Classifying the Southern Tibetic Languages

Carl Bodnaruk

University of Sydney

Tibetic Languages

- Languages descended from Old Tibetic, following Tournadre and Suzuki (2023)

  - a.k.a. Bodic, Bodish, Central Bodish, Gangjung Bhoti

  - Not including East Bodish, West Himalayish, other languages that have sometimes been grouped under these labels

- Precise internal classification of the Tibetic languages remains difficult

  - Various sources generally agree on a set of broad groups, though do not propose specific relations between them

  - Of these, "Southern Tibetic" is a commonly proposed subgroup (Tournadre and Suzuki 2023: 472), although generally tentatively

  - Covers the Tibetic languages of Bhutan, Sikkim, and the Chumbi Valley

Southern Tibetic

- Bhutan: Dzongkha, Chocangaca, Dur Brokkat, Brokpa, Lakha**

- Sikkim: Denjongke

- China: Dromo

* No data is available for Lakha anywhere that I can find

Southern Tibetic cont.

- Dzongkha: National language of Bhutan, natively spoken in the Western half of the country

- Chocangaca: Spoken in eastern Bhutan by about 20,000 speakers (Tournadre and Rigzin 2015)

- Dur Brokkat: Spoken in a single village (Dur, central Bhutan) by a community of former yakherds

- Brokpa: Also spoken by yakherds, far eastern Bhutan, also called Merak-Sakteng (main villages)

- Dromo: Spoken in the Chumbi Valley, between Bhutan and Sikkim governed by China. Very limited data available, also called Groma.

- Denjongke: Historically main language of Sikkim, language of the Tibetan population in Tibet, also called Sikkimese, Lhoke

Problems with Southern Tibetic

- No internal classification has really been proven

    - GvD suggests that Chocangaca is a "sister" language to Dzongkha, and that Chocangaca speakers are descended from an offshoot of early Dzongkha* speakers

    - He claims to have sound changes to support this but does not give them anywhere that I can find (I admittedly have not asked him)

- Its status as a single subfamily has not really been proven

    - "It is not possible to list common phonological features to all the dialects of the S section." (Tournadre and Suzuki 2023: 561)

    - Features that do occur across the whole group also occur more widely in Tibetic languages

        - Are not shared innovations of the Southern group

Data Sources

- Chocangaca: Original data

  - Collected for a field methods class at Usyd, recorded by me and Gwen Hyslop

- Dzongkha: Hyslop et al (2025)

  - Has Dzongkha words by pronunciation and not only spelling, which can be archaic

- Dur Brokkat: Nishida (2014)

  - In Japanese, but with a short English-glossed wordlist

- Brokpa: Pema Wangdi (2021), Tshering Leki et al (2020)

- Denjongke: Yliniemi (2021)

- Dromo: Tournadre and Suzuki (2023)

  - Essentially no actual data from Dromo, but comments on similarities and sound changes

Evidence

- Good (I think) but very meagre
  - Problematised by a general lack of data
    - e.g. Lakha not able to be categorised, data on Dur Brokkat very limited
- Two sound changes
  - Neither is totally consistent, one is much more so

Reflex of Old Tibetan Cr clusters

- Old Tibetan had a set of plosive+/r/ clusters in onsets
  - Possible with all 3 voicing contrasts, 3 PoA (bilabial, alveolar, velar)
    - (and sr, mr, but I won't deal with them here)
  - In many contemporary varieties, all PoAs reduced to retroflex plosives
    - e.g. Lhasa [ɖɛ̀] < OT 'bras 'rice'
    - Preservation of Cr clusters is uncommon outside of Western Tibetic (Tournadre and Suzuki 2023)

Reflex of Old Tibetan Cr clusters

- Two patterns in Southern Tibetic:

  - Cr preserved in bilabials, reduced to retroflex in alveolar (Chocangaca, Brokkat, Brokpa)

    - OT dpral.ba > Cn prewa 'forehead', OT 'brog > Br,DB brok 'nomad', OT brag > Cn bra, Br pra 'cliff'

    - OT drug > Cn ɖu, Br ʈuk 'six', OT 'khrud > Cn ʈʰu, Br ʈʼ 'wash', OT khrag > Cn, DB, Br ʈʰa 'blood' (however, Dz ʈʰa, Dn kʰya), OT skra > Cn ʈa 'hair' (Br, DB ra̰ < East Bodish?)

