linguistic forms to their new homes. Speakers' groups that migrated out of this centre took the current stage of the language with them. Still, due to its pivotal political and cultural function, the centre kept innovating at a faster pace than the centrifugally spreading "peripheries". Needless to say, recently conquered regions did not turn over night to Tibetan-speaking areas, but the form of Tibetan that was to predominate there in the future was the one the conquerors brought with them from the central provinces (Three or Four Horns in the administrative nomenclature of the Tibetan Empire). This is in agreement with the general observation that "[i]n terms of demography alone there is nothing like warfare for the mass movement and relocation of populations" (Scott 2017: 193). When Tibetan colonising power settled and local administration could be effectively carried out, the "peripheries" became their own centres of innovation, changing independently from the current linguistic affairs elsewhere and moderately integrating elements of local non-Tibetic languages over time. I presume that the earliest dialectal diversification had its immediate source in this innovating force of the central political provinces and centrifugal migrations of Tibetan speakers all over the plateau.³

It is an axiom of the comparative method that only shared innovations warrant grouping two or more languages together. Since in the first centuries of the expansion, the central provinces remained the primary source of linguistic innovation, it is also non-participation in the innovations that characterises Tibetan dialects that split off during a previous phase of Old Tibetan. For example, the first group that split off from the core, Ancient Old Tibetan-North-East (AOT-NE), migrated northeast after the conquest of the Sum-pa (Bialek 2021b: 352ff.) The AOT-NE group, recruiting from soldiers and administrators, did not have its own cultural or linguistic identity beyond what it inherited from the core. In the meantime, the central provinces introduced innovations (characteristic of Early Old Tibetan) in which the separated group, by definition, could not participate

3. The hypothesis of the centrifugal spread of Tibetan culture and Tibetic languages from the Four Horns is buttressed by local rites and festivals still celebrated in Amdo, which recall wars against the Chinese during the Tibetan Empire. Songs accompanying the events allude to men being sent away from their homes in Central Tibet to serve as border guards in Amdo (Buffetrille 2008: 41ff.; Tournadre & Suzuki 2023: 500).

any more. The language of the AOT-NE group reflected the stage of the central language at the point of the group's migration. Of course, as time passed, these single proto-languages also started innovating, leading to the forms we find in modern vernaculars. However, the reconstruction attempted here focuses on the language stages when their carriers split from the central language. The approach is supplemented by the assumption that languages do not borrow archaic forms and, therefore, any such identified in a language must have been inherited from the proto-language. Accordingly, the first level of reconstruction focuses on defining linguistic innovations that (i) characterise Ancient, Early, Middle, and Late Old Tibetan and (ii) were shared by all proto-languages that descended from the respective phase. By way of illustration, dialect groups that did not participate in changes that occurred in Middle Old Tibetan (MOT) are classified as Early Old Tibetan dialects (EOT) and further subdivided into EOT-NW and EOT-S in agreement with their linguistic traits introduced sometime after their separation. Abbreviations (-E, -NW, -SW, etc.) used in addition to the language labels refer to the geographic location of the respective dialect group relative to the central provinces; e.g., EOT-NW denotes a dialect group that (i) separated from the core during the period of Early Old Tibetan (EOT), (ii) attests to sound changes of EOT but not to those of MOT, and (iii) its descendants are nowadays located roughly to the north-west (NW) from the central provinces.

Needless to say, a small-scale project such as this one cannot provide a complete description of all the identifying features of the reconstructed languages. Its main objective is to add another small piece of information to our understanding of the historical development of the Tibetic languages. Although in this paper I concentrate on modern equivalents of OLT *ybras*, insights gained from sound changes in reflexes of Proto-Tibetic *sl-, *zl-, *sr-, and *zr- (see Bialek 2018b), as well as some other presumably early changes, are likewise included into the analysis and have an impact on the proposed dialect subgrouping. Important early sound changes identified thus far, which can be used for subgrouping and determining the relative chronology of Old Tibetan, are provided below; they are shared by all the languages that separated from the core within the respective period. In addition, Table 1 lists linguistic traits that can be treated as diagnostics for the preliminary classification of a dialect. The table verifies the occurrence of the most conservative features in a given group, which might co-exist with more innovative forms, but the latter have not been considered.⁴ Language classification incorporating archaisms runs the risk of lacking regularity and being incomplete because it is almost impossible to identify *all* archaic forms, especially in under-

4. For a similar list, see Chart 3 in Tournadre & Rigzin (2015: 61). If not otherwise stated, the dialectal data are quoted after Bielmeier et al. (2013; 2018).

studied languages such as Tibetic. Modern linguistic studies concentrate on regular patterns and recursive phenomena, but it is the least common structures and lexical units that tend to contain the most conservative linguistic material.⁵ The result is that archaisms are only seldom reported in linguistic literature. To counteract or minimise the risk of including archaic forms into the analysis, the priority in factoring linguistic material is given to shared linguistic innovations following the requirements of comparative reconstruction.

Proto-Tibetic

- Palatalisation of alveolars * $[tj-]$, * $[dj-]$, * $[nj-]$, * $[zj-]$, * $[sj-]$, and * $[lj-]$:⁶
Proto-Bodic * sja > Proto-Tibetic * $\text{c}a$ > OLT *śa* ‘flesh’; Proto-Bodic * lej > Proto-Tibetic * $\text{z}ej$ > OLT *zín* ‘field’
- Excrescence $C^{[+nasal]}C^{[+liquid]}$ > $C^{[+nasal]}CC^{[+liquid]}$ (see § 7):
Proto-Tibetic * ral > IPFV $[^N]+[ral]$ > *ydral* $[^Ndral]$ ‘to tear asunder’
- Affrication $C^{[+nasal/r-/l-/s-]}C^{[+fricative]}$ > $C^{[+nasal/r-/l-/s-]}C^{[+affricate]}$ (see § 7):
Proto-Tibetic * $\text{c}ar$ > IPFV $[^N]+[\text{c}ar]$ > *yčhar* $[^N\text{t}car]$ ‘to rise’

Ancient Old Tibetan (AOT)

- Voicing assimilation of prefixes and excrescent consonants to the root consonant (see § 7):
Proto-Tibetic * lag > IPFV $[^N]+[\text{lag}]$ > OLT PFV *plags* (Classical Tibetan *bklags*) ‘to read’
- Merger $[tr-]$ + $[kr-]$ ⁷

Early Old Tibetan (EOT)

- Palatalisation $[h]$ > $[j]$ / $\#_{-}V^{[+front]}$ (see § 4.1):
Proto-Tibetic * hin (cf. Semkyi-Nyida/Rongdrak $[h\text{ĩ}]$)⁸ > EOT *yin* $[jin]$ ‘to be’
- Metathesis $[zl-]$ > $[lz-]$, $[sl-]$ > $[ls-]$:
Proto-Tibetic * zla > EOT * lza (cf. Balti $[lza]$) ‘moon’
- Excrescence $C^{[+fricative]}C^{[+rhotic]}$ (see § 7):
Proto-Tibetic * $srin$ > EOT * $strin$ (cf. Balti $[strin]$) ‘insect’

5. Cf. “[A] catalogue of the ‘irregularities,’ the nonproductive forms of a language, is at the same time a catalogue of the most archaic features of that language and those which are most valuable as tokens of its past” (Watkins 1973: 107).

6. On the phonemic status of alveolo-palatal in Old Tibetan, see Hill (2010: 118).

7. Judging from the rare occurrence of the clusters *tr-* and *thr-* in OLT (see Footnote 30), these were the first complex onsets with medial *-r-* to merge with the velar series. Deciding which Literary Tibetan forms come from Proto-Tibetic **tr-* and which come from Proto-Tibetic **kr-* would require a thorough comparative study, including data from cognate languages.

8. Tournadre & Suzuki (2023: 330); Suzuki (Forthcoming).

Middle Old Tibetan (MOT)

- Devoicing of obstruents in plain onset:⁹
Proto-Tibetic *dar > Themchen [tar] 'ice'
- Merger [dr-] + [gr-]:
Proto-Tibetic *drug > Themchen [tʃsəç] 'six' and Proto-Tibetic *gru > Themchen [tʃə] 'boat'
- Merger *sdr- + [zgr-]:¹⁰
Proto-Tibetic *sdrun (cf. Balti [zdrun], OLT *sgrun*) > Chabcha [əɖzun] 'story' and Proto-Tibetic *sgra > Chabcha [əɖza] 'sound'

Late Old Tibetan (LOT)

- Loss of preconsonantal *g-*, *b-*, and *r-*:
Proto-Tibetic *bdun (cf. Balti [bdun]) > Thewo Tö [dē] 'seven'¹¹
- Merger [ʰN] + [m] / #_C-:¹²
Proto-Tibetic *^Ndu (cf. Themchen [ndə], OLT *γdu* [^Ndu]) > Thewo Tö [ndə] 'to gather'¹³ and Proto-Tibetic *mda (cf. Themchen [mda]) > Thewo Tö [ndə] 'arrow'¹⁴

Classifications of Tibetan dialects proposed in previous works such as Bielmeier et al. (2013; 2018) or Tournadre & Suzuki (2023) differ from the one put forward here in that their primary objective is the subgrouping of contemporary vernaculars. In contrast, the present article attempts to reconstruct the earliest days of

9. Devoicing of fricatives [z-] and [z-] in absolute onset seems not to be completed in dialects descending from MOT-NE. Hiroyuki Suzuki (p.c. 2024-06-11.) kindly informed me that likewise Sharkhog, Thromjekkhot, Thewo Tö, and Pälkyi descending from LOT-E partially preserve voiced equivalents of Literary Tibetan *z-* and *z-*. However, this is not borne out in Lin's data (2014) for Thewo Tö. Therefore, devoicing might have reached fricatives later than plosives and affricates, spreading over the whole lexicon after LOT-E had separated from the core. The *terminus ante quem* for the start of devoicing, in general, can be set to 675/6 when the demonym Dru-gu (< Old Turkic *tür(ü)k*) occurs for the first time in the *Old Tibetan Annals* (IOL Tib J 750: 64; Bialek 2018b: 29). Bialek & Suzuki (Forthcoming) present a detailed analysis of devoicing in Tibetic.

10. The merger might have started in EOT but was completed first in MOT. This might explain the fact that the outcome of this merger was accepted in the orthography because, with very few exceptions, the etymological *zr-* is written as *sgr-* in OLT (Bialek 2018b).

11. Lin (2014: 262).

12. Throughout the paper <^N> indicates a reconstructed preconsonantal nasal sound, not the voiced uvular nasal of the IPA. The exact phonetic characteristic of this sound, usually represented in the script by *ṙ* *γ*, remains a bone of contention among Tibetologists; see §7 *infra*.

13. Lin (2014: 251).

14. Lin (2014: 247).

Table 1. Archaic features in single dialect groups

Linguistic trait	AOT		EOT		MOT	LOT					EMT	
	NE	NW	S	NE	SW	E	N	S	C	W	E	
Metathesis sL>Ls-	-	+	+	+	+	+	+	+	+	+	+	+
sDR-		+		+	+	+	+	+	+	+	+	+
Voiced onset	+	+	+	± ^a	-	-	-	-	-	-	-	-
Kr- & Tr-	+	+	-	+ ^b	-	-	-	-	-	-	-	-
sdr- & sgr-		+		-	-	-	-	-	-	-	-	-
kr- & gr-	+	+	+	+ ^b	-	-	-	-	-	-	-	-
gC-, dC-, bC-, rC-	+	+	-	+	-	-	-	-	-	-	-	-
pr- & br-	+	+	+	+ ^b	+	-	-	-	-	-	-	-
^N C- & mC-	+	-	-	+	-	-	-	-	-	-	-	-
kj- & kr-	-/+	+	+	+ ^b	-	-	-	-	-	-	-	-
pj- & bj- ^c	-	+	+	+ ^d	-	-	-	-	-	-	-	-
-Vs & -Vd		+	+	± ^e	-	-	-	-	-	-	-	-

a. The voice distinction is preserved only in fricatives.

b. This distinction is preserved only in Sastod (Sun 2025).

c. Suzuki (2020b) pointed to reflexes of Proto-Tibetic *[Cr-] and *[Cj-] as important indicators for advanced sound changes in southern Khams dialects.

d. This distinction is partly preserved but only in Sastod (Sun 2025).

e. Final [s] and [d] are preserved only word-internally in Sastod (Sun 2025).

dialect formation, working first of all from shared lexical and phonetic innovations while taking into account retentions and archaic traits. However, we shall keep in mind that innovations might skew the historical perspective since innovative traits have a strong tendency to spread across dialects or even across dialect groups as we observe nowadays, for example, between the descendant languages of EOT-NW and EMT-W or between AOT-NE and EMT-E (see Map 4). Consequently, two neighbouring languages may share some innovations despite not being historically closely related. Their distinct provenance may be demonstrated by revealing rare archaic traits present in one language but not in the other. Of course, even the complete absence of archaic traits must not mean a language is more recent. The ultimate goal of studying the history of dialect formation in Tibetic must be to complement the insights gained from both approaches. As long as this is not possible, we shall bear in mind that, for example, the language I call EOT-NW might not be identical with Proto-Western Archaic Tibetan (in Biemeier et al.'s (2013; 2018) terminology) or Proto-North-Western section

(according to Tournadre & Suzuki 2023) as reconstructed only from modern vernaculars.

Although the article is primarily concerned with the reconstruction of one word from dialects for which enough data is available, I can conclude that modern equivalents of OLT *ybras* accord with the typical sound changes encountered in those dialects. Whenever this should not be the case, a separate remark will clarify the historical position of the diverging form. It is worth noticing, however, that in the corpus of 103 dialects examined below, only one dialect, Drenjong (LOT-S) in Sikkim, possesses synchronically two equivalents of OLT *ybras*: [bja] and [rɛ:], the latter being a borrowing from another dialect. This would suggest that sound changes spreading from neighbouring, more innovative dialects were only seldom introduced through direct borrowings but might have instead been induced in regular contact situations.

As a subsidiary goal, the article looks at the potential cognates of OLT *ybras* in neighbouring Trans-Himalayan languages to (i) buttress the reconstruction of Proto-Tibetic *mbras and (ii) suggest its immediate ancestor form. I make no claim concerning the higher-order classification of the respective languages or their exact relationship to Tibetic. Such an approach would necessitate a more comprehensive investigation, including a much broader sample of data. This I hope to present in the near future with a thorough examination of reflexes of Proto-Bodic *ml- and *mr- onsets (Bialek Forthcoming b).

Typographical conventions

Literary Tibetan forms are transliterated following Bialek (2020b) and written in italics. Romanisations and unacademic transcriptions are quoted using a standard typeface. The standard typeface is also used for reconstructed forms that are preceded by an asterisk <*>. When discussing sounds rather than letters, the International Phonetic Alphabet (IPA) symbols are used. Phonemes are enclosed in slashes // and sounds in brackets []; thus, *ybras* = transliteration, *mbras = reconstructed form, [mbras] = phonetic transcription, /mbras/ = phonemic transcription. For legibility's sake, brackets are omitted from tables.

As a corollary, this article proves the necessity of making a strict division between *transliteration* of written forms and *transcription* of words arrived at by comparative reconstruction. The use of phonetic transcriptions to forms attested in script (see Jacques 2012) is misleading and ahistorical, as demonstrated by the case study of *ybras*. The transcription of འབྲས་ *ybras* as ¹[^Nmbras] projects a very concrete phonetic interpretation onto the first letter. The present study reveals that this phonetic reading cannot be earlier than the late 8th century. The original reading at the time of the script invention was *mbras /mbras/ although at the

same time ࠠདྲལ་ *ydral*, written with the same prescript ࠠ, was pronounced as $*^N$ dral (< $*^N + \sqrt{\text{ral}}$).¹⁵

In this paper, transcriptions of linguistic data strictly adhere to the sources; no attempt at unification has been undertaken. This calls for special caution in interpreting the data, especially as the word under examination attests to the complex onset with the series *Cr-*. Spoken equivalents of Literary Tibetan *Cr-* have been variously treated in literature and, as perceptively observed by Suzuki (2024c: 108f.), “the difference between retroflex plosives and affricates corresponding to the *r*-glide series is due to the scholarly tradition of descriptions than to an actual phonetic discrepancy”. Neither shall we forget that linguistic studies of Tibetan dialects have been carried out for over eighty years (150 if we start counting from Jäschke). This implies that data collected earlier might reflect a different stage of the language (especially if elicited from older community members) than data surveyed nowadays. And the fact is that we have no means to double-check it. To minimise the risk of discussing spurious data, I have checked the author’s transcription policy whenever available in the original source and whenever the data was relevant for the discussion.

3. *ybras* in Old Literary Tibetan

The simple word *ybras* is attested in OLT in two senses: 1. ‘crop; grain’, and 2. ‘rice’, as exemplified by (1) and (3)–(4) respectively:

- (1) *ñen_kar=nī dog=dañ ñe/ ybras drug=ni si_lī_li//*
 Ñen_kar(ABS)=FOC valley=COM be_close grain six(ABS)=FOC rustle
 ‘Ñen-kar lies near the valley. [Its] six crops rustle, rustle.’

[Pelliot tibétain 1287: 422]

The passage comes from a song sung by queen Sad-mar-kar married to the *Žaṅ-zuṅ* ruler Lig Myi-rhya, probably in the 620s (Bialek 2021b: 356, Footnote 38). This is the oldest roughly datable attestation of *ybras* in the meaning ‘crop; grain’. *ybras drug* ‘six crops’ seems to have had a deeper cultural meaning, for the phrase recurs in OLT records several times. Unfortunately, the six crops are not particularised in any text, but the term must have referred to crop grains cultivated in the southern valleys of the Tibetan Plateau – the cradle of Tibetan culture; cf.:

15. The situation complicates if we consider that Proto-Tibetic had an inherited onset cluster $*dr-$ and a possibly productive derivation by means of prefix $*m-$. A logical conclusion is that Proto-Tibetic might have had forms in $*mdr-$ and $*^Ndr-$ that are both represented by ࠠདྲ- in OLT.

- (2) *gyaṅ+mo=nī* *byaṅ=na* *γphan//*
 sheep+NMLZ.F(ABS)=FOC north=INESS be_prolific
ybras drug=nī *lho=na* *mod//*
 grain six(ABS)=FOC south=INESS COP.EMPH
 'Sheep breed in the north.
 Six crops are abundant in the south.' [IOL Tib J 738: 3v79–80]

Ñen-kar mentioned in (1) was located in Lo valley, a side valley of the Skyi-čhu river (Modern Tibetan Skyid-čhu). The meaning 'grain' must have also underlain the compound *rgya ybras* (< **rgya bayi ybras*) 'buckwheat', lit. *'grain which is broad'.

In (3), the mention of *ybras* together with *nas* 'barley' supports its reading as referring to a plant species. The passage states that rice was harvested in the Rkoñ area at the end of the 8th century.¹⁶ Example (4) confirms rice cultivation on the Tibetan Plateau by the end of the 7th century – it quotes words uttered by Mgar Khri-ybriñ Bcan-brod, a grand councillor from the year 685/6 onwards who fled to China in 698/9.

- (3) *scaṅ+dphyay nas=daṅ//* *ybras gaṅ* *ybul=yaṅ* *ruṅ//*
 grain+tax barley=COM rice what(ABS) IPFV:give=ADD be_suitable
ybul+bayī *sa_chigs//* *da_ltar=gyī=las myī bsriṅ*
 IPFV:give+NMLZ:GEN station(ABS) now=GEN=DEL NEG DPASS:lengthen
 'Whatever grain-taxes, [in] barley or rice, are to be paid, the staging post for the delivery shall not be moved to a more distant place than it is now.'
 [Rkoñ-po inscription 19; trslr. *apud* Iwao et al. (2009: 16)]

- (4) *nas+ybras* *thaṅ gaṅ* *skyes=pa=yaṅ/* *raṅ_tag*
 barley+rice(ABS) plain full(ABS) grow:PFV=NMLZ(ABS)=ADD mill
gčhīg=gi *naṅ=du* *čhib=bo//*
 one=GEN inside=TERM be_equal=FNL
 'Even barley and rice that overgrew whole plains fit in a single mill.'
 [Pelliot tibétain 1287: 503–504]

Notably, in (3) and (4), rice is mentioned together with barley – until today, the main staple crop on the Tibetan Plateau. This co-occurrence indicates the significance of rice in the diet of local populations and confirms that (4) describes

16. For the most recent dating of the Rkoñ-po inscription, see Bialek (2021a: 21f.)

the situation of the southern Tibetan Plateau, too.¹⁷ Example (3) also adds to our knowledge by applying *scañ* as a hypernym for *ybras* and *nas*.

The more concrete meaning ‘rice’ is additionally confirmed in the compound *ybras skyems* ‘rice-beverage’ occurring side by side with *dgun* (read: *rgun*) *skyems* ‘grape-beverage’, *sbrañ skyems* ‘honey-beverage’, *khre yu* ‘millet-beverage’, and *gro skyems* ‘wheat-beverage’ (Pelliot tibétain 1042: 72–73, 76). Compare also the following passage with *ybras* used for a concrete plant species:

- (5) *nas khal gčig=dañ/ gro khal drug=dañ/ khri* (read: *khre*) *khal*
 barley khal one=COM wheat khal six=COM millet khal
gsum=dañ/ ybras bre gañ/
 three=COM rice bre full(ABS)
 ‘one *khal* of barley, six *khal* of wheat, three *khal* of millet, and one full *bre* of
 rice’ [Or.8210/S.5721: 4]

4. Equivalents of OLT *ybras* in modern Tibetan dialects

Data on modern equivalents¹⁸ of OLT *ybras* are abundant and allow for a preliminary clustering of dialects into eleven major sub-groups that will be discussed

17. On the cultivation of rice in Tibet, see Stein (1972: 24f.) Nowadays, rice continues to be cultivated in the Rja-yul (Chinese Cháyú 察隅) and Me-tog (Chinese Mòtuō 墨脱) Counties on the Tibetan Plateau (Song et al. 2018: 2021f.), as well as in the southern part of Yunnan Tibetosphere (Hiroyuki Suzuki, p.c. 2024-06-16). Due to the scantiness of archeological excavations from the southeastern parts of the Tibetan Plateau, no additional support can be quoted for the textual evidence adduced above. The diversity of ecologies and vegetation types over the geographical region known as the Tibetan Plateau precludes any extrapolation from archaeological sites on the fringes of the area to any specific locality. For example, the only site from the Tibetan Plateau surveyed in d’Alpoim Guedes (2011) is Mkhar-ro in Čhab-mdo (Chāmdü Kāruò 昌都卡若; ca. 2700–2300 BC), whereas d’Alpoim Guedes et al. (2014) discuss Yphren-go (Chāngguò 昌果; Goñ-dkar County) for the early 1st millenium BC and Khyuñ-luñ (Mñay-ris) for the 1st millennium AD. At the latter site and in the neighbouring Zebang, Song et al. (2018) reported rice grains for the period 455–700 CE. Wang et al. (2024) restricted to excavation layers from the prehistorical period. Since rice does surface in many early Tibetan documents, frequently side-by-side with barley, I do not see any reason to doubt the correctness of the written sources. The fact that rice remains largely invisible to archaeological studies might be related to the scattered data, many geographical gaps, and time span targeted in the excavations. Requiring considerably more human labourers than other crops (Scott 2009: 64ff.), rice may have started being cultivated only in the imperial period when military conquests secured enough manpower, and it might have formed part of the diet of the political elites only.

18. I use the word “equivalent” instead of “reflex” because written forms should not be directly equated with etymological ones that can be reconstructed to Proto-Tibetic.

further below. Table 2 lists forms collected thus far from secondary sources; the data within the main groups are arranged according to the similarity of sound correspondences. Modern dialects generally agree on the meaning '(paddy) rice'. The only dialect that attests to the meaning 'grain' is Phongpa of southern Khams; the meaning 'rice' is housed in this dialect by the equivalent of OLT *drus ma*.¹⁹ These confirm the OLT data and allow us to reconstruct the meaning *¹'grain; ²'rice' to Proto-Tibetic *mbras. Sound changes described below for single dialect groups take the reconstructed *mbras as their starting point (on this reconstruction, see § 5. Proto-Tibetic *mbras).

Table 2. Modern equivalents of OLT *ybras*^a

OT	Group	TS	Language	Transcription	Source
AOT	NE	17	Sprosnang	ᵐbrɛ:	Suzuki (2007: 36)
		*mbras	22	Phongpa	ᵐbre:
	Mthachu		ᵐbɛ:	Suzuki (2012a: 110)	
	Zhollam		ᵐbɛʰ:		
	Gyälthang		ᵐḡur, ᵐḡur:		
	Rgyalbde		ᵐḡur:		
	Yangthang		ᵐḡur:		
	Daan		ᵐḡur:		
	Skadrag/Mtshongu		ᵐḡi:	Suzuki (2012a: 110)	
	Maoniuping		ᵐḡi:		
	Nyishar		ᵐḡi		
	Byagzhol		ᵐḡi:		
	Qidzong		ᵐḡe:		
	Phuri		ᵐḡi:	Suzuki (2012a: 110)	
	Lamdo		ᵐḡe:	Suzuki (2012a: 111)	
	G.yaglam		ᵐḡe:	Suzuki (p.c. 2022-03-26)	
	Gsertshang		ᵐḡe:		
	Khrezhag		ᵐḡe:	Suzuki (2024b: 8)	
	Mbacug	ᵐḡe:	Suzuki (2012a: 110)		
EOT	NW	46–49	Balti	bras~blas	
			*mbras	*bras	
			Turtuk	bras	Bielmeier et al. (2013: 5964)
			Kargil	bras	
			Tshangra	bras	

19. This information has been kindly supplied by Hiroyuki Suzuki (p.c. 2022-03-26).

Table 2. (continued)

OT	Group	TS	Language	Transcription	Source
			Parkachik	bras	
			Thuwina	bras	
			Sod	bras	
			Darket	bras	
			Hanu	bras	
			Chiktan	bras	Bielmeier et al. (2013:5964)
			Sapi	bras	
			Shargol	bras	
			Mulbek	bras	
			Lamayuru	bras	
			Wanla	bras	
			Khalatse	bras~ɖas ^b	
			Lingshet	ɖas	
			Dibling	ɖas	
			Nurla	ɖas	Bielmeier et al. (2013:5964)
			Nyemo	ɖas	
			Leh	ɖas	
			Khalsar	ɖas	
	S	75	Choca-ngaca	bra:	Tournadre & Rigzin (2015:53)
	*bras	76	Brokpa	bre:	Wangdi (2021:935, Footnote 3)
MOT	NE	23–34	Rkangtsha	mɖzi	
*mbras	*mbras		Themchen	mɖzwi	
			Chabcha	nɖzi	
			Bayan	nɖzi	
			Shando	nɖze	
			Labrang	mɖze	
			Rmastod	nɖzi	Bielmeier et al. (2013:5964)
			Mdzorganrabar	nɖzi	
			Rngaba	nɖzi	
			Ndzorge	nɖzi	
			Arik	mɖzi	
			Mkharmar	mɖzwi	
			Golok Sertha	mdai:	Sprigg (1979:58)
			Lhagang	^m ɖe	Suzuki & Wangmo (2017:110)
			Sastod	ⁿ bri	Sun (2025:743)

Table 2. (continued)

OT	Group	TS	Language	Transcription	Source
LOT	SW	65.4	Kyirong	brɛː	Bielmeier et al. (2013: 5964)
* ^N bras	* ^N breː			bɛ̃ɛ̃	Webster (2022: 30)
		65.3	Tsum	bre	Donohue & Dhakal (2016: 4)
				bre	Webster (2022: 30)
		65.6	Yolmo	brɛː	Bielmeier et al. (2013: 5964)
E		36	Thewo Tö	ⁿ dzɛː	Lin (2014: 260)
* ^N dzɛː		37	Thewo Mä	ⁿ ɖɛ	Suzuki (2022a: 74)
		35	Cone	ndzɛː ^L	Jacques (2014: 276)
		44	Zhongu	ndzɛ	Sun (2003: 823)
N		9–10	Gertse	ndzɛː (lw)	Bielmeier et al. (2013: 5964)
* ^N bas			Hor	bɛʔ	
			Amdo	bɛʔ	Bielmeier et al. (2013: 5964)
			Nakchu	bɛʔ	
			Bachen	baː	
S		72	Dzongkha	bdzaː~bzaː	Yliniemi (2021: 55, 260)
* ^N bjas		70	Drenjong	bja/reː (lw) ^c	
C		5.3	Lhasa	ʈʂɛʔ	
* ^N dzas		4.2	Shigatse	ʈʂiɛ	Bielmeier et al. (2013: 5964)
		1.6	Gar	nʈʂɛː	
		1.1	Ruthog	nʈʂɛː	
		1.8	Purang	nʈʂɛː	Bielmeier et al. (2013: 5964)
		1.2	Gergye	nʈʂɛː	
		1.4	Tshochen	nʈʂɛː	
		1.16	Dingri	ʈɛː	Bielmeier et al. (2013: 5964)
		61	Humla (Til)	ɖɛ̃	
			Humla (Muchu)	ɖɛ	
			Humla (Yalbang)	ɖrai	de Vries (2020: 54)
			Humla (Kermi)	dai	
			Humla (Yakpa)	ɖa ⁱ	
		64	Southern Mustang	ɖɛ~ʈɛ draä	Bielmeier et al. (2013: 5964) Nagano (1983: 481)
		65.2	Nubri	ʈɛː	Bielmeier et al. (2013: 5964)
			Nubri(Sama)	ɖiʈɛ	
			Nubri(Lho)	ʈiɛ	Webster (2022: 30)
			Nubri(Prok)	ʈiɛ~ɖɛ	

Table 2. (continued)

OT	Group	TS	Language	Transcription	Source
		68	Lhomi	ʈe	Vesalainen (2016: 125)
		67	Jirel	ɖa	Bielmeier et al. (2013: 5964)
		66	Sherpa	ɖa	Kelly (2004: 288)
EMT	W	53–60	Trashitongde	ɖe	
* ^N ɖzas	* ^N ɖzas		Trangtse	nɖe	
			Man-Merak	nɖe	
			Tabo	nɖe	
			Nako	nɖe	Bielmeier et al. (2013: 5964)
			Namgya	nɖe	
			Nesang	nɖe	
			Tholing	nɖza:	
	E	11–15	Kandze	nɖzɛ:	
	*nɖzɛ:		Derge	nɖzɛ̃:	
			Dartsendo	nɖzɛ̃:	Bielmeier et al. (2013: 5964)
			Bathang	nɖzɛ̃:	
			Lithang	nɖzɛ̃:	
			Sangdam	ˈnɖi:	Suzuki (2012b: 53)
			Dgudzong	ˈnɖi:	Suzuki (2011: 624)
			Khyungpo	ˈmɖɛ:	Suzuki (2010: 106)

a. The transcriptions strictly follow the sources. The third column (TS) provides the number of a dialect group in Tournadre & Suzuki (2023: 1022ff.) Shaded cells indicate dialects, my classification of which diverges from that of Tournadre and Suzuki.

b. This I understand as synchronically alternating variants. No further explanation is provided in the source. The data seem to have been collected during editors' fieldwork.

c. [re:] is used in the meaning 'rice kernel'.

The subgrouping presented in Table 2 refines the one proposed in Bialek (2018b: Appendix B), which was based mainly on Bielmeier et al. (2013; 2018). For the present study, information from dialects not included in Bielmeier et al. (2013; 2018) was additionally supplied, allowing for more fine-grained distinctions, especially concerning Khams and Amdo dialect groups. Dialects in Table 2 are arranged according to the phase of Old Tibetan (leftmost column), from which their proto-languages seem to have separated. In the remaining part of this section, I will comment on the changes observed in the single dialect groups, attempting to reconstruct their proto-forms. Needless to say, the subgrouping remains tentative.

4.1 Ancient Old Tibetan dialect group

It has been argued that all documented Tibetic languages go back to Early Old Tibetan, so, by definition, they cannot contain information on language stages preceding the script invention (Bialek 2018b; Hill 2019: 5). However, the identification and documentation of Semkyi-Nyida, Derong-Njol, Gyälthang, and Rongdrak by Hiroyuki Suzuki seem to render this view obsolete and support Biellemeier’s (1982: 406) assumption that “certain linguistic phenomena in modern dialects can, in terms of relative chronology, be more archaic than their equivalents in, e.g. W[ritten] T[ibetan] (= my Literary Tibetan – JB)”. Accordingly, I deem Ancient Old Tibetan the last common ancestor of all Tibetic languages from which AOT-NE separated. The latter, thus far the only dialect group classified as AOT, was the language of the military elites who conquered and colonised the Sum-pa in the early 7th century (see Map 1), before the script introduction.²⁰ The geographic distance between, on the one hand, Derong-Njol, Semkyi-Nyida, Gyälthang and, on the other hand, Rongdrak suggests that they might be remnants of an older language stratum and were separated in space by later migrating speakers of more innovating forms of Tibetan, first MOT-NE, and then LOT-E and EMT-E (see Map 2).

4.1.1 AOT-NE *mbras

AOT-NE encompasses dialect groups nowadays located far away from each other: Derong-Njol/Semkyi-Nyida/Gyälthang (south-western Kham) and Rongdrak (far-eastern Kham; see Map 2).²¹ The most important characteristic that relates these groups to each other is that, despite the considerable geographical distance between Rongdrak and Zhollam (Semkyi-Nyida), the two share uniquely one grammatical innovation that must have been introduced to their last common ancestor and is not encountered anywhere else:

- Ablative marker *das /de/²² \bowtie *ydas*, v2 of *yday* [^Nda] ‘to pass over’²³

Before grammaticalising *nas* and *las* as the elative and delative markers, Old Tibetan used *se* as a general ablative marker (Bialek 2026: 426ff.). The latter is

20. The Sum-pa were the first non-Tibetic speaking group known to have been entirely conquered by the Tibetan Empire (Bialek 2018b: 25f.; 2021b: 352ff.)