  - Cr → Cy (merged with OT Cy clusters) for bilabial, velar; retroflex for alveolar (Dzongkha, Dromo, Denjongke)

    - OT dpral.ba > pyewa > Dz pcew 'forehead', OT brag > Dz bya, Dn pya 'cliff', OT 'khrud > Dz chu, Dn kʰju 'wash', OT skra > Dz ca 'hair'

    - OT dri.ma > Dz ɖim, Dn ʈʼim 'odour, smell'

    - No data for Dromo, but same as Denjongke for bilabials according to Tournadre and Suzuki (2023)

Reflex of Old Tibetan Cr clusters
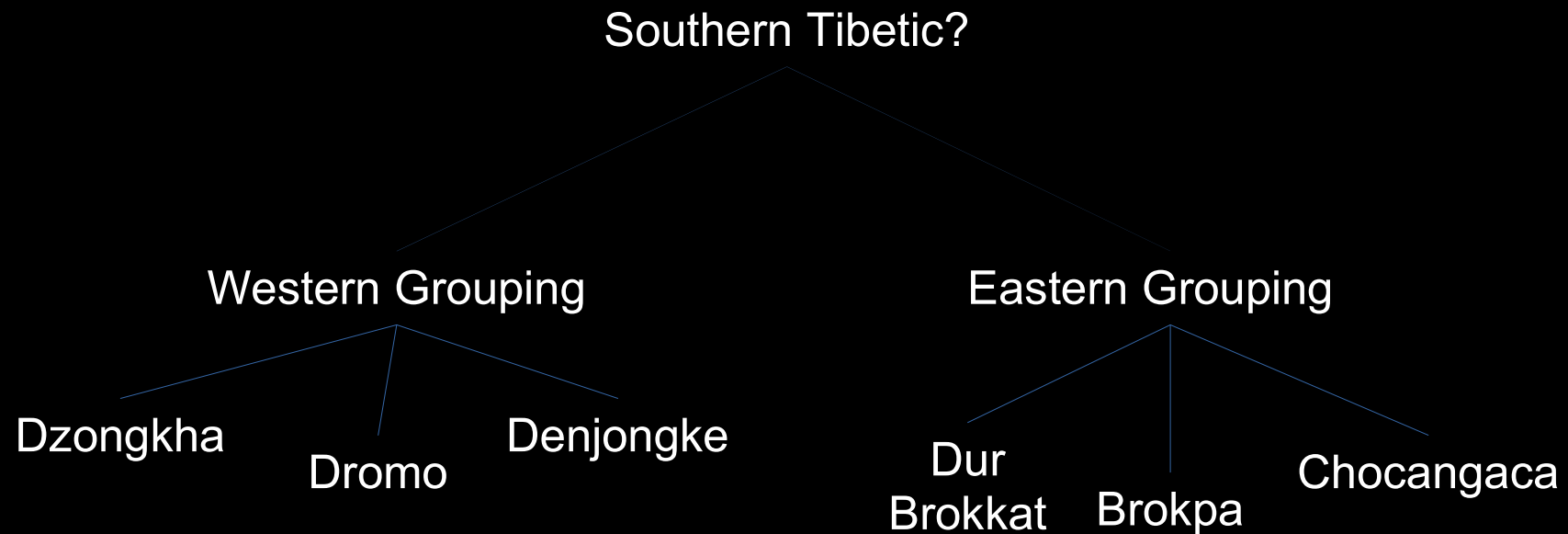
- Clear split between two groups
  - Consistent split with bilabial+r
  - Both retroflex with alveolar+r
  - Inconsistent with velar+r in Western group (sometimes becomes Cy, other times becomes retroflex)
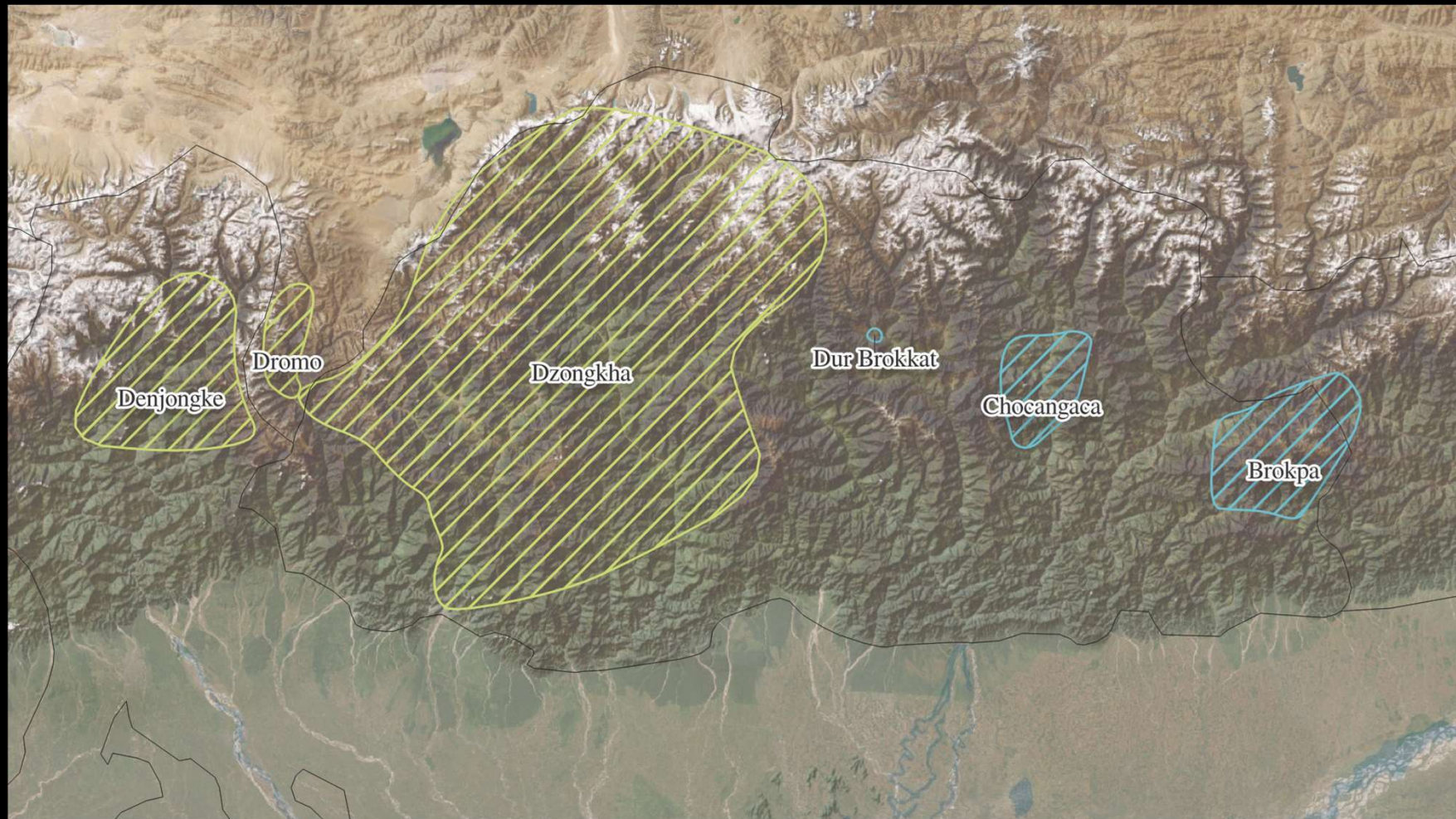
My initial proposal
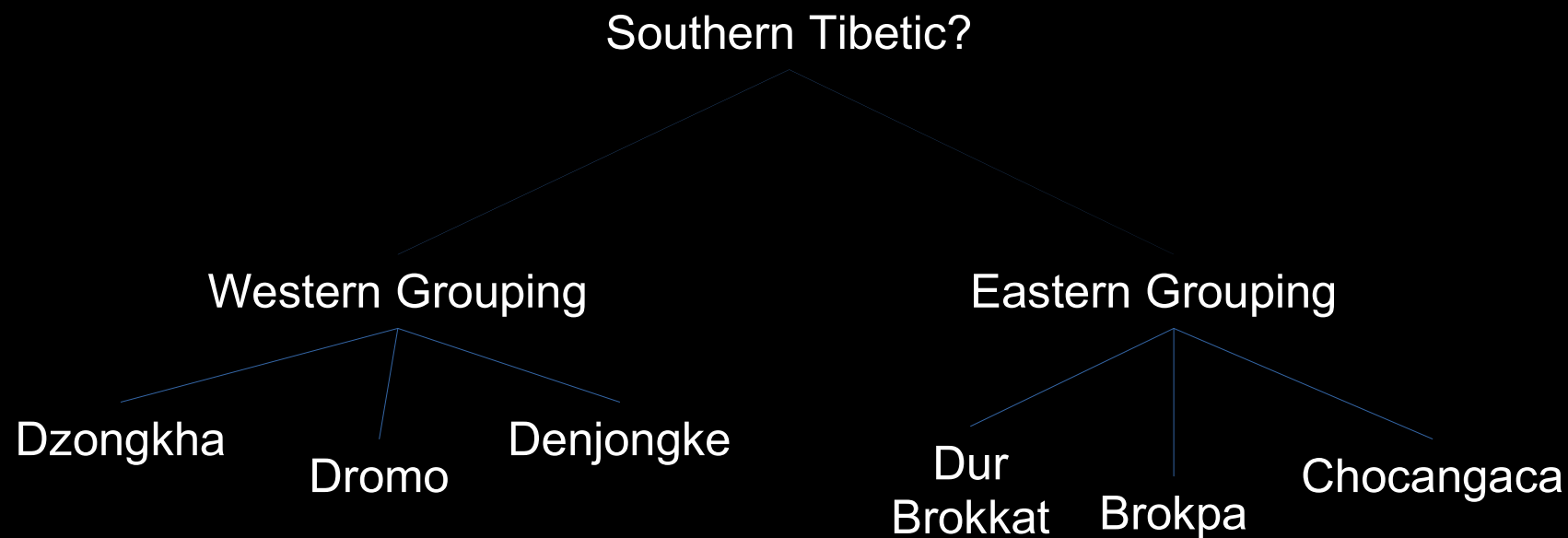
My initial proposal