21. Suzuki (2024a); cf. also Tournadre & Suzuki (2023: 479ff.) who characterised Rongdrak as “considerably different from any Tibetic varieties spoken in Kham” (ibid., p. 482).

22. Tournadre & Suzuki (2023: 315, 493).

23. According to Kuteva et al. (2019: 98, 251f.), verbs ‘to come from’ and ‘to leave, to abandon’ are two most common sources of ablative markers cross-linguistically.

found only in ritual texts that contain a fair number of archaisms. It seems, therefore, less likely that AOT-NE would have inherited the ablative *das /de/ from AOT, which at that time had a distinct morpheme with this function: *se*.

Moreover, Semkyi-Nyida and Rongdrak attest to the palatalised *h- before front vowels (Tournadre & Suzuki 2023: 330; Suzuki Forthcoming), which feature predates the known orthography rules and supports the view that AOT-NE separated before or roughly simultaneously with the script invention that perpetuated in the script the innovated form *yin*: AOT *h- > OLT *y*- [j]- / _#-V^[+front].²⁴ Concerning other archaic features, Derong-Njol has [ʰl-] in the equivalents of OLT *zl*- (Suzuki 2009: 227; cf. also Melung [ʰlɛ] for *zla* in [ʰlɛ gɛ: mō] ‘moon’ ≠ OLT *zla dkar mo* (?)),²⁵ indicating that its ancestor language did not participate in

24. Cf. OLT *yin* ‘to be’ vs Semkyi-Nyida/Rongdrak [hĩ] (Tournadre & Suzuki 2023: 330) and Tamang [hin] (Michailovsky & Mazaudon 1994: 552); Chantyal [hin] (Noonan 1999: 40); Bunan [fien] (Widmer 2017: 75); Rongpo [hwən] (Zoller 1983: 335). Beside such orthographical variants as *thems* ~ *ěhems* ‘testament’, *mthiin* ~ *mčhiin* ‘azure’, or *mthil* ~ *mčhil* ‘bottom’ still found in OLT ritual texts, the AOT onset [hV^{+front}] indicates that the palatalisation before front vowels was not fully completed by the early historical period.

25. Tournadre & Suzuki (2023: 805). The same observation might be true for Minyak Rabgang, another dialect group from eastern Khams. According to *Bing*-series *Xīfānguǎn Yìyǔ* 西番館譯語, a Chinese-Tibetan glossary edited in the 16th century and preserving historical data on Minyak Rabgang (Suzuki 2014), the local pronunciation of Literary Tibetan *zla ba* ‘moon’ contained the sound [l] (modern pronunciation is [ʰda wa] in the Lhagang variety of Minyak Rabgang (Suzuki 2014) which might be a later borrowing from an Amdo dialect). The glossary provides Chinese transcription of the respective word as 老瓦 [law-wa] (Nishida 1963: 148). This pronunciation seems to be confirmed by the spoken equivalent of Literary Tibetan *zla ba zlum po* ‘round moon’ transcribed as 老瓦輪卜 [law-wa lum-pu] (ibid., p. 150), where Literary Tibetan *zl*- is twice rendered with an [l]. However, the phonetic interpretation of the data contained in those Chinese glossaries is not without a problem, for we also find *la- (*lá* 刺) in a local equivalent of Literary Tibetan *zla ba* in *Yi*-series *Xīfānguǎn Yìyǔ* 西番館譯語 dealt with by Miller (1955: 289). Since this work most probably contains reading pronunciations (see Footnote 34 *infra*), I feel reluctant to base any reconstruction of Tibetan historical dialects solely on these Chinese-Tibetan glossaries. The suspicion is additionally substantiated by data on numerals in *Bing*-series *Xīfānguǎn Yìyǔ*: [ɳi tɕu] for Literary Tibetan *ñi su* ‘20’, [zɳi tɕu] for Literary Tibetan *bzi bču* ‘40’, [ɳja tɕu] for Literary Tibetan *lha bču* ‘50’, or [gu tɕu] for Literary Tibetan *dgu bču* ‘90’ (Nishida 1963: 173) to name the most striking examples. [ɳi tɕu] would be an equivalent of **gnis bču*, a form never attested, whereas [zɳi tɕu], [ɳja tɕu], and [gu tɕu] lack traces of the otherwise omnipresent word-internal [b]. Should further data on Minyak Rabgang nevertheless confirm that it indeed preserves some archaic features, the language could be another descendant of AOT-NE, which is not impossible judging from its geographic proximity to the Rongdrak group. I wish to thank Hiroyuki Suzuki for drawing my attention to and providing secondary literature on glossaries known as *Xīfān Yìyǔ*. These important sources of information on historical Tibetan dialects were prepared by Chinese governmental offices

the metathesis [zl-] > [lz-] postulated for EOT (Bialek 2018b: 34). An additional piece of evidence comes from Geshiza (a West Rgyalrongic language in direct contact with the Rongdrak group), where we find the borrowing [zlæç^hɔ] ⌘ Literary Tibetan *zla sól* 'leap month' (Honkasalo 2019:727), that displays no traces of the metathesis.²⁶ Concerning lexical peculiarities, as the only Khams dialects, Melung (Semkyi-Nyida) and Rongdrak retain equivalents of OLT *byi la* 'cat', whereas the remaining dialects of the region exhibit more innovative forms (Suzuki 2023:232). Finally, several languages of the Rongdrak group retain reflexes of Proto-Tibetic *zo 'milk' (OLT *zó*), a word that was elsewhere (except for Choca-ngaca) replaced by equivalents of OLT *yo ma*.²⁷ All these are the oldest features found thus far in Tibetic languages, to which we may add remnants of voiced onset in Phongpa (cf. /beʔ/ ⌘ OLT *byed*; Suzuki 2020a:34), as well as some singular forms displaying what appears to be archaic traits: Phongpa [ʰgru: ma] ⌘ *grog ma* 'ant' with a reflex of a lost prefix;²⁸ Sprosnang [ʰkru] ⌘ OLT v3 *bkru* 'to wash' (v1 *ykhrud*; Suzuki 2007:35); and the Pomtserag free form [ʰpa: pa] (Suzuki 2009:227) ⌘ OLT *-lpags* (?) 'skin'.²⁹ It is generally acknowledged that Tibetan dialects have not preserved any traces of the passive v3-stems from OLT conjugations. Sprosnang [ʰkru], if underpinned by further examples, would prove the existence of the respective forms in the spoken language.

Concerning other linguistic features that support the early formation of this group, data presented by Suzuki (2022b) indicate that at least Tsharethong, Snyingthong, Sakar, Lothong, Phongpa, Zhollam, and Melung from the Semkyi-Nyida group have segmental equivalents of OLT prescribed letters *d-*, *g-*, and *b-*, whereas Pomtserag from the Derong-Njol group displays traces of all complex clusters in onset, albeit reduced to prenasalisation or preaspiration assimilated in the feature [voice] to root consonants (Suzuki 2009). Their proto-language(s)

between the 14th and 18th century, and they contain Chinese transcriptions of foreign vocabulary (Suzuki 2014).

26. Furthermore, Geshiza [sræn] ⌘ Literary Tibetan *sram* 'otter' (Honkasalo 2019: 727) confirms that the donor language did not participate in the D-epenthesis – another innovation that occurred in EOT after the script invention (Bialek 2018b: 34).

27. Tournadre & Suzuki (2023: 675). This lexical retention is interesting in the light of OLT data where *yo ma* is virtually unknown, and only *zó* is used for 'milk' (cf. Pelliot tibétain 1068:99; Pelliot tibétain 1285: r163, r169, r170; Pelliot tibétain 1289: r2.2). This once again proves the conservative character of the written language.

28. Suzuki (2020a: 33); cf. in this context Gyalrong *sk^hrow/sk^hre/sk^hru*, Stau *sk^hşau/sk^hro*, and Khroskyabs *skhráy* 'ant' (*apud* STEDT).

29. Cf. also Golok [ʃfak]~[fak] 'eyelid' (Bielmeier et al. 2013:4787). In OLT, the form *-lpags* is used only as the second syllable of a compound, the first constituent of which is an open syllable (cf. Bialek 2018a: vol. 1, page 245, Footnote 2); the free form is *pags* (*pa*).

must have separated from the core group before the respective consonants got lost in LOT (see § 4.4). The available data also indicate that Sprosnang (Rongdrak), Myigzur, Choswateng, Khyimphyuggong, Rtswamarteng, G.yaglam, Gyälthang, Mbacug (Gyälthang), Mbalhag (Derong-Njol), Dngo, Lamdo, Phongpa, Zhollam, Daan, and Mthachu (Semkyi-Nyida) possess distinct equivalents for OLT *Kr-* and *Tr-* series (Suzuki 2024c: 109, Table 3), which means that their proto-language(s) separated from the common stock before the introduction of the merger [Tr-] + [Kr-] in MOT.³⁰ As expected, the languages of these groups likewise display distinct equivalents for the series [Kr-] and [Pr-].³¹

In light of the data from Table 2, AOT-NE attests to the greatest internal variety in the equivalents of OLT *ybras*, which Suzuki (2012a) already thoroughly discussed. The following discussion is much indebted to his analyses in a series of papers on the Semkyi-Nyida dialects.³² In general, we can divide the data in Table 2 into five groups:

- a. [ᵐbr-] Sprosnang, Phongpa
- b. [ᵐb-] Mthachu, Zhollam
- c. [ᵑg-] Gyälthang, Rgyalbde, Yangthang, Daan, Skadgrag, Maoniuping, Nyishar, Byagzhol, Qidzong
- d. [pɣ-] Phuri, Lamdo, G.yaglam, Gsertshang, Khrezhag
- e. [ᵐdz-] Mbacug³³

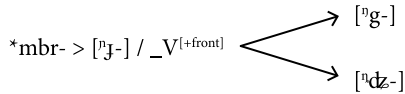
30. However, the merger must have started with [tr-] already before AOT because the combination is not attested in any modern dialect (WAT [tr-] in equivalents of OLT *dr-* (see Bielmeier et al. 2013) seems to be an ongoing change introduced after the split-off of EOT-NW) and surfaces in OLT only in a few proper names: Mal-t(h)ro (toponym), Tre-goñ, and Tre-mye-slebs (personal names). There is one puzzling exception that concerns the numeral *drug* ‘six’, the equivalents of which in all dialects that did not merge [Kr-] + [Tr-] display regular outcomes characteristic of [Tr-]. However, other Trans-Himalayan languages point to a velar onset: Dakpa [krò], Dzala [d̥oʔ], Kurtöp [d̥ò:], Bumthap [grok], Khengkha [gro], Chali [d̥oʔ], Phobjip [d̥o] (Hyslop 2014: 169), Trung [k̥uʔ55], Atong [korok] (Gerber & Grollmann 2025: 12), Written Burmese *khrok*, Middle Chinese *ljuwk* < Old Chinese *k.ruk (Hill 2019: 61). This problem requires a separate in-depth study.

31. The same distinction might have been made in 16th century Minyak Rabgang, but the data are ambiguous (see Footnote 25), sometimes showing [r] with labials, sometimes retroflex affricates instead (see Nishida 1963: 148ff.).

32. See Suzuki (2012a) and, most recently, Suzuki (2024a).

33. Groups (b)–(e) and Phongpa are part of Semkyi-Nyida (Suzuki 2024a).

Group (a) has preserved the most remarkable for the most complex onset [mbr-], which also forms the basis for reconstructing Proto-Tibetic *mbras.³⁴ Both Sprosnang and Phongpa display a reduction in rime characteristic of the neighbouring dialects, so that this feature might have resulted from language contact. Group (b) most probably underwent changes analogous to those of LOT-N, merging reflexes of Proto-Tibetic *mbr- with those of *mb-, but we lack sufficient data to test this hypothesis. Groups (c)–(e) underwent the most radical changes that can be sketched as follows:



Divergent vowel reflexes suggest that the proto-language had [-as] coda that was subsequently simplified when single sub-groups started to form out of AOT-NE. The reduction of the rime to a front vowel must have occurred before *mbr- > [ʱj-] so that it could trigger the latter change. As indicated by Suzuki (2012a: 113f.), some of the dialects that have [ʱg-] for OLT *ybras* seem to have merged equivalents of Proto-Tibetic *mbr- with [bgj-].

4.2 Early Old Tibetan dialect groups

EOT stands out by three sound changes:

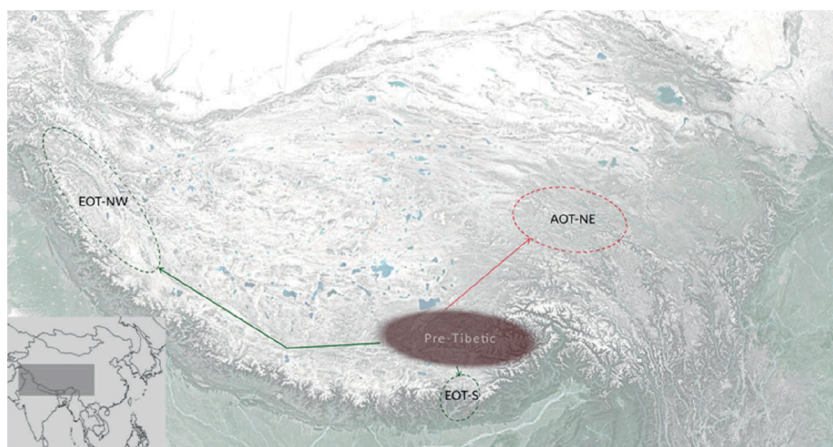
- Palatalisation: [h] > [j] / _# _V^[+front]
- Metathesis: [sL-] > [Ls-]
- D-excrecence: [zr-] > [zdr-] & [sr-] > [str-]³⁵

Metathesis and D-excrecence must have come into being after the script introduction in the 630s. In addition, EOT continued the assimilation of prefixes to the root consonant in the feature [voice] from the previous phase. Since these innovations, but not those of MOT, are shared by EOT-NW and EOT-S, I conclude that the groups split from EOT. Map 1 illustrates the first splits from the core language

34. Phongpa and Sprosnang pronunciations support Miller’s *mbras (1955: 286) quoted after *Yi-series Xifānguān Yiyǔ* (see also Footnote 25 *supra*). This reading can be dated as late as the 15th century. However, the transcriptions *gro for Literary Tibetan *gro* ‘wheat’, *sbrul for Literary Tibetan *sbrul* ‘snake’, and *mrgja for Literary Tibetan *brgyad* ‘eight’ (Miller 1955: 290) suggest an attempt at a transliteration rather than transcription. The status of the data and their phonetic interpretation remain, therefore, uncertain.

35. Bialek (2018b: 34) discussed the last two changes.

of Central Tibet; the splits accompanied the military expansion of the Tibetan Empire and the ensuing mass migrations.



Map 1. Tibetan dialects in the early second half of the 7th century (Image Landsat/Copernicus 01.2021)

4.2.1 EOT-NW *bras

This group is defined by one shared innovation:

- Proto-Tibetic *e > EOT-NW [ja] / C_C^[+velar]³⁶

The group has lost preconsonantal [m-] in its reflexes of Proto-Tibetic *mbras but has retained the rime [-as]. Some of the most eastern varieties innovated retroflex [ɖ] in onset: [br-] > [ɖ] / _#-. This change is a subsequent phase in the general merger of the equivalents of OLT *dr-*, *gr-*, and *br-*; [ɖ] has spread to EOT-NW most probably from EMT-W.³⁷ Forms such as Balti, Kargil, Tshangra [drul] ‘to walk’ (Bielmeier et al. 2018: 256) for Classical Tibetan *γgrul* (OLT *γdrul*) confirm that EOT-NW did not participate in the merger [gr-] + [dr-] of MOT.

The EOT-NW form [bras] was early borrowed into other neighbouring languages, for we find its traces in: Kinnauri [brass] ‘an inferior grain’ (Bailey 1911: 321a), Rongpo [bhəs] ‘a type of grain out of which a loaf is made which is considered sacred’ (Sharma 2001: 237a),³⁸ and Burushaski [bras] ‘rice’ (Lorimer

36. For this sound change, see Jacques (2009).

37. The same group of dialects also simplified the inherited [ldz-] to [ld-] in equivalents of OLT *zl-* (Bialek 2018b: Appendix B).

38. Modern Rongpo does not permit the consonant cluster [br] in onset (Zoller 1983: 14), but this innovation might postdate the borrowing.

1938:86a; [bras] in Berger 1998:180).³⁹ Since direct contacts between the Tibetans and the Burushaski people (OLT *bru źa*) were most intensive during the Tibetan Empire (especially in the early 8th century), we may assume that the word was also borrowed in that early period.⁴⁰ Modern Burushaski allows cluster [mb] only word-internally (Berger 1998:23f.) so that this borrowing cannot be used to date the loss of preconsonantal [m-] in EOT-NW. Nowadays, Kinnauri and Rongpo are spoken closer to the area of the EMT-W group. Still, the form [brass] indicates that the word must have been borrowed before the arrival of EMT-W speakers in this region, i.e. before the early 10th century.

4.2.2 EOT-S *bras

I provisionally group Choca-ngaca and Brokpa together for purely heuristic reasons as they belong to one geographic area of Eastern Bhutan. Although not enough data is available on Choca-ngaca to evaluate its onsets equivalent to OLT *sl-*, *zl-*, and *sr-*, it is evident that the language did not participate in devoicing of plain consonants in onset – an innovation of MOT. Likewise, Brokpa retains the voiced consonants (Wangdi 2021:101ff.) and so must have split off from the common stock before devoicing occurred. It is feasible that two other Tibetic minority languages of Bhutan, Lakha and Brokkat, developed from EOT-S too. However, today, they display much more advanced changes in their reflexes of Proto-Tibetic, most probably due to the long-standing influence from Dzongkha and Bumthang, respectively (van Driem 2001:2.866f.)

In their reflexes of Proto-Tibetic *mbras, Choca-ngaca and Brokpa, as the only languages of the region, have retained the complex onset [br-] while losing the pre-consonantal [m-]. However, the languages show distinct developments in rime; Choca-ngaca lengthened the vowel in response to the loss of coda [-s], whereas Brokpa reduced the rime, fronting the vowel. Choca-ngaca and Brokpa seem to have been isolated from other Tibetic languages for quite a long time, and their reflexes of Proto-Tibetic *mbras might have received influences from the surrounding non-Tibetic languages. For instance, the reduction of [-as] to [-e] occurred in Dzala and Dakpa.⁴¹ In general, however, data available on Choca-ngaca and Brokpa is very scanty and partly contradictory.⁴²

39. The word was later borrowed into Domaaki (Indo-Aryan language spoken in the same region as Burushaski) as bräs (Kuiper 1966:61).

40. The contact between the groups certainly continued in later times, but it was in 737/8 that the Tibetan army defeated Burushaski, and their king became a vassal of *bcan po* Khri Lde-gug-rcañ (cf. IOL Tib J 750:277–8).

41. For data, see Hyslop (2014:159); Hyslop & d'Alpoim-Guedes (2021:575, Figure 3).

Since the cluster [mbr-] has been preserved in MOT (see below), the loss of [m-] in EOT-NW and EOT-S was an independent change, a case of parallel development, without further consequences for the subgrouping. The change must have occurred after the groups had separated from EOT, their last common ancestor, and should be considered innovation occurring independently in both EOT-NW and EOT-S.

4.3 Middle Old Tibetan dialect group

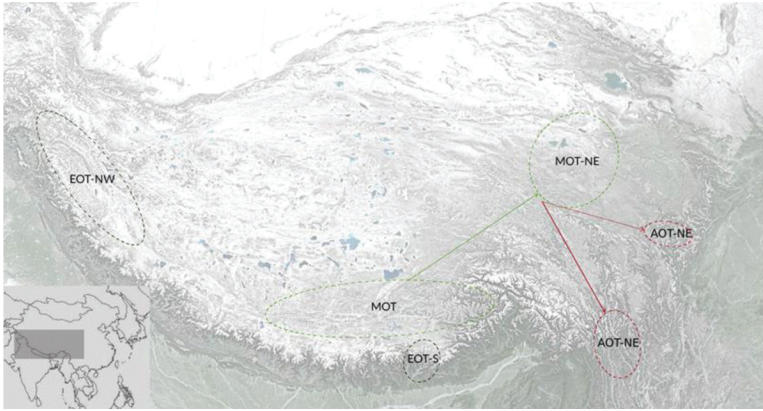
MOT started after the separation of EOT-NW and EOT-S from the common stock and witnessed three further sound changes:

- Devoicing of plain obstruents in onset
- Merger [dr-] + [gr-]
- Merger [sdr-] + [sgr-]⁴³

The lack of data makes it unclear whether EOT-S could still have participated in the last change. Since Brokpa possesses reflexes of the merger [Tr-]+[Kr-] > [d] that must have preceded [sdr-] > [sgr-], at least the ancestor of this dialect might have separated a bit later during the EOT but still before the devoicing of plain onsets. The merger [Tr-] + [Kr-] must have occurred after (i) the separation of AOT-NE and EOT-NW, (ii) the borrowing of the demonym *dru gu* (see Bialek 2026: 1002), and (iii) the devoicing of [dr-] (otherwise the demonym could have been spelt as *'tru gu*, more in accordance with its source).

42. For Choca-ngaca, Tournadre & Rigzin (2015) remains our sole source. Brokpa has received more attention in recent years. Still, the analysis of its phonological system encounters serious problems with huge divergences even in the phoneme inventory: 27 consonants and 5 vowels were distinguished by Dondrup (1993:1), 32 + 5 by Bodt (2012:309), 35 + 11 by Funk (2020:23), and 39 + 8 by Wangdi (2020:396). The studies of Funk (2020) and Rüfenacht & Waldis (2020) are based only on one informant in his 20's. The young age of the informant may explain the lack of certain archaic features in his pronunciation as compared with data provided by Wangdi (2020; 2021). The data at hand suggest considerable influence on Brokpa first of all from the *čhos skad*, the liturgical language of Tibetan Buddhism. The impact is most visible in the random ascriptions of tone, as if arbitrary and not resulting from internal historical development.

43. See Bialek (2018b: 27f.)



Map 2. Tibetan dialect proto-groups at the turn of the 7th and 8th century (Image Landsat / Copernicus 01.2021)

All Tibetic dialects (besides those that directly evolved from AOT and EOT) lack voiced consonants in their reflexes of Proto-Tibetic simple voiced plosives in onset. I take this as an indication that their proto-languages formed after the devoicing in MOT. The first group to separate after the said three sound changes was MOT-NE following the conquest of the Υ a- \acute{z} a in the early second half of the 7th century. Map 2 illustrates the linguistic situation on the Tibetan Plateau after the split of MOT-NE.

4.3.1 MOT-NE *mbras

MOT-NE introduced rhotacism:

- Proto-Tibetic [l] > MOT-NE [r]/_#-C⁴⁴

Although the vast majority of dialects in this group attest to retroflexion of post-consonantal [br] in their equivalents of OLT *ybras*, we find two intriguing exceptions in Golok and Sastod: [mdɿi:] and /ⁿbri/, respectively. Following Sprigg's detailed survey (1968: 309f.), the given Golok form should be interpreted as [mq̄zi:], but Suzuki's data (p.c. 2024-06-13) show the word to be pronounced with the onset [mq̄r-], along with [mq̄-] and [mq̄z-]. Although Sun does not discuss his transcription, Sastod /ⁿbri/ for Literary Tibetan *ybras* can be juxtaposed with [χt̚ə] for Literary Tibetan *dpri* 'colostrum' (Sun 2025: 731), showing that he uses a different convention for retroflex affricates (see *ibid.*, p. 706: Example (4)).

The inclusion of Sastod in MOT-NE might be deemed controversial because the dialect does not share many of the group's specific sound changes (see Sun

44. Bialek (2018b: 6).

2025: §4), but on the other hand, it certainly underwent rhotacism (LC- > RC-)⁴⁵ not seen, for instance, among the neighbouring languages that descended from LOT-E (see *infra*). The preservation of the complex [br-] in Sastod confirms that the proto-group did not participate in the merger [Kr-] + [Pr-]. Generally, Sastod shows distinct equivalents of OLT preconsonantal *ɣ*- and *m*-, except for the merger before velars and labials. A possible solution would conceive Sastod as separated from the core of the MOT-NE right after the rhotacism but before all the other changes. The merger [N-] + [m-] might have been triggered by the contact with LOT-E. It is conceivable that the migration of LOT-E speakers to the region was the direct cause for the separation of Sastod from MOT-NE early in its history (see Map 3).

The Sastod data allow us to reconstruct the onset *mbr- to MOT-NE. Forms such as Rkangtsha [mɔ̃zɪ], Themchen [mɔ̃zwi], Arik [mɔ̃zɪ], Mkharmar [mɔ̃zwi], and Lhagang [mɔ̃ɖe] confirm the temporal precedence of retroflexion over the assimilation in this group. Sastod prenasalisation is homorganic, so Sun's (2025) /^mbri/ represents [m^bbri].

MOT-NE dialects attest to the regular sound change Proto-Tibetic *[-as] > [-i] that is also seen in Table 2. However, dialects with [-e] may suggest that MOT-NE inherited *mbras from MOT (see also §5).

4.4 Late Old Tibetan dialect groups

Sound changes that characterise LOT include:

- Loss of preconsonantal *g*-, *b*-, and *r*-⁴⁶
- Merger [N-] + [m-]

The situation of dialects and dialect groups derived from LOT is complex because many of them remained very innovative in later periods, for instance, by further reducing the structure of their syllables and developing tones. The most difficult to disentangle is the situation in the historical region known as Khams, where we have to reckon with several waves of intra- and extra-regional migration in the past millennium or so. Despite the tremendous progress achieved in recent years, mainly due to the painstaking work of Hiroyuki Suzuki, the vast majority of the dialects spoken in this region are still poorly documented, so the conclusions reached below must be treated as preliminary.

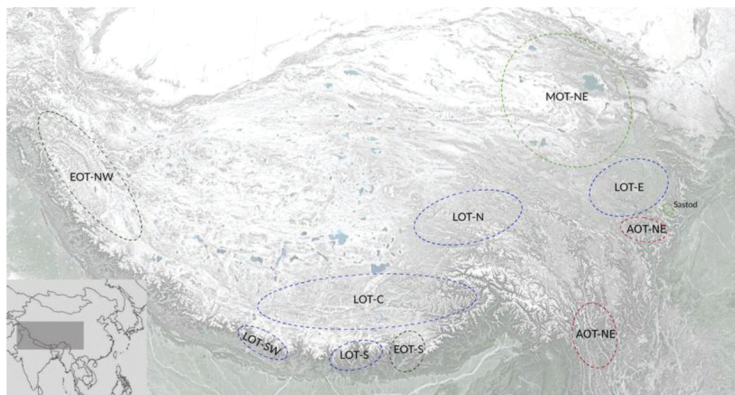
45. Bialek (2018b:6). Sastod has [sl] in equivalents of Literary Tibetan *slob* (Sun 2025:732, s.v. 'disciple'), but this I deem to be a reading pronunciation.

46. Bialek (2018b:35).

Although the dialects of the sub-groups LOT-SW, LOT-S, and LOT-N have not preserved any traces of the preconsonantal *m- in their reflexes of *mbras (exception: Gertse), a nasal consonant has to be reconstructed to their proto-languages for it prevented devoicing that otherwise affected the bare onset [br-] (cf., e.g. modern equivalents of OLT *brag* 'rock' in Biemeier et al. 2013:5723). The loss of *m- in these groups was not an innovation of a common ancestor but rather a change occurring independently in each group, as suggested by their distinct reflexes of Proto-Tibetic *zr-, *sr-, *zl-, and *sl- (Bialek 2018b). OLT orthography indicates that the merger [N-] + [m-] might have first occurred in onset clusters with labial plosives (see §5). Accordingly, I reconstruct LOT *^Nbras.

The data collected in Table 2 suggest yet another sound change: the loss of the final [-s] accompanied by a vowel change. However, since exceptions are also found (see LOT-E, or Jirel and Sherpa) I consider it to be an independent change introduced after the respective proto-languages had separated from LOT.

Map 3 presents the reconstructed situation on the Tibetan Plateau after all the hitherto identified LOT dialect groups separated from the core. The formation of single groups cannot be dated, but it seems plausible that it occurred before the emergence of the tone system in Central Tibet at the turn of the 8th and 9th centuries.



Map 3. Tibetan dialect proto-groups at the turn of the 8th and 9th century (Image Landsat / Copernicus 01.2021)

4.4.1 LOT-SW *^Nbrɛ:

This group did not participate in the retroflexion of [br-] shared in the area; cf. LOT-C Jirel [tʰak] and Southern Mustang [tak] ≠ OLT *brag* 'rock' (Biemeier et al. 2013:5723). Biemeier (1982; 1985:16f.) noticed certain archaic phonetic

features of Kyirong that distinguish it from Central Tibetan varieties. In addition, as the only dialect Kyirong has a low tone equivalent of Literary Tibetan *sre moñ* (≠ OLT *zre mo*; Bialek 2018b: 21), a feature that confirms that the language (or rather LOT-SW) went its way, distinct from other dialects of the area. These primary findings receive support from the recent Bayesian phylogenetic analysis (Dhakal et al. 2024) that shows a closer historical relationship between Yolmo, Kagate, and Tsum as against other surveyed languages of the region (Northern Mustang, Nubri, Gyalsumdo, Jirel, Sherpa) which I subsume under LOT-C (see *infra*).

4.4.2 LOT-E *^Ndʒe:

This group delivers a relatively simple picture in its equivalents of LOT *^Nbras. It shares the merger [ʰ-] + [m-] postulated for LOT but adds retroflexion of the cluster [br-]. LOT-E merged all the clusters [Cr-], partly even with [Cj-].⁴⁷ There is not enough data published on Thewo Mä to confirm that the dialect deaffricated all reflexes of *[Cr-], but I consider it a relatively recent change.⁴⁸ Because of its geographical position, LOT-E might have separated before LOT-N and was moved to its current location by the migrating LOT-N speakers and then by Amdo groups migrating southward into Khams several centuries later.

4.4.3 LOT-N *^Nbas

Four dialects of this group (Hor, Amdo, Nakchu, Bachen) lost the post-consonantal [-r-] after the merger with the equivalents of OLT *ɣb-* [ʰb-]/[mb-]; cf. Gertse, Nakchu [bɑʔ], and Amdo [bɑk] ≠ OLT *ɣbag* ‘mask’ (Bielmeier et al. 2013: 5872). Although the data in Bielmeier et al. (2013) does not display nasalisation, Hiroyuki Suzuki confirmed (p.c. 2024-06-11) that all these words have prenasalised onset. The changes in the rime allow us to reconstruct *-as with the loss of the coda and lengthening in Bachen, and fronting in the remaining dialects. Accordingly, I reconstruct *^Nbas to LOT-N. Gertse [ndʒe:] does not fit well with the remaining equivalents. However, the form is a borrowing from an EMT-E

47. The fact that Thewo (as all the dialects descending from LOT-E) participated in sound changes common to LOT (see Table 1) falsifies the hypothesis of its pre-Old Tibetan origins (Yang et al. 2024). In addition, Thewo has lost *v3* stems, its equivalents of OLT *sl-*, *zl-*, and *sr-* onsets parallel those of MOT-NE and LOT-E (Bialek 2021c: xx), and it requires auxiliaries to express TAME (Yang et al. 2024: 718). These are likewise features that locate its proto-language in the late 8th century at the earliest. For a critical evaluation of the “voicing rule” as recently postulated by Yang et al., see Bialek & Hiroyuki (Forthcoming).

48. On deaffrication in Tibetic, see LOT-C *infra*.

dialect (see *infra*) after groups migrating from Yul-śul (modern southern Qinghai) settled in Mñay-ris (Tournadre & Suzuki 2023: 481).

4.4.4 LOT-S *^Nbjas

This group merged the Proto-Tibetic *br- reflexes with *bj-, with subsequent palatalisation in Dzongkha; cf. Dzongkha [bdz̥aŋ] ≠ OLT *ybyañ* ‘to be cleansed’ (Bielmeier et al. 2018: 892). The lengthening of the vowel compensates for the loss of [-s]. Interestingly, Drenjong has two equivalents of OLT *ybras* that go back to two different sources: [bja] is a reflex of the inherited *^Nbjas, but [re:] must have been borrowed from a dialect that had *br̥e:. The borrowing occurred after *br- > *bj- in LOT-S, so the borrowed cluster *br- was simplified to [r-].

4.4.5 LOT-C *^Ndzas

This is another highly divergent group of tonal dialects with four distinct reflexes for Proto-Tibetic *mbras in onset:

- a. [nʈs-] Gar, Ruthog, Purang, Gergye, Tshochen
- b. [ʈs-] Lhasa, Shigatse
- c. [t-] Dingri, Nubri, Lhomi
- d. [d-] Humla, Southern Mustang, Jirel, Sherpa

Groups (a)–(b) pattern together geographically (Mñay-ris vs. Dbus-Gcañ) and linguistically in their equivalents of OLT *br-*. The remaining two groups intersect geographically because they are both spoken in eastern Nepal’s border area, north and south of the current Nepal-China frontier. Although we often read that Sherpas migrated in several waves from Khams to the Solukhumbu area in Eastern Nepal first in the 16th–18th century,⁴⁹ their reflex of *mbr- suggests the local origin of the word for ‘rice’. Voiced [d-] can only be accounted for if we reconstruct *^Ndz̥- to LOT-C, an onset presumably inherited also by EMT-W speakers (see next section). The subsequent changes were twofold:

- I. *^Ndz̥- > [nʈs-] devoicing that accompanied the development of a full-fledged tonal system > [ʈs-] loss of preconsonantal nasal [ⁿ-] > [t-] deaffrication
- II. *^Ndz̥- > *^Nd- deaffrication > [d-] loss of preconsonantal nasal [ⁿ-]

49. Tournadre et al. (2009: 19). To my knowledge, no linguistic evidence has ever been presented to substantiate this hypothesis. Sherpa does not display any features characteristic of Khams dialects, and the recently published Bayesian phylogenetic analysis carried out on several Tibetan dialects from Nepal clearly shows that Sherpa is most closely related to Jirel (Dhakal et al. 2024).