- Two groups, Western and Eastern
- Western are more closely related, potentially a historical continuum
- Unclear if these two groups share an exclusive common ancestor

Southern Tibetic?

Western Grouping

Eastern Grouping

Dzongkha

Dromo

Denjongke

Dur Brokkat

Brokpa

Chocangaca

My initial proposal

Other evidence

- Of course, it's never that simple

- I have one other sound change that begins to support this
  - Dropping of final vowels is very common in Dzongkha, occasional in Denjongke
  - OT char.pa > Dz thap, Dn tsʰaːp, Cn tɕarpa 'rain'
  - OT dmar.po > Dz map, Dn maːp (marpo given as a song lyric however), Cn marpo 'red'
  - BUT OT sngon.po > Dz höm, Dn hompu/ŋ̊ompu, Cn ŋonpo

- I have, however, found some that complicate this

Reflex of Old Tibetan -d codas

- OT coronal codas are often deleted in Tibetic languages, preceeding vowel fronted
  - Specifically Southwestern, Central, Kham languages
  - Similar process to Germanic umlaut
  - Lhasa [dʑɛ̀~ɟɛ̀], Lhagang (Kham) [ʱdʑě̃ʔ] < OT brgyad 'eight'
  - Lhagang [tṹmu] < OT dudpa 'smoke'