Deaffrication is the most problematic stage in both chains because it is typologically a relatively uncommon process.⁵⁰ The fortition affricate > plosive is not unknown in Tibetic languages as dialects of three distinct dialect groups attest synchronically to alternation between affricate [dz-] and plosive [d-] in their reflexes of Proto-Tibetic *zl-: Hor (LOT-N), Nangchen (EMT-E), and Golok, Themchen, Arik, Mkharmar, Rkangtsha, Rngaba, Bayan, and Ndzorge (MOT-N; Bialek 2018b: Appendix B). But the situation is more complicated because affricates inherited from Proto-Tibetic and expressed in the script by letters *č, čh, ĵ, c, ch*, and *j* can have fricative but not plosive reflexes in modern Tibetan dialects (Liu 2024: 53). When analysing the complexity of the processes Liu did not take into account the verb stem alternation of Tibetan conjugations. If we leave out verb alternations and irregular patterns displayed in isolated nominal lexemes, we notice that fricative equivalents of Literary Tibetan affricates are rare in modern languages and concern mainly voiced consonants:⁵¹

- OLT *mĵ-* and *yĵ-* have fricative equivalents in several lexical items in EOT-NW
- Labrang and Chabcha (MOT-NE) regularly have [z] for Literary Tibetan *rĵ-/lj-* and [z] for *rj-*
- Deaffrication to fricatives is an ongoing process in EMT-W, most advanced in equivalents of OLT *gc-*, *rc-*, and *rj-*, spreading at times to Leh (EOT-N) in the last case

A certain degree of regularity is encountered only in dialects descending from MOT-NE and EMT-W. These plus LOT-C are also the dialect groups with deaffrication to plosives in their equivalents of *ybras*.⁵² The processes, however, concern two historically distinct sounds; deaffrication to fricatives occurs on postconsonantal alveolar and palatal affricates inherited from Proto-Tibetic, whereas deaffrication to plosives takes place with secondary retroflex affricates that emerged from the consonant clusters [Cr-] and are found either in absolute onset (_#Cr-) or in a postconsonantal position (_#CCr-):

C^[+affricate] inherited: _#CC^[+affricate]- > _#CC^[+fricative]-
 C^[+affricate] < Cr: _#(C)C^[+affricate]- > _#(C)C^[+plosive]-

The exact motivation behind this divergent behaviour requires further study. Still, it might be that retroflex affricates are realised with acoustic features that make

50. Kümmel (2007:96ff.) provides examples of strengthening affricate > plosive from Indo-European, Afro-Asiatic, and Uralic language families.

51. The survey is based on Bielmeier et al. (2013:2266–2873, 6525–7076).

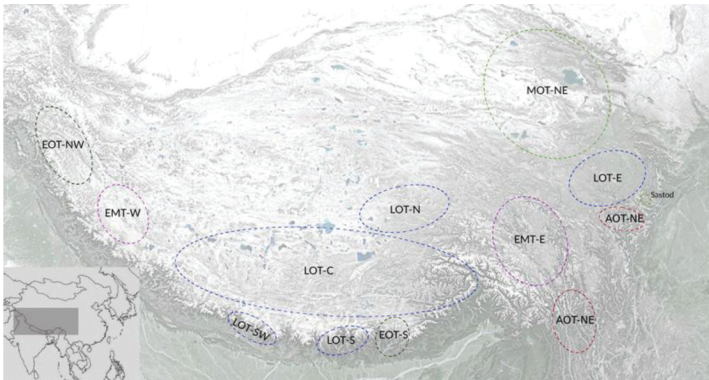
52. Deaffrication to plosives in equivalents of Literary Tibetan *ybr-* happens in dialects distinct from those where we find the same process in reflexes of Proto-Tibetic *[zl-]. That the processes occur independently can be explained by the fact that the latter seems to be a case of merger [ldz-] + [ld-] > [ld-] as argued by Liu (2024:61ff.)

them resemble a sequence of a plosive and a fricative rather than one unitary segment.⁵³ Since cross-linguistically the fortition affricate > plosive takes place most commonly in absolute onset, we can hypothesise that the sound change in Tibetic started in this environment and later spread to the postconsonantal position.⁵⁴

In their reflexes of Proto-Tibetic *mbras, Jirel and Sherpa attest to vowel [a] as against fronted [e] and [ɛ] of the remaining dialects. On this basis, I assume that the loss of [-s] and fronting of the vowel postdate the LOT-C stage.

4.5 Early Middle Tibetan dialect groups

Proto-groups stemming from Central Tibetan dialects in EMT must have inherited some basic tones. For now, only two such groups have been identified: EMT-W and EMT-E. Map 4 approximates the linguistic situation on the Tibetan Plateau shortly after the fall of the Tibetan Empire.



Map 4. Tibetan dialects in the early 10th century (Image Landsat/Copernicus 01.2021)

4.5.1 EMT-W *^Nd̥zas

EMT-W must have split off approximately at the turn of the 9th and 10th century. The data on Trashitongde is scanty, amounting to roughly three dozen words (see Bielmeier et al. 2013: D-351); therefore, I leave aside this dialect. The remaining dialects seem to have undergone the development of the onset parallel to the second path of LOT-C: [^Nd̥z̥] > *^Nd̥-, with the important distinction that Tholing

53. It is, however, true that some discrepancies may be due to diverging strategies of scholars in transcribing the sounds.

54. Kümmel (2007: 97f.) lists examples of deaffrication in postconsonantal position only after a fricative.

has preserved the onset [^Ndʑ-]. The same dialect also attests to vowel [a], indicating that the loss of final [-s] and fronting of the vowel postdate the formation of EMT-W. This allows us to reconstruct the EMT-W form as *^Ndʑas. Since the formation of EMT-W was most probably related to the migration from Central Tibet towards the west and the foundation of the first Western Tibetan kingdom by Skyid-lde Ńi-ma-mgon in the early 10th century (Bialek 2018b: 31, Footnote 79), the form *^Ndʑas can likewise be dated to this period.

4.5.2 EMT-E *ndʑē:

Despite their spread over a relatively large area, the majority of the dialects of this group have the same reflex of Proto-Tibetic *mbras: [ndʑē:]. The remaining dialects on which I have data (Sangdam, Dgudzong, Khyungpo) attest to a simplified onset. Whether EMT-E, splitting off from EMT later than EMT-W, inherited *ndʑē: from its parent language or innovated it later cannot be ascertained based on the data available at the moment. What strikes mostly is the high tone in some of the dialects, which accords with the general pattern for the equivalents of OLT *ybr-* in this group.

5. Proto-Tibetic *mbras

The data quoted from OLT confirm two basic meanings of *ybras*: (1) ‘crop; grain’, and (2) ‘rice’, which receive strong support from modern dialects. Suzuki & Wangmo (2016: 75) proposed that *ybras* ‘rice’ spread among Tibetan dialects as a cultural word. However, it is clear from the present investigation that the word has been inherited from the common ancestor language. The word might have been preserved in the dialects due to the ritual significance of rice, which was also a substitute for other white objects that could not be delivered (ibid.) This ritual function of rice must predate Buddhism since dialect formation started already in the early 7th century. Rice could have risen to the status of a food staple in the unique socio-historical environment of the Tibetan Empire. As Examples (3) and (4) have demonstrated, rice was an important crop that gave way only to barley in this period. Rice is a labour intense cultivator and involves more manpower than barley (Scott 2009: 64ff.), so it might have become a popular grain only with enough human labourers available for its cultivation. The additional manpower was supplied by peoples subjugated during military expeditions. With the fall of the Tibetan Empire and the decline of its economy, rice could not be cultivated

on the same scale, so it became a minor crop.⁵⁵ Finally, due to climatic conditions on the Tibetan Plateau, rice can be cultivated only on its southern fringes, which is the cradle of the Tibetan civilisation. However, the expansion of the Empire has brought speakers of Tibetic languages to all corners of the Plateau, where rice, as a food staple, had to be replaced by another cereal grain.

In terms of its etymological form, the word can be reconstructed as *mbras for Proto-Tibetic; the complex onset [mbr-] is still attested in Sprosnang and Phongpa (AOT-NE),⁵⁶ whereas descendants of EOT-NW have preserved the rime [-as].

The reconstructed form *mbras is additionally buttressed by a few, apparently very early borrowings into Rgyalrongic: Munashan Chuchen [mbras], Sengge Btsanlha [mbras], Maerkang Situ [mbrəs], Japhug [mbrɤz].⁵⁷ Further interesting forms borrowed from Tibetic include: West Rgyalrongic Stau and Geshiza [mbrɛ] (Suzuki & Wangmo 2016: 73, Table 1; Honkasalo 2019: 188, Table 3.41), Wobzi [mbræ] (Lai 2016: 150), and Qiangic nDrapa [mbzɛ¹³] (Huang et al. 1992: 0439.14). All these suggest that the original Tibetan form transmitted to the region was *mbras, and it must have been brought there by either AOT-NE or MOT-NE speakers.⁵⁸

One important question remains to be addressed: If Proto-Tibetic had *mbras and this form was transmitted up to MOT, why do all our OLT records agree on the orthography རྩལ་ *ybras* instead of the expected རྩལ་ *mbras*? Not only in this case but in general, the prescribed *m*- letter is not allowed with labial plosives, so that **mp*-, **mph*-, and **mb*- are prohibited in the Literary Tibetan orthography from its earliest witnesses. The working hypothesis presumes that by the time our oldest preserved documents have been composed (i.e. by the second half of the 8th century), the dialects of the Tibetan administrative centres and chancelleries

55. Another factor is the climate change in the first half of the 9th century (Beckwith 2009: 158ff.), which might have negatively impacted rice cultivation in large portions of southern Tibetan Plateau.

56. Beyer (1993: 24, Footnote 23) quoted [mbri] as a modern pronunciation in Čhab-mdo but without giving a source for this form.

57. Data quoted after Suzuki & Wangmo (2016: 73, Table 1), Jacques (2021: 1516), and STEDT.

58. Jacques (2016: 176f.) argued that Proto-Rgyalrongic had an epenthetic [-b-] inserted between [m-] and [-r-]: *mr- > *mbr-. Since his reconstruction is based on one lexeme, in the light of the Tibetic data presented here, an early borrowing from Tibetic seems a more plausible explanation. Whereas Japhug retained the complex onset [mbr-], Proto-Tangut would have turned [-r-] to [-j-]: *mbr- > *mbj-. There is no reason to assume that the innovation *mr- > mbr- occurred twice in the history of the Trans-Himalayan languages, especially as the only languages that attest to a bilabial plosive in reflexes of PTH *mr- are those with a well-documented history of contacts with Tibetic.

merged reflexes of [m-] and [N-], starting in pre-labial contexts. This change, initiated as it seems in LOT, was perpetuated in the Literary Tibetan orthography of our documents.

6. Cognates and Proto-Bodic *mras

Further support for the reconstructed Proto-Tibetic *mbras can be adduced from other presumably cognate Trans-Himalayan languages. Table 3 provides forms tentatively identified as cognate with Proto-Tibetic *mbras. Caution is required, for it is evident that no firm sound correspondences can be established by comparing one word. The following analysis sets the stage for a more thorough study (see Bialek Forthcoming b).

Table 3. Cognates of Proto-Tibetic *mbras

	Language	Dialect	Form	Source
EB	Kurtöp		mras	Hyslop & d'Alpoim-Guedes (2021: 575, Figure 3)
			^L mra:ʔ	Michailovsky & Mazaudon (1994: 548)
	Bumthap		mras/mrat	Hyslop & d'Alpoim-Guedes (2021: 575, Figure 3)
		Ura	mras	Dzongkha Development Commission (2018: 65)
		Choekhor	^L mrat	Michailovsky & Mazaudon (1994: 548)
		Chume	^L mras	
Khengkha		mras	Hyslop & d'Alpoim-Guedes (2021: 575, Figure 3)	
KGT	Tamang		mrasi	Owen-Smith (2015: 308)

OLT *ybras* was already related to the East Bodish (EB) lexemes by Michailovsky & Mazaudon (1994: 548). In Bumthap, the final [s] shows a tendency to be replaced by [t], especially in the Choekhor dialect (Michailovsky & Mazaudon 1994: 548):⁵⁹ Kurtöp [nas], Bumthap [nat] ‘barley’ ≠ OLT *nas*. Kurtöp, Bumthap, and Khengkha form a sub-group within East Bodish languages (Hyslop 2014: 158, Figure 2). Their last common ancestor inherited *mras from PEB. The remaining East Bodish languages replaced their inherited PEB *mras by borrowings from other languages.⁶⁰

59. Four main dialects are distinguished in Bumthap: Chume, Choekhor, Ura (Michailovsky & Mazaudon 1994: 545), and Tang (van Driem 2015: 20).

60. Dakpa and Dzala [dep] (Hyslop & d'Alpoim-Guedes 2021: 575, Figure 3; Dzongkha Development Commission 2017: 38) might go back to OLT *ybras bu* ‘fruit’ (< *‘harvest’?), for

Tamangic [mra:si] 'rice grain' (Owen-Smith 2015: 308) is the only occurrence of the word in literature on this group of Trans-Himalayan languages. No other sources know of it or its possible variants. Unfortunately, Owen-Smith does not discuss the word's etymology so that the source of the final -i remains unknown. Owen-Smith's glossing indicates that it belongs to the root. A disyllabic word suggests a compound, maybe *mras-si 'rice-fruit', with -si related to the popular word si/se 'fruit' (see STEDT #1019).

Worth considering in this context is the inclusion of Bjokapakha [mrat] 'pimples' (Grollmann 2020:485a), Central Tsangla [mrash] 'pimples', and Motuo Tsangla [mras] 'spot' (*apud* STEDT) as potential cognates. Bjokapakha does not permit [-s] in coda (Grollmann 2020: 45f.), which may account for its substitution by [-t]. Both meanings 'grain' and 'pustule' are housed, for instance, by Literary Tibetan *ybrum* (cf. Jäschke 1881:401b), but one particular passage in an Old Tibetan text suggests that *ybras* might indeed have been used to denote a surface change in skin condition. In (6), *bya ybras* is coordinated with *chag ma* 'dark spots'; both are apparently skin diseases typical of animals carrying loads on their backs. Thus, *ybras* could have had the meaning very much resembling Tshangla 'pimples'.

- (6) *bžon=gyi myi* *myed=kyañ* *skar_ma* *gyen=gyis* *bžon=te*
 ride=GEN human(ABS) NEG.EXIST=ADD star division=ERG ride=GER
rgyab *bya+ybras=dañ* *chag+ma/* *chīg=par*
 back(ABS) bird+grain=COM dark_spots+NMLZ(ABS) burn=NMLZ:TERM
śog=śig/
 AUX.IMP=IMP
 'Even though there is no human to ride [you], being ridden by stars' divisions
 may *bya ybras* and dark spots burn [on your] back!' [IOL Tib J 731: r92–3]

which compare Kurtöp drebu 'effect, fruit, outcome, result', indeed a borrowing from a Central Tibetan dialect (cf. Biemeier et al. 2013:5971). The same lexical unit may be suspected behind drä in Dzala dräsi 'scented cooked rice served in occasions (sic)' (Dzongkha Development Commission 2017:59), historically identical with Bumthap drä se 'cooked rice mixed with sugar, dried grapes, and saffron coloured with turmeric powder served to the guest during specials occasions' (Dzongkha *ybras ser/sil*; Dzongkha Development Commission 2018:61); this informative description of Bumthap drä se with its focus on the yellowish colour of the dish indicates that the word is a loan from Tibetic **ybras ser*, lit. 'yellow rice'. Chali [təaŋbu] is another loanword from Tibetic; cf. OLT *ybras gi lč(h)añ pa* (Pelliot tibétain 1134: 117, 182; Pelliot tibétain 1136: 27; Pelliot tibétain 1194: 61; IOL Tib J 731: r45) 'seedling of rice' and Kardze [tchō:pā] 'seedling' (Biemeier et al. 2013:2501). The source of Hengkhe [sem] is apparently Dzongkha [səm] 'lentils, beans, dried beans, small beans' (Biemeier et al. 2013:8908), Literary Tibetan *sran ma* 'pease, beans, lentils' (Jäschke 1881:580b).

Since Trans-Himalayan languages spoken near Tshangla attest to the meaning ‘rice’ for their reflexes of Proto-Bodic **mras*, it seems that PTsh **mras* was inherited from the last common ancestor of Proto-Bodic and PTsh. This argument is strengthened by Tshangla [mrok] ‘open grazing patch in the forest’ (Bodt 2023: 116), an obvious cognate of OLT *ybrog* ‘summer pasture’ (Old Tibetan Dictionary: <http://otdict.com>) < Proto-Bodic **mrög*. The latter etymon calls to mind Bunan *rwak* ‘to graze’ (Widmer 2017: 751b) and Darma *hro(k)* ‘to herd’ (Willis Oko 2019: 537), but I leave this discussion for another occasion (see Bialek Forthcoming b).

Proto-Tibetic **mbras*, PEB **mras*, PKGT **mras*, and PTsh **mras* provide a relatively straightforward picture, allowing for reconstructing their last common ancestor as **mras*. The Proto-Tibetic form is easily explained by recourse to the well-documented consonant insertion: **mras* > Proto-Tibetic **mbras* (see §7). The corresponding preliminary Stammbaum is diagrammed in Figure 1; Tables 4 and 5 list the established sound correspondences. Regarding the meaning of Proto-Bodic **mras*, the prevalence of the sense ‘rice’ in modern languages would suggest the semantic narrowing from the original ‘grain’ to, on the one hand, ‘rice’ and, on the other hand, ‘pustule’.

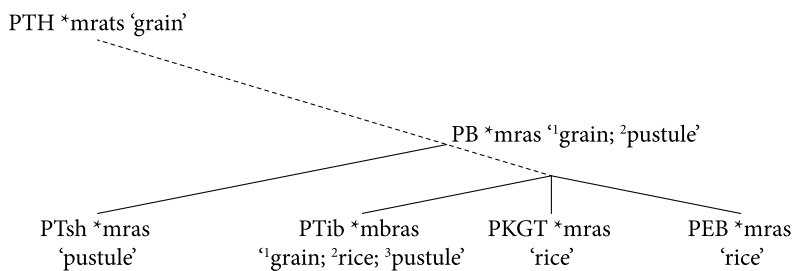


Figure 1. Reflexes of Proto-Bodic **mras*

Table 4. Onset correspondences for PTH **mr-*

PTH	PB	PTsh	PTib	PKGT	PEB
* <i>mr</i>	* <i>mr</i>	* <i>mr</i>	* <i>mbr</i>	* <i>mr</i>	* <i>mr</i>

Table 5. Rime correspondences for PTH **-ats*

PTH	PB	PTsh	PTib	PKGT	PEB
* <i>-ats</i>	* <i>-as</i>	* <i>-as</i>	* <i>-as</i>	* <i>-as</i>	* <i>-as</i>

Apart from the data quoted from the so-called Bodic branch of the Trans-Himalayan family, we find at least two other forms that might be cognate:

- Lhokpu *rat* ‘rice (paddy)’ (Grollmann & Gerber 2018:79, Table 32, no. 135)
- Chinese *li* 糲 ‘coarse rice’; Old Chinese *[r]ʰat (Baxter & Sagart 2014a), *rât~*râts~*rats (Schuessler 2009:236)

The comparison OLT *sbrul* (< Proto-Tibetic *s+mbrul) ‘snake’ ≠ Lhokpu *ri:pu* ‘big constrictor’ (Grollmann & Gerber 2018:77, Table 32, no. 191) confirms the loss of the preconsonantal *m-* in Lhokpu. Lhokpu *rat* and Old Chinese *rât(s) add to our reconstruction, allowing us to posit PTH *mrats (see Figure 1 and Tables 4–5). I tentatively reconstruct *mr- since nominal derivation with prefix [m-] has not been demonstrated for the languages included in the Bodic group, but the consonantal cluster [mr-] is well-attested in other branches of the Trans-Himalayan family and seems to have been inherited from PTH.⁶¹ The reconstructed form *mrats suggests the simplification of the final *-ts to *-s as a potential innovation of Proto-Bodic and would be the first strong confirmation for the inclusion of Tibetic, East Bodish, Tshanglic, and Kaike-Ghale-Tamangic in one sub-branch.⁶²

When queried for ‘rice’, STEDT and Lexibank⁶³ provide further words that could be potential cognates, but due to their highly simplified phonetic form (CV), they cannot yet be compared with the words studied in this section. In general, we can divide these lexemes into two groups: (i) [mV] (or even only [m]) and (ii) [rV], with the prevailing open central vowel in both types. If not representing distinct etymons, mV and rV could be outcomes of differentiated treatment of the original *mr- in distinct sub-branches of Trans-Himalayan.

7. Previous hypotheses

The etymology of Tibetic *ybras* was suggested several times in earlier works, although their authors usually restricted to listing its potential cognates without attempting a thorough historical explanation. Matisoff (2003:437) included *ybras* together with Chinese *li* 糲 (Old Chinese *lâd~l̥jad~lât)⁶⁴ and Mizo [raʔ] ‘fruit,

61. E.g. in Lepcha (Mainwaring & Grünwedel 1898:301; Plaisier 2007:31) and Tani (Sun 1993:116f.); see STEDT for other examples. Benedict (1972:37) did not posit *mr- for PTH, but Matisoff (2003:80) included *mr- in his reconstruction.

62. The correspondence between OLT -s and the Old Chinese *-ts cluster was first observed by Hill (2015:170).

63. <<https://lexibank.cld.org/parameters/2026#3/18.48/103.54>>; (Accessed 2024-05-27.)

bear fruit' under the PTB etymon *b-ras (see also STEDT: #2071 PTB *b-ras RICE/FRUIT/BEAR FRUIT/ROUND OBJECT).⁶⁵ Likewise Sagart (2003:129) related the word to Chinese *li* 糲 and Mizo [raʔ] but he reconstructed the Old Chinese form as *b^mmə-rat-s 'coarse (=unpolished) husked rice'. In addition, he reconstructed Proto-Tibetic *m-ras "with parasitic /b/ arising between /m/ and /r/", equating it also with his PST *m-ras (ibid., p.133, Table 1). Sagart's hypothesis included PST *m-ras and Proto-Austronesian *beRas 'husked rice' as cognates derived from one proto-language of his East-Asian superfamily (ibid., p.134).⁶⁶ In a later publication of Baxter & Sagart (2014a), Old Chinese loses the first syllable, becoming *[r]^hat that also accords with Schuessler's reconstructions *rats (2007:352) and *rât~*râts~*rats (2009:236). It is unclear whether Sagart still conceives this Old Chinese *[r]^hat as cognate with OLT *ybras*. Building upon Sagart's idea, Ferlus (2010:66) identified Chinese *li* 糲 (his Old Chinese *m.rats 'coarse grain, millet or rice') and Literary Tibetan *ybras* 'rice' as borrowed from Proto-Austroasiatic *c.rac 'rice (general)'. Two points remain unaddressed in this connection; firstly, was Proto-Austroasiatic *c.rac borrowed into PTH or independently into Proto-Chinese and Proto-Tibetic, and secondly why do Proto-Chinese and Proto-Tibetic as the only languages in his broad sample have the nasal prefix *m-? Ferlus (2010:71f.) explained the etymological meaning of *c.rac as 'the cereal collected by tearing off the grains along the stem with the hand', deriving it from the verb *rac 'to collect by tearing off the grain (of rice) along the stem with the hand'. This meaning appears to be at odds with the more general sense *'grain' reconstructed for *mrats in the present article. As against the reconstruction of a bilabial nasal, Peiros & Starostin (1996:84) proposed PST *phriā(s) 'rice' on the evidence from Chinese *bài* 稗 (Old Chinese *bhrēs) 'fine rice', Literary Tibetan *ybras* [ãbras] 'rice', Proto-Kiranti *rə 'rice'⁶⁷ and Mizo [raʔ] 'fruit, bear fruit'. The existence of historical relationship between *li* 糲, *bài* 稗, and *ybras* was also considered by Schuessler (1998:155, 352) who added to them Mru [rai].⁶⁸ The reconstructed Old Chinese vowel and semantics (*bài* 稗: LH be^c, OCM *brêh 'fine rice'; Middle Chinese *beaH*, Old Chinese *C.[b]^hre-s 'weed resembling grain'; Schuessler 2009:128; Baxter & Sagart 2014b:232) speak against the direct comparison between the word and Proto-Tibetic *mbras.

64. This reconstruction was first proposed by Karlgren (1957:100, 340g).

65. Due to their diverging semantics, I have excluded the Mizo (formerly Lushai) comparanda from the above discussion (see §6).

66. A historical relationship between, on the one hand, Chinese, Tibetan, Mizo words and, on the other hand, Austronesian forms was earlier suggested by Kuiper (1966:61).

67. This reconstruction is not further explained.

68. Löffler (1966:132). Löffler compares Mru [rai] with Mizo [tai], but his other comparanda show Mru [r-] ~ Mizo [hr-]/[h-] as well.

Hill (2019: 28f.) devoted some more space to explaining the origin of OLT *ybras* based on sound correspondences between Old Chinese *mr- and Literary Tibetan *br-* first suggested by Simon (1930: 41, § 86) and therefore dubbed Simon’s law. But the Tibetan-Chinese comparanda adduced by Hill concern the Literary Tibetan onset *ybr-* and not *br-*:⁶⁹

<i>ybras</i> ‘rice’ < * <i>hmras</i>	⊗ Chinese <i>lì</i> 糲; Middle Chinese <i>lat</i> ; Old Chinese *(mǎ-)r ^s at ⁷⁰
<i>ybri mo</i> ‘female yak’ < * <i>hmri mo</i>	⊗ Chinese <i>lí/máo</i> 犛; Middle Chinese <i>maew</i> ; Old Chinese *mr ^s u
<i>ybrog</i> ‘pasture’ < * <i>hmrok</i>	⊗ Chinese <i>mù</i> 牧; Middle Chinese <i>mjuwk</i> ; Old Chinese *mæk

The Tibetan reconstructions proposed by Hill are bound to three problems:

- Literary Tibetan *ybr-* (< Hill’s Proto-Tibetic **hmri-*) corresponds to three distinct onsets in Old Chinese: *r^s-, *mr^s-, and *m-
- The source of *y-* in *ybr-* remains unaddressed⁷¹
- *y-* is taken to represent /ɣ/⁷²

The last issue is related to Hill’s interpretation of the 23rd letter of the Tibetan alphabet: *ṛ*. He accepts its phonetic realisation as [N-] in pre-consonantal contexts for the period reconstructible from Tibetan dialects but analyses it synchronically for the period of the script invention as one phoneme with *y* written in absolute onset and coda. According to this hypothesis, phoneme /ɣ/, written in all positions as *ṛ*, very early developed two allophones: [N] before consonants and [ɣ] in other positions.⁷³ Several arguments can be raised against this hypothesis:

69. I consider *sbr-* in *sbrul* and *sbran* as derived in Tibetic through the prefix *s-* (*s+mbr-) and, therefore, not directly comparable with Chinese words. As Hill (2011: 453; 2019: 29) admits, words in *smr-* pose a serious obstacle to Simon’s law.

70. The bracketing of the syllable *(mǎ-) in the latest version of the Baxter-Sagart notation system indicates that the authors cannot provide any positive evidence for the existence of the segment (2014b: 7). In fact, Simon (1930: 41, § 86) compared Literary Tibetan *ybras* with *mǐ* 米 (Middle Chinese *mejX*, Old Chinese *(C.)m^s[ejjʔ] which is even less compelling.

71. *y-* is not known as a nominal prefix in OLT; it only surfaces as a derivational and inflectional prefix in the verb system (Bialek 2026: 394ff.). Should it be understood as inherited (a question Hill does not address), no languages considered for this study attest to any segment that could be compared with the preconsonantal *y-* [ɣ].

72. Hill (2019: 28, Footnote 36).

73. Hill’s interpretation is isolated, for it is more commonly assumed that the letter *ṛ* represented two distinct phonemes (cf. Coblin 2002: 181ff.; Sun 2003: 779, Footnote 14; Jacques 2012: 91f.)

- No linguistic data justifies the reconstruction of [ɣ] to preconsonantal contexts. Even modern dialects that, according to the proposed model, are descendants of AOT-NE and, therefore, split off before the script invention have nasals in their equivalents of OLT *ɣ* in preconsonantal position (cf. Table 2 and §4.1).
- Labial consonants never receive prescript *m*- in the script. The lack of the cluster 'mb- already points to a merger, the outcome of which is always written with *ɣ*- although the resulting sound must have been nasal: *ɣ*- [ɣ-] + *m*- [m-] > *ɣ*- / _C^[+labial]. Thus, the use of the letter *ɣ* cannot be applied to the reconstruction of historical phonology in this case.
- Examples of rhinoglottophilia provided by Matisoff (1975) and recalled by Hill (2009: 126) all concern the historical change glottal (not velar!) > nasal, which occurs word-internally and preponderantly in contexts with low vowels. In all quoted cases, nasalisation develops as co-articulation and does not replace the fricative.⁷⁴ Neither Matisoff's examples nor his articulatory explanations support the change [ɣ-] > [N-] in onset. At the same time, nasalisation of fricatives typically occurs "in positions where nasality spreads from a nasal consonant or a nasalised vowel in the neighbourhood" (Ladefoged & Maddieson 1996: 132). Therefore, the literature does not support the spontaneous change [ɣ-] > [N-] presumed for EOT by Hill.⁷⁵
- In OLT, the letter *ɣ* is not used in coda with any regularity, which disrupts the presumed pattern observed with other prescripts written in three positions: pre-consonantal, absolute onset, and coda. This challenges the unitary phonemic interpretation of the letter *ɣ*.
- The archaic consonant clusters *snr*- and *smr*- preserve the original onsets **ɣr*- [N-r-] and **mr*- [m-r-].⁷⁶
- The Tibetan alphabet was devised as phonetic and not phonemic.⁷⁷
- Prefix *ɣ*- [N-], primarily a derivational marker of decreased transitivity (cf. OLT INTR *ɣpho* 'to move' vs TR *spo* 'to change'), is historically related to the PTH anticausative nasal prefix *N-, confirming its reconstructed phonetic value.⁷⁸

74. Matisoff (1975: 265) speaks of "nasalization of vowels in the environment of 'laryngeals'".

75. The complementary distribution of [h] and [ŋ] in English (adduced anecdotally by Matisoff 1975: 269) does not lead anybody to posit one phoneme for them.

76. Bialek (2026: 436).

77. Bialek (2026: §4.4).

78. See Bialek (Forthcoming a). The problem of the historical source of the preconsonantal *ɣ*- in not addressed by Hill.

But I think that the strongest argument against Hill’s interpretation (2019: 28f.) of $yC-$ as $/yC-/$ comes from the current discussion of $ybras$:

Hill: Proto-Bodic *mras > Proto-Tibetic * $\text{h}mras$ > $ybras$ [^Nbras]

Bialek: Proto-Bodic *mras > Proto-Tibetic *mbras > AOT *[mbras] > LOT
*[^Nbras] $\not\approx$ OLT $ybras$

Whereas Hill’s reconstruction involves two processes otherwise unknown to Tibetic (nominal prefixation with * $\text{h}-$ / $\text{y}-$ / and fortition * $\text{mr}-$ > $\text{br}-$), the author’s proposal is more parsimonious for it involves one sound change well-documented and acknowledged in Tibetic: consonant insertion before a postconsonantal liquid. Proto-Tibetic underwent two similar but distinct changes that have been treated uniformly in previous works (see Bialek 2018b:9, Footnote 25; Hill 2019: 17f.):

- I. Excrescence $C^{[+nasal]}C^{[+liquid]}$ > $C^{[+nasal]}CC^{[+liquid]}$
 $^{*N}l_1-$ > $^{*N}t_1-$ (excrescence) > $^{*N}l_1t-$ (metathesis) > EOT $lt-$ [$lt-$]:
 v1 $ltuñ$, v2 $lhuñ$ ‘to fall’
 $^{*N}l-$ > $^{*N}dl-$ (excrescence) > $^{*N}ld-$ (metathesis) > EOT $ld-$ [$ld-$]:
 v1 $ldañ$, v2 $lan(s)$ ‘to rise’⁷⁹
 $^{*N}r-$ > $^{*N}dr-$ (excrescence) > EOT $ydr-$ [^Ndr-]:
 v1 $ydral$, v2 ral INTR ‘to tear asunder’⁸⁰
 $^{*mr}-$ > $^{*mbr}-$ (excrescence) > EOT $ybr-$ [$mbr-$]:
 Proto-Bodic *mras > OLT $ybras$ ‘rice’⁸¹

Another variant of excrescence concerns the cluster $C^{[+fricative]}C^{[+rhotic]}$:

- *sr- OLT $sr-$ > EOT [$str-$] (excrescence)
 $^{*zr}-$ OLT $zr-$ > EOT $^{*zdr}-$ (excrescence) > [$zgr-$] (merger with $/sgr-/$)

79. The first two changes apply only to syllables with non-front vowels. Those with the front vowels [i] and [e] underwent palatalisation to [tɕ] and [dʒ] from [t] and [l] respectively.