Lhagang data: Suzuki and Sonam Wangmo (2016)

Reflex of Old Tibetan -d codas

- Consistently coalesced into preceeding vowel in Dzongkha, Denjongke, sometimes Brokpa
   - OT brgyad > gä (Dz), gɛʔ (Dn), gæ (Br) 'eight'
    - No Dromo data
- Preserved in Chocangaca (mostly), sometimes in Brokpa?
    - Nishida also gives 'eight' as gɛː
    - Cn dʑat 'eight', brat 'scratch' vs Br bræ < OT 'brad.pa
    - Cn dutpa, Br durpa, DB dau, Dz dupa, Dn tuwa 'smoke' < OT dudpa
    - Cn gatpu, Br gatpu, Dz gep, Dn gɛmpo/gɛːpo 'old man' < OT rgad.po
    - Cn ʈʰu, Br ʈý 'wash' < OT 'khrud Consonant deleted in both but sound change only occurs in Br
- Brokkat data is a little confusing here, as Nishida (2014) gives a sound change of OT skad > ke
    - But the language name clearly has skad > kat
    - Nishida says GvD gave Brokkat as the native name, but that he didn't come across it
        - Possible more recent change?
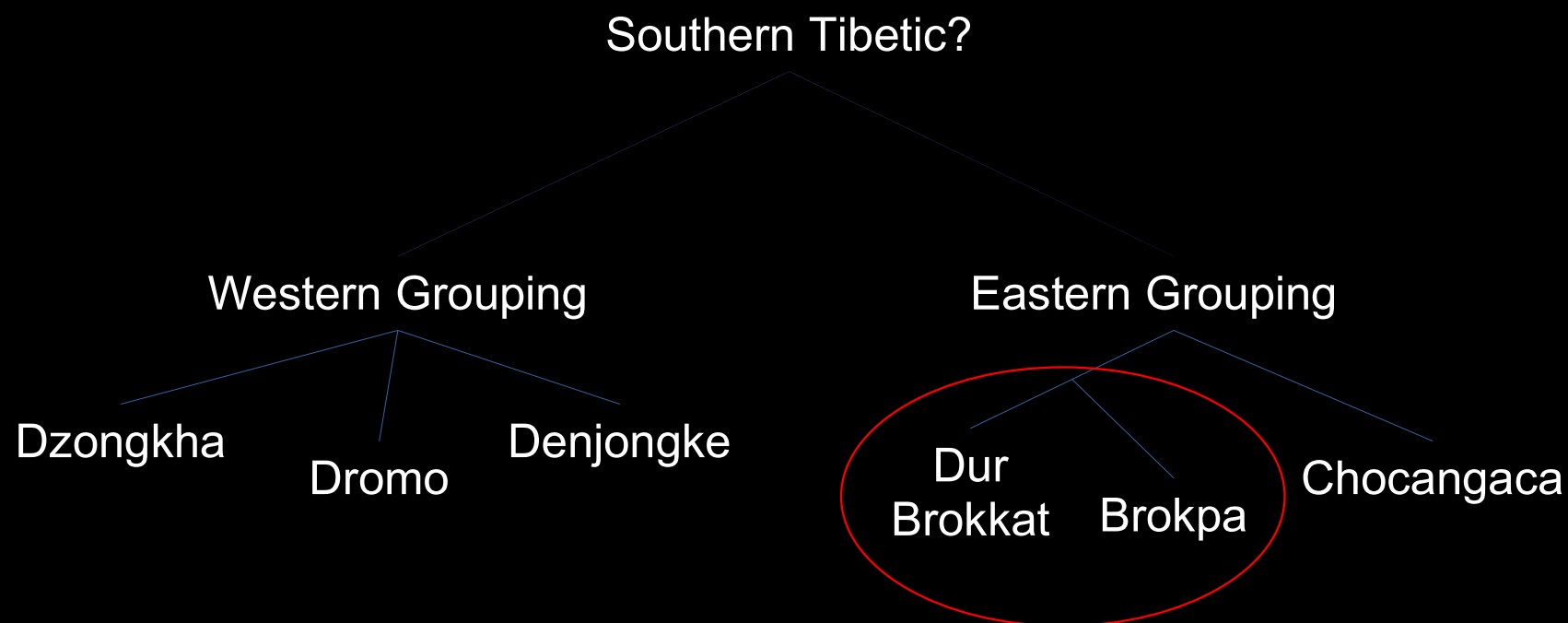
Reflex of pre-Tibetan -d codas