80. For this conjugation, see Bialek (2020a: 304).

81. As rightly observed by Hill (2019: 19), the outcome of Proto-Bodic *ml- in Tibetic differs from that of ^NL- so that a complete distinct path of sound changes has to be reconstructed for the former, although likewise here we observe distinct changes in Proto-Tibetic for syllables with front and non-front vowels. That Proto-Bodic reflexes of *mr- and *ml- differ in Proto-Tibetic does not need to surprise us, for Bialek (2018b) already demonstrated with divergent reflexes of Proto-Tibetic *sL- and *sR- clusters that liquids were not a natural class in EOT. The question of modern reflexes of Proto-Bodic *mr- and *ml- will be addressed from the comparative perspective in Bialek (Forthcoming b).

II. Affrication $C^{[+nasal/r-/l-/s-]}C^{[+fricative]} > C^{[+nasal/r-/l-/s-]}C^{[+affricate]}$

*^Ns- > [^Nts-] OLT *yc-*

*^Nz- > [^Ndz-] OLT *ɣj*⁸²

*^Nç- > [^Ntç-] OLT *ɣč-*

*^Nz- > [^Ndz-] OLT *ɣj̃-*

*rs- > [rts-] OLT *rc-*

*rz- > [rdz-] OLT *rj-*

*rz- > [rdz-] OLT *rj̃-*

*ss- > [sts-] OLT *sc-*

*sç- > [stç-] OLT *sč-*

Another variant of affrication concerns the cluster $C^{[+fricative]}C^{[+lateral]}$:

*sl- OLT *sl-* > EOT *ls- (metathesis) > [lts-]

*zl- OLT *zl-* > EOT [lz-] (metathesis) > [ldz-]⁸³

Both excrescence and affrication were still productive in EOT. The regular affrication (II) outcome is an affricate consonant as part of a bi-segmental syllable onset. In contrast, the primary outcome of excrescence (I) consists of three segments. Insertion of a plosive into the cluster of a nasal and a liquid is a cross-linguistically well-attested sound change.⁸⁴ It is likewise well-documented in Tibetic with the nasal prefix [^N-]. The inserted consonant has the same place of articulation as the preceding nasal (Bybee 2015: 480) or fricative as with Proto-Tibetic *sr- and *zr-. Accordingly, there are no words in *'mdr-* in OLT⁸⁵ because Proto-Bodic *mr- became Proto-Tibetic *mbr-.

The fact that a preconsonantal nasal in modern equivalents of OLT *ɣbr-* is found in several very distinct dialect subgroups (AOT-NE, MOT-NE, LOT-E,

82. According to Schiefner's law, formulated by Hill (2019: 26ff.), Proto-Tibetic did not have inherited voiced fricatives [z] and [ʒ]. However, pairs of alternating verbs such as \sqrt{zad} 'to decline' \bowtie \sqrt{sad} 'to kill', \sqrt{zig} 'to collapse' \bowtie $\sqrt{çig}$ 'to destroy', and \sqrt{zu} 'to come off' \bowtie $\sqrt{çu}$ 'to peel off', as well as the existence of words *jo mo*, *jo bo*, *je ba*, or *ju tig* in OLT (see Bialek 2021c), demonstrate that the law is not exceptionless and might require additional stipulations to be generally acknowledged.

83. For the last two sound changes, see Bialek (2018b).

84. Cf.: Proto-Indo-European *a-mr̥t-os > Greek *ambrotos*, Ancient Greek *anr-os* (GEN SG) > Greek *andros*, Old English *bræmle* > Present-Day English *bramble*, Old English *θymel* > Present-Day English *thimble*, Old English *θunrian* > Present-Day English *thunder* (cf. German *Donner*), Latin *camera* > French *chambre*, Latin *hominem* > Spanish *hombre* (cf. French *homme*), Spanish *venirá* > *vendrá* '3SG FUT'. Data *apud* Campbell (2004: 34f.); Bybee (2015: 480).

85. *mdra* in Pelliot tibétain 981: r203 should be emended to *mday* 'arrow'.

LOT-C, EMT-W, EMT-E), as well as in the borrowings into Rgyalrongic, justifies the reconstruction of a nasal consonant to Proto-Tibetic. Hence, I dismiss Simon's law as moot and tentatively conclude that OLT nominal forms in *ybr-* (Proto-Tibetic *mbr-) come from Proto-Bodic *mr-.⁸⁶ I hope to further substantiate this hypothesis in another paper in the near future (see Bialek Forthcoming b).

8. Conclusions

This article has proposed a new historical subgrouping of Old Tibetan dialects, refining the one in Bialek (2018b). It demonstrated for the first time that Tibetan dialects started to differentiate in late pre-historical times, shortly before the introduction of the script. The emergence of the first dialects was bound to the military expansion of the Tibetan Empire, a polity that was to conquer the Tibetan Plateau within roughly 250 years. The present reconstruction of the dialect formation supplemented earlier insights on early sound changes by studying modern equivalents of OLT *ybras*, the Proto-Tibetic form of which could be reconstructed as *mbras. I have shown that the latter is cognate with several other lexical units from neighbouring languages. The comparison of the cognates permitted the reconstruction of *mras to Proto-Bodic and, tentatively, *mrats to Proto-Trans-Himalayan. The discussion has provided the first tangible proof for the inclusion of Tibetic, East Bodish, Tshanglic, and Kaike-Ghale-Tamangic in one sub-branch of the Trans-Himalayan language family termed Bodic, but the results arrived at in this paper will have to be continuously refined and updated when more research has been done on the history of the Trans-Himalayan language family.

Funding

This project has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101149388—TOMB: Tibetan Obsolete Mortuary practices and afterlife Beliefs. Language conservatism of religious writings in the service of Proto-Bodish reconstruction.

86. Worth quoting in this context is Wolfenden's intuition (1936:401, Note 1) expressed towards Simon's Tibetan-Chinese comparanda: "I do not believe that more than a fraction of the equations there set up will prove to be true". With the present analysis, one of Simon's equations is hereby rejected.

Acknowledgements




This paper owes a lot to the work of Hiroyuki Suzuki, whose attention to detail in the description of Tibetan dialects of Khams brings Tibetan linguistics to a new level. I wish to thank him for his invaluable advice, generosity in sharing data, and readiness to help with their interpretation. I thank Lingzi Zhuang for an engaging discussion of the first draft of the paper. I also thank the anonymous reviewers for their numerous comments and corrections.







List of abbreviations

!	form not attested or violating established rules
#	word-boundary
*	reconstructed form
×	cognate with
ABS	absolutive
ADD	additive
AOT	Ancient Old Tibetan
C	¹ consonant; ² centre
COM	comitative
COP	copula
D	excrecent alveolar
DEL	delative
DPASS	dynamic passive
E	East
EB	East Bodish
EMPH	emphatic
EMT	Early Middle Tibetan
EOT	Early Old Tibetan
F	feminine
FNL	final
FOC	focus
FUT	future
GEN	genitive
INESS	inessive
INTR	intransitive
IPFV	imperfective
K	velar consonant
KGT	Kaika-Ghale-Tamangic
L	lateral consonant
LH	Late Han Chinese
LOT	Late Old Tibetan
lw	loanword
MOT	Middle Old Tibetan
N	North

NE	North-East
NEG	negation
NMLZ	nominaliser
NW	north-west
OMC	Old Minimal Chinese
OLT	Old Literary Tibetan
OT	Old Tibetan
PB	Proto-Bodic
PEB	Proto-East Bodish
PFV	perfective
PKGT	Proto-Kaike-Ghale-Tamangic
PST	Proto-Sino-Tibetan
PTam	Proto-Tamangic
PTB	Proto-Tibeto-Burman
PTH	Proto-Trans-Himalayan
PTib	Proto-Tibetic
PTsh	Proto-Tshanglic
R	rhotic consonant
S	South
SG	singular
STEDT	The Sino-Tibetan Etymological Dictionary and Thesaurus (see Matisoff 1987–Ongoing)
T	alveolar consonant
TERM	terminative
trslr.	transliteration
V	vowel
v1, v2, v3, v4	verb stems
W	West
WAT	Western Archaic Tibetan

References



-  Bailey, T. Grahame. 1911. Kanauri vocabulary in two parts: English-Kanauri and Kanauri-English. *Journal of the Royal Asiatic Society of Great Britain and Ireland (New Series)* 43(2), 315–364.
- Baxter, William H. & Sagart, Laurent. 2014a. *Baxter-Sagart Old Chinese reconstruction*. (<https://sites.lsa.umich.edu/ocbaxtersagart/wp-content/uploads/sites/1415/2025/03/BaxterSagartOCbyGSR2014-09-20.pdf>) (Accessed 2025-12-18.) (Version 1.1, 20 September 2014.)
-  Baxter, William H. & Sagart, Laurent. 2014b. *Old Chinese: A new reconstruction*. Oxford: Oxford University Press.
-  Beckwith, Christopher I. 2009. *Empires of the Silk Road: A history of Central Eurasia from the Bronze Age to the present*. Princeton: Princeton University Press.
- Benedict, Paul K. 1972. *Sino-Tibetan: A conspectus*. Cambridge: Cambridge University Press.

- Berger, Hermann. 1998. *Die Burushaski-Sprache von Hunza und Nager, teil I: Grammatik*. Wiesbaden: Harrassowitz.
- Beyer, Stephan V. 1993. *The Classical Tibetan language*. Delhi: Sri Satguru Publications.
- Bialek, Joanna. 2018a. *Compounds and compounding in Old Tibetan: A corpus based approach*. Marburg: Indica et Tibetica. (2 volumes.)
-  Bialek, Joanna. 2018b. The Proto-Tibetan clusters sL- and sR- and the periodisation of Old Tibetan. *Himalayan Linguistics* 17(2). 1–50.
-  Bialek, Joanna. 2020a. Old Tibetan verb morphology and semantics: An attempt at a reconstruction. *Himalayan Linguistics* 19(1). 263–346.
- Bialek, Joanna. 2020b. Towards a standardisation of Tibetan transliteration for textual studies. *Revue d'Etudes Tibétaines* 56. 70–88.
- Bialek, Joanna. 2021a. Kinterms: New potential indicators for dating Old Tibetan documents. *Revue d'Etudes Tibétaines* 60. 6–48.
- Bialek, Joanna. 2021b. Naming the empire: from *Bod* to *Tibet* – A philologico-historical study on the origin of the polity. *Revue d'Etudes Tibétaines* 61. 339–402.
- Bialek, Joanna. 2021c. Comments on Jacques' "The directionality of the voicing alternation in Tibetan". *Journal of the Southeast Asian Linguistics Society* 14(1). xii–xxiii.
-  Bialek, Joanna. 2026. *Old Literary Tibetan: A comprehensive text grammar based on the Old Tibetan Annals*. Leiden: Brill.
- Bialek, Joanna. Forthcoming a. Causative alternation in Tibetic and Trans-Himalayan. (Accepted by *Linguistics of the Tibeto-Burman Area*, to be published in 2026, in issue 49(2).)
- Bialek, Joanna. Forthcoming b. Proto-Bodic *mr- and *ml- and their modern reflexes.
- Bialek, Joanna & Suzuki, Hiroyuki. Forthcoming. Devoicing phenomena across Tibetic: Patterns in plosives and fricatives.
- Bielmeier, Roland. 1982. Problems of Tibetan dialectology and language hi-story with special reference to the *sKyid grong* dialect. *Zentralasiatische Studien* 16. 405–425.
- Bielmeier, Roland. 1985. A survey of the development of western and southwestern Tibetan dialects. In Aziz, Barbara Nimri & Kapstein, Matthew (eds.), *Soundings in Tibetan civilization*, 3–19. New Delhi: Manohar.
- Bielmeier, Roland & Haller, Felix & Häsler, Katrin & Huber, Brigitte & Volkart, Marianne. 2013. *Comparative dictionary of Tibetan dialects*. (Draft.)
-  Bielmeier, Roland & Häsler, Katrin & Haller, Chungda & Haller, Felix & Hein, Veronika & Huber, Brigitte & Volkart, Marianne & Preiswerk, Thomas & Ngawang Tsering & Widmer, Manuel & Zemp, Marius. 2018. *Comparative dictionary of Tibetan dialects (CDTD), volume 2: Verbs*. Berlin: De Gruyter Mouton.
-  Bodt, Timotheus A. 2012. *The new lamp clarifying the history, peoples, languages and traditions of Eastern Bhutan and Eastern Mon*. Wageningen: Monpasang Publications.
- Bodt, Timotheus A. 2023. East Bodish revisited. *Bulletin of Tibetology* 54(1). 49–212.
- Buffetrille, Katia. 2008. Some remarks on mediums: The case of the *lha pa* of the musical festival (*glu rol*) of Sog ru (A mdo). *Mongolo-Tibetica Pragensia* 1(2). 13–66.
-  Bybee, Joan L. 2015. Articulatory processing and frequency of use in sound change. In Honeybone, Patrick & Salmons, Joseph (eds.), *The Oxford handbook of historical phonology*, 467–484. Oxford: Oxford University Press.

- Campbell, Lyle. 2004. *Historical linguistics*. Cambridge: The MIT Press.
- [doi](#) Coblin, W. South. 2002. On certain functions of 'a-chung in early Tibetan transcriptional texts. *Linguistics of the Tibeto-Burman Area* 25(2). 169–184.
- [doi](#) d'Alpoim Guedes, Jade. 2011. Millets, rice, social complexity, and the spread of agriculture to the Chengdu plain and Southwest China. *Rice* 4 (3–4). 104–113.
- [doi](#) d'Alpoim Guedes, Jade & Lu, Hongliang & Li, Yongxian & Spengler, Robert N. & Wu, Xiaohong & Aldenderfer, Mark S. 2014. Moving agriculture onto the Tibetan plateau: The archaeobotanical evidence. *Archaeological and Anthropological Sciences* 6(3). 255–269.
- de Vries, Klaas. 2020. A sociolinguistic survey of Humla Tibetan in Northwest Nepal. *Electronic Survey Report* 2020(13). 1–173.
- [doi](#) Dhakal, Dubi Nanda & List, Johann-Mattis & Roberts, Seán G. 2024. A phylogenetic study of South-Western Tibetic. *Journal of Language Evolution* 9(1–2). 14–28.
- Dondrup, Rinchin. 1993. *Brokeh language guide*. Itanagar: Directorate of Research, Arunachal Pradesh Government.
- Donohue, Mark & Dhakal, Dubi Nanda. 2016. *A Tsum lexicon*. München: Lincom.
- Dzongkha Development Commission. 2017. *Dzongkha-English-dZalakha lexicon*. Thimphu: Dzongkha Development Commission.
- Dzongkha Development Commission. 2018. *Bumthangkha-Dzongkha-English lexicon*. Thimphu: Dzongkha Development Commission.
- Ferlus, Michel. 2010. The Austroasiatic vocabulary for rice: Its origin and expansion. *Journal of the Southeast Asian Linguistics Society* 3(2). 61–76.
- [doi](#) Funk, Damian. 2020. A preliminary phonology of Brokpa. *Himalayan Linguistics* 19(1). 22–41.
- [doi](#) Gerber, Pascal & Grollmann, Selin. 2025. The phylogenetic position of Tshangla: A reevaluation. *Cahiers de Linguistique Asie Orientale* 54(2). 231–257.
- [doi](#) Grollmann, Selin. 2020. *A grammar of Bjokapakha*. Leiden: Brill.
- [doi](#) Grollmann, Selin & Gerber, Pascal. 2018. Linguistic evidence for a closer relationship between Lhokpu and Dhimal: Including some remarks on the Dhimalish subgroup. *Cahiers de Linguistique Asie Orientale* 47(1). 1–96.
- [doi](#) Hill, Nathan W. 2009. Tibetan <h-> as a plain initial and its place in Old Tibetan phonology. *Linguistics of the Tibeto-Burman Area* 32(1). 115–140.
- [doi](#) Hill, Nathan W. 2010. An overview of Old Tibetan synchronic phonology. *Transactions of the Philological Society* 108(2). 110–125.
- [doi](#) Hill, Nathan W. 2011. An inventory of Tibetan sound laws. *Journal of the Royal Asiatic Society* 21(4). 441–457.
- Hill, Nathan W. 2015. Tibetan *-as > -os. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 12. 165–175.
- [doi](#) Hill, Nathan W. 2019. *The historical phonology of Tibetan, Burmese, and Chinese*. Cambridge: Cambridge University Press.
- Honkasalo, Sami. 2019. *A grammar of Eastern Geshiza: A culturally anchored description*. Helsinki: University of Helsinki (Doctoral dissertation).
- Huang, Bufan & Xu, Shouchun & Chen, Jiaying & Wang, Huiyin & Dai, Qingxia. 1992. *Zangmianyuzu yuyan cihui*. Beijing: China Minzu University Press.