- While it's very unlikely that Cn would have regained -d after the branch lost it, it was never lost in East Bodish (Hyslop 2014)
    - PEB *gʲat > Kurtöp ɟat > Cn dʑat 'eight'?
    - Kurtöp brat > Cn brat 'scratch'? cf Br bræ, Dz ʈe
- Could Chocangaca have reborrowed them from Kurtöp (which is spoken in the same valley)?
    - Was retained in Proto-East-Bodish, has since been lost in some languages (probably under Tibetic influence?)
    - Challenges to this:
        - Cn has 2.SG pronoun tɕʰot < OT khyod, but PEB 2.SG is *i > Kp wi-

Evidence for Southern Tibetic

- Do we have any shared innovations across all of these languages that could be attributed to proto-Southern Tibetic?
  - None that I have found, but I am by no means an expert on all Tibetic varieties and sound changes, so I very well could have missed something
  - Could we be looking at two entirely separate branches?

Limitations

- Probably missing the bigger picture here
  - Looking at sound changes only within ST
    - Although I haven't systematically checked, none of the forms I've seen in ST compared to OT have stood out as potentially unique to ST
  - Fairly limited data
- Haven't looked at vocabulary systematically
  - Some roots are mixed throughout the region
    - e.g. 'forehead' is about 50/50 between OT dpral.ba and OT thod.pa 'skull', not following expected lines
    - 'night' is all over the place

References

- Hyslop, Gwendolyn, Kuenga Lhendup & Pema Chhophyel. 2025. *Kurtöp - Dzongkha - English Dictionary*. Thimphu: Centre for Bhutan and GNH Studies.

- Nishida, Fuminobu. 2014. A Phonology of Brokkat, a language spoken in the village of Dhur in Bumthang valley in central Bhutan. *アルテス リベラレス* 94. 1–17.

- Pema Wangdi. 2021. *A Grammar of Brokpa: a Trans-Himalayan language of Bhutan* (PhD Thesis). Cairns: James Cook University.

- Suzuki, Hiroyuki & Sonam Wangmo. 2016. Lhagang Tibetan of Minyag Rabgang Khams: Vocabulary of Two Sociolinguistic Varieties. *Asian and African Languages and Linguistics* 10. 245–286.

- Tshering Leki, Damian Funk, Pascal Gerber, Selin Grollmann, Corinne Mittaz, Sara Rüfenacht & Sereina Waldis. 2020. Brokpa texts, glossary and verb stems: Appendices to Aspects of Brokpa Grammar. *Himalayan Linguistics* 19(1).

- Tournadre, Nicolas & Karma Rigzin. 2015. Outline of Chocha-Ngacha. *Himalayan Linguistics* 14(2).

- Tournadre, Nicolas & Hiroyuki Suzuki. 2023. *The Tibetic Languages: an introduction to the family of languages derived from Old Tibetan*. Paris: LACITO-Publications.

- Yliniemi, Juha. 2021. A descriptive grammar of Denjongke. *Himalayan Linguistics* 20(1).

Appendix: 'night'

	Root 1	Root 2	Root 3	Root 4
Chocangaca				preme
Brokpa		zhaʔ	num	
Brokkat			numa	
Dzongkha	tshem	zha	numu	
Denjongke		ɕàʔ		
Old Tibetan	mtshan.mo	zhag	nub.mo?	
Kurtöp (East Bodish)	san?	zha		

Also Dn *pʰiːru* 'at night'