- Hyslop, Gwendolyn. 2014. A preliminary reconstruction of East Bodish. In Hill, Nathan W. & Owen-Smith, Thomas (eds.), *Trans-Himalayan linguistics: Historical and descriptive linguistics of the Himalayan area*, 155–179. Berlin: De Gruyter.
-  Hyslop, Gwendolyn & d'Alpoim-Guedes, Jade. 2021. Linguistic evidence supports a long antiquity of cultivation of barley and buckwheat over that of millet and rice in Eastern Bhutan. *Vegetation History and Archaeobotany* 30(4). 571–579.
- Iwao, Kazushi & Hill, Nathan W. & Takeuchi, Tsuguhito (eds.). 2009. *Old Tibetan inscriptions*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
-  Jacques, Guillaume. 2009. Le développement du tibétain ancien *-e-* dans les dialectes occidentaux. *Études mongoles et sibériennes, centrasiatiques et tibétaines* 40. 1–5.
-  Jacques, Guillaume. 2012. A new transcription system for Old and Classical Tibetan. *Linguistics of the Tibeto-Burman Area* 35(2). 89–96.
- Jacques, Guillaume. 2014. Cone. In Sun, Jackson T.-S. (ed.), *Phonological profiles of little-studied Tibetic varieties*, 269–375. Taipei: Institute of Linguistics, Academic Sinica.
-  Jacques, Guillaume. 2016. *Esquisse de phonologie et de morphologie historique du tangoute*. Leiden: Brill. (Hardback published in 2014.)
- Jacques, Guillaume. 2021. *A grammar of Japhug*. Berlin: Language Science Press.
- Jäschke, Heinrich August. 2003[1881]. *A Tibetan-English dictionary dictionary: With special reference to the prevailing dialects; to which is added an English-Tibetan vocabulary*. New York: Dover Publications. (First published by the Secretary of State for India in Council, in London.)
- Karlgren, Bernhard. 1957. *Grammata serica recens*. Stockholm: The Museum of Far Eastern Antiquities.
- Kelly, Barbara. 2004. A grammar and glossary of the Sherpa language. In Genetti, Carol (ed.), *Tibeto-Burman languages of Nepal: Manange and Sherpa*, 193–324. Canberra: Pacific Linguistics.
- Kuiper, Franciscus B. J. 1966. The sources of the Nahali vocabulary. In Zide, Norman H. (ed.), *Studies in comparative Austroasiatic linguistics*, 57–81. The Hague: Mouton & Co.
- Kümmel, Martin Joachim. 2007. *Konsonantenwandel: Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen für die vergleichende Rekonstruktion*. Wiesbaden: Reichert.
-  Kuteva, Tania & Heine, Bernd & Hong, Bo & Long, Haiping & Narrog, Heiko & Rhee, Seongha (eds.). 2019. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Ladefoged, Peter & Maddieson, Ian. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
-  Lai, Yunfan. 2016. Causativisation in Wobzi and other Khroskyabs dialects. *Cahiers de Linguistique Asie Orientale* 45(2). 148–175.
- Lin, You-Jing. 2014. Thebo. In Sun, Jackson T.-S. (ed.), *Phonological profiles of little-studied Tibetic varieties*, 215–267. Taipei: Institute of Linguistics, Academic Sinica.
-  Liu, Kitty Wenying. 2024. The development of *zl* in Tibetic languages. *Cahiers de Linguistique Asie Orientale* 53(1). 32–67.

- Löffler, Lorenz G. 1966. The contribution of Mru to Sino-Tibetan linguistics. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 116(1). 118–159.
- Lorimer, David Lockhart Robertson. 1938. *The Burushaski language, volume 3: Vocabularies and index*. Oslo: H. Aschehoug & Co.
- Mainwaring, George Byres & Grünwedel, Albert (eds.). 1898. *Dictionary of the Lepcha language*. Berlin: Unger Brothers.
- Matisoff, James A. 1975. Rhinoglottophilia: The mysterious connection between nasality and glottality. In Ferguson, Charles Albert & Hyman, Larry M. & Ohala, John J. (eds.), *Nasálfest: Papers from a symposium on nasals and nasalization*, 265–287. Stanford: Stanford University.
- Matisoff, James A. (ed.). 1987–Ongoing. *The Sino-Tibetan etymological dictionary and thesaurus*. Berkeley: University of California. (<http://stedt.berkeley.edu/~stedt/cgi/rootcanal.pl>) (Accessed 2017-11-07.)
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. Berkeley: University of California Press.
- Michailovsky, Boyd & Mazaudon, Martine. 1994. Preliminary notes on the languages of the Bumthang group. In Kværne, Per (ed.), *Tibetan Studies, Proceedings of the 6th Seminar of the International Association for Tibetan Studies, Fagernes 1992*, 545–557. Oslo: The Institute for Comparative Research in Human Culture.
-  Miller, Roy Andrew. 1955. Notes on the Lhasa dialect of the early ninth century. *Oriens* 8(1). 284–291.
-  Nagano, Yasuhiko. 1983. A historical study of gLo Tibetan. *Bulletin of the National Museum of Ethnology* 7(3). 472–513.
- Nishida, Tatsuo. 1963. On the T'ien-ch'üan Tibetan dialect of Hsi-K'ang in the sixteenth century: A study of the Chinese-Tibetan vocabulary, *Hsi-Fan-Kuan I-yu*. *Kyoto Daigaku Bungakubu Kenkyu Kiyo* 7. 84–174.
-  Noonan, Michael. 1999. *Chantyal dictionary and texts*. Boston: Mouton de Gruyter.
-  Owen-Smith, Thomas. 2015. *Grammatical relations in Tamang, a Tibeto-Burman language of Nepal*. London: SOAS University of London. (Doctoral dissertation.)
- Peiros, Ilia & Starostin, Sergei A. 1996. *A comparative vocabulary of five Sino-Tibetan languages, fascicle V: Velars, uvulars and laringals*. Melbourne: Department of Linguistics and Applied Linguistics, the University of Melbourne.
- Plaisier, Heleen. 2007. *A grammar of Lepcha*. Leiden: Brill.
- Róna-Tas, András. 1966. *Tibeto-Mongolica: The Tibetan loanwords of Monguor and the development of the Archaic Tibetan dialects*. Budapest: Akadémiai Kiadó.
- Róna-Tas, András. 1985. *Wiener Vorlesungen zur Sprach- und Kulturgeschichte Tibets*. Wien: Arbeitskreis für Tibetische und Buddhistische Studien, Universität Wien.
-  Rufenacht, Sara & Waldis, Sereina. 2020. Diachronic and areal aspects of Brokpa phonology. *Himalayan Linguistics* 19(1). 42–53.
- Sagart, Laurent. 2003. The vocabulary of cereal cultivation and the phylogeny of East Asian languages. *Bulletin of the Indo-Pacific Prehistory Association* 23. 127–136.
-  Schuessler, Axel. 1998. Another note on Old Tibetan *rje-blas*. *Linguistics of the Tibeto-Burman Area* 21(2). 3–4.

- Schuessler, Axel. 2007. *ABC etymological dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.
- Schuessler, Axel. 2009. *Minimal Old Chinese and later Han Chinese: A companion to Grammata Serica Recensa*. Honolulu: University of Hawai'i Press.
- Scott, James C. 2009. *The art of not being governed: An anarchist history of Upland Southeast Asia*. New Haven: Yale University Press.
-  Scott, James C. 2017. *Against the grain: A deep history of the earliest states*. New Haven: Yale University Press.
-  Sharma, Suhnu Ram. 2001. A sketch of Rongpo grammar. In Nagano, Yasuhiko & LaPolla, Randy J. (eds.), *New Research on Zhangzhung and Related Himalayan Languages*, 195–270. Osaka: National Museum of Ethnology.
-  Simon, Walter. 1930. *Tibetisch-Chinesische Wortgleichungen: Ein Versuch*. Berlin: Walter de Gruyter.
-  Song, Jixiang & Lu, Hongliang & Zhang, Zhengwei & Liu, Xinyi. 2018. Archaeobotanical remains from the mid-first millennium AD site of Kaerdong in western Tibet. *Archaeological and Anthropological Sciences* 10(8). 2015–2026.
- Sprigg, Richard Keith. 1968. The role of *R* in the development of the modern spoken Tibetan dialects. *Acta Orientalia Academiae Scientiarum Hungaricae* 21(3). 301–311.
-  Sprigg, Richard Keith. 1979. The Golok dialect and written Tibetan past-tense verb forms. *Bulletin of the School of Oriental and African Studies* 42(1). 53–60.
- Stein, Rolf Alfred. 1972. *Tibetan civilization*. Stanford: Stanford University Press.
- Sun, Jackson T.-S. 1993. *A historical-comparative study of the Tani (Mirish) branch in Tibeto-Burman*. Berkeley: University of California. (Doctoral dissertation.)
- Sun, Jackson T.-S. 2003. Phonological profile of Zhongu: A new Tibetan dialect of Northern Sichuan. *Language and Linguistics* 4(4). 769–836.
- Sun, Jackson T.-S. 2025. A phonological and lexical profile of Sastod: A non-Amdo Tibetan language of Khrochu (Sichuan). *Language and Linguistics* 26(4). 695–751.
- Suzuki, Hiroyuki. 2007. /r/ as a glide of the consonant cluster in sProsnang [Zhonglu] Tibetan. *Tokyo University Linguistic Papers* 26. 31–47.
-  Suzuki, Hiroyuki. 2009. Khams Tibetan sPomtserag [Benzilan] dialect: Phonetic analysis. *Asian and African Languages and Linguistics* 4. 219–258.
-  Suzuki, Hiroyuki. 2010. Khams Tibetan Khyungpo/Khromtshang dialect: Phonetic and dialectal analysis. *Journal of Asian and African Studies* 79. 95–120.
-  Suzuki, Hiroyuki. 2011. Phonetic analysis of dGudzong Tibetan: The vernacular of Khams Tibetan spoken in the rGyalrong area. *Bulletin of the National Museum of Ethnology* 35(4). 617–653.
- Suzuki, Hiroyuki. 2012a. A propos du terme 'riz' et de l'hypothèse du groupe dialectal de Sems-kyi-nyila en tibétain du Khams. *Revue d'Etudes Tibétaines* 23. 107–115.
-  Suzuki, Hiroyuki. 2012b. Khams Tibetan Sangdam dialect: Phonetic and dialectal analysis. *Journal of Asian and African Studies* 83. 37–58.
- Suzuki, Hiroyuki. 2014. *Xifanguan Yiyu*, a Chinese-Tibetan vocabulary edited in the 16th century, as a dialectal material of Minyag Rabgang Tibetan. (Paper presented at the 24th Annual Meeting of the Southeast Asian Linguistics Society, Myanmar, 27–31 May 2014.)

- Suzuki, Hiroyuki. 2020a. Kamu Chibettogo Sasage hachi [Phongpa] hōgen no watari-on /r/. *Nidaba* 49. 31–38.
- Suzuki, Hiroyuki. 2020b. Geolinguistic significance of the Phongpa dialect in the history of Yunnan Tibetan. In Fukushima, Chitsuko (ed.), *Proceedings of the Second Annual Meeting of the Geolinguistic Society of Japan*, 25–29. (<https://zenodo.org/records/4515013>) (Accessed 2025-09-25.)
- Suzuki, Hiroyuki. 2022a. Sound correspondences of the 'Azha dialect of Thewo-smad Tibetan with Literary Tibetan and a Japanese-'Azha word list: With an introduction to Tibetic languages in Thewo County. *Journal of Kijutsuken* 14. 65–114. (<http://id.nii.ac.jp/1422/00004411/>) (Accessed 2025-07-23.)
-  Suzuki, Hiroyuki. 2022b. Dialectal classification of the Phongpa Dialect of Khams Tibetan. *Tokyo University Linguistic Papers* 44. e133–e147.
-  Suzuki, Hiroyuki. 2023. Geolinguistic approach to migration history in the south-eastern edge of the Tibetosphere: A case study of Sangdam Tibetan and methodological remarks. *Asian Languages and Linguistics* 4(2). 224–250.
- Suzuki, Hiroyuki. 2024a. Shaping rGyalthagang: A historical account of Yunnan Khams. In Nohara, Masaki & Ikeda, Takumi (eds.), *Grammatical phenomena of Sino-Tibetan languages 6: Typology and historical change*, 87–108. Kyoto: Institute for Research in Humanities, Kyoto University.
-  Suzuki, Hiroyuki. 2024b. Dialectal characteristics of Khams Tibetan Nyishe/Khrezhag dialect: Accompanied with its vocabulary. *Asian and African Languages and Linguistics* 18. 1–39.
-  Suzuki, Hiroyuki. 2024c. Evolution of dorsal fricatives in rGyalthagang varieties of Khams Tibetan. *Journal of the Phonetic Society of Japan* 28(2). 107–118.
- Suzuki, Hiroyuki. Forthcoming. Irregular lexical innovation of *yin* in south-eastern Tibetic languages.
- Suzuki, Hiroyuki & Sonam Wangmo. 2016. Cultural contexts of the expansion of a Tibetan word 'bras 'rice' in the easternmost Tibetosphere. In Endo, Mitsuaki (ed.), *Papers from the Third International Conference on Asian Geolinguistics*, 72–79. Fuchu: Research Institute for Languages and Cultures of Asian and Africa, Tokyo University of Foreign Studies.
-  Suzuki, Hiroyuki, & Sonam Wangmo. 2017. Vocabulary of Shingnyag Tibetan: A dialect of Amdo Tibetan spoken in Lhagang, Khams Minyag. *Asian and African Languages and Linguistics* 11. 101–127.
-  Tournadre, Nicolas Laurent & Rigzin, Karma. 2015. Outline of Chocha-Ngacha. *Himalayan Linguistics* 14(2). 49–87.
- Tournadre, Nicolas & Sherpa, Lhakpa Norbu & Gyurme Chodrak & Oisel, Guillaume. 2009. *Sherpa-English and English-Sherpa dictionary with literary Tibetan and Nepali equivalents*. Kathmandu: Vajra Publications.
- Tournadre, Nicolas & Suzuki, Hiroyuki. 2023. *The Tibetic languages: An introduction to the family of languages derived from Old Tibetan*. Villejuif: LACITO.
- Uray, Géza. 1954. Review of *Tibetische Dialekte von Amdo*. *Anthropos*, vol. 47 (1952), no 1–2, pp. 193–202, by Hermanns, Mathias. *Acta Orientalia Academiae Scientiarum Hungaricae* 4(1-3). 308–314.

- van Driem, George. 2001. *Languages of the Himalayas: An ethnolinguistic handbook of the Greater Himalayan region containing an introduction to the symbiotic theory of language*, vol. 2. Leiden: Brill.
-  van Driem, George. 2015. Synoptic grammar of the Bumthang language: A language of the central Bhutan highlands. *Himalayan Linguistics Archive* 6. 1–77.
- Vesalainen, Olavi. 2016. *A grammar sketch of Lhomi*. Dallas: SIL International.
-  Wang, Yanren & Yang, Qi & Yang, Jishuai & Sharma, Shalini & Huang, Yunzhe & He, Wei & Tstring, Tinley & Hu, Shihua & Chen, Songtao & Sun, Qingli & You, Ting & Gu, Zhengquan & Tsho, Yang & Wangdue, Shargan & Chen, Zujun & Gao, Yu & Yang, Xiaoyan & Chen, Fahu. 2024. The evolution of cropping structure in prehistoric Xizang. *Science Bulletin* 69(24). 3959–3967.
-  Wangdi, Pema. 2020. Segmental and suprasegmental features of Brokpa. *Himalayan Linguistics* 19(1). 393–422.
-  Wangdi, Pema. 2021. *A grammar of Brokpa: A Trans-Himalayan language of Bhutan*. Cairns: James Cook University. (Doctoral dissertation.)
- Watkins, Calvert. 1973. Language and its history. *Dædalus* 102(3). 99–111.
- Webster, Jeff. 2022. A sociolinguistic survey of the Tibeto-Burman dialects of North Gorkha District, Nepal. *Journal of Language Survey Reports* 2022(2). i–114.
-  Widmer, Manuel. 2017. *A grammar of Bunan*. Berlin: de Gruyter Mouton.
- Willis Oko, Christina. 2019. *A grammar of Darma*. Leiden: Brill.
-  Wolfenden, Stuart N. 1936. On certain alternations between dental finals in Tibetan and Chinese. *Journal of the Royal Asiatic Society of Great Britain and Ireland (New Series)* 68(3). 401–416.
-  Yang, Daxue & Sangsrgyas Tshering & Gates, Jesse P. 2024. A voicing rule for non-continuant obstruents in Thebo Tibetan. *Language and Linguistics* 25(4). 710–746.
-  Yliniemi, Juha Sakari. 2021. A descriptive grammar of Denjongke. *Himalayan Linguistics Archive* 10. i–690.
- Zoller, Claus Peter. 1983. *Die Sprache der Rang Pas von Garhwal (Raṅ Pə Bhāsa): Grammatik, Texte, Wörterbuch*. Wiesbaden: Otto Harrassowitz.

Address for correspondence

Joanna Bialek
 Trinity Centre for Asian Studies
 Trinity College Dublin
 Room 2012, Arts Building
 College Green
 Dublin 2
 Ireland
 bialekj@tcd.ie
 <https://orcid.org/0000-0001-8542-4272>

Publication history

Date received: 25 June 2024

Date accepted: 11 March 2025

Published online: 26 February 2026