THE TRANS-HIMALAYAN COPULA #NI AND ITS SECONDARY FUNCTIONS

Scott DeLancey

University of Oregon
delancey@uoregon.edu

1. Trans-Himalayan copulas

Copula is an unstable category in Trans-Himalayan (Sino-Tibetan). There is no shared copula root attested in every subbranch, and many languages have innovative copulas which are not shared with other languages or branches. There are a few etyma which are widely attested, and have been proposed as reconstructable for the proto-language. The best-known such proposal is *way (Thurgood, 1982; Matisoff, 1985). On the STEDT website[1], Matisoff suggests several others besides *way (#450): *C-nə:ŋ (#79), *g-na-s (#84), *g-ray (#449), *s/g-ray ⋇ way (#1821), *s-ri-t ⋇ s-rut (#2608), *nay (#5741), *ley (#5755), and *duŋ (#5789). Lowes (2006) presents evidence for #yoC, #d/tuC, #d/ta(C), #ni, and #la, all attested in several different subbranches.[2]

The large number of forms which are attested with copular function in one or another language suggests widespread replacement of original copulas with innovative forms. And many other attested languages have copulas which cannot be ascribed to any of these etyma. Several copular roots seem to be reconstructable at deeper levels. Thurgood and Matisoff have presented evidence for *way and a few others. In this paper I will present further evidence for the Proto-Trans-Himalayan (PTH) provenance of Lowes' #ni, which is mostly but not entirely equivalent to Matisoff's *nay, concentrating on evidence from secondary grammaticalizations of its original copular function.

First we will quickly look at evidence for another well-attested example, which will be peripherally relevant to our main topic. This is an innovative copula in the Eastern branch of the family, which Jacques and Pellard reconstruct as *ŋwa.

> [T]he affirmative copula found in Rgyalrongic languages (Japhug *ŋu*, Khroskyabs *ŋɛ́*, Lai 2017, p. 247) is related to proto-LB *ŋwa[1] 'be the case' (Bradley, 1979, #698). This verb was originally an adjectival stative verb 'be true', a meaning still marginally preserved in Japhug (Jacques, 2014a, p. 61). No verbal cognate is found outside of Burmo-Rgyalrongic, but Tibetan *ŋo.bo* 'true nature' shares the same root. Rgyalrongic and Lolo-Burmese here share a unidirectional semantic innovation 'be true' > copula (Jacques & Pellard, 2020, p. 16)

This root is well-attested across the Eastern or Burmo-Qiangic branch; some evidence is presented in Table 1.

Table 1: A sampling of evidence for *ŋwa in Eastern Trans-Himalayan

Burmese-Ngwi	PLB	*ŋwa[1] 'be the case'
	Nuosu	*ŋe*
	Youle Jino	*ŋu*[55]
	Zaiwa	*ŋvut*[5]
Qiangic	Northern Qiang	*ŋuə*
	Munya	*ŋo*
Rgyalrongic	Kyomkyo Situ	*ŋos*
	Zbu	*ŋoʔ*
	Khroskyabs	*ŋɛ́*
	Geshiza	*ŋuə*

Evidence for *ŋwa outside of the Eastern branch is sparse. Jacques and Pellard mention Tibetan ངོ་བོ *ŋo.bo* 'true nature', and another possible Tibetan

---

[2] I use # to mark tentative reconstructions whose phonological detail remains to be demonstrated. Reconstructed forms marked with * are taken from STEDT or another cited source.

comparator is དངོས *dngos* 'reality, existence', but although the semantic connection is plausible, these words are nouns, not copulas. Better comparators are Kadu (Asakian) *ngā* 'existential/locational' (Sangdong, 2012, pp. 226-228) and Jinghpaw *ŋà*, 'exist', also marking continuous aspect (Kurabe, 2016, p. 285), and perhaps also the second element of Baram (Newaric) *ŋiŋa*, 'possessive/locative' and its negated form *mə-ŋa* (Kansakar et al., 2011, p. 142). The Jinghpaw-Asakian forms could plausibly represent borrowing from Burmic, in which case we can follow Jacques and Pellard and assume that the shift of *ŋwa to copula function is an Eastern innovation.

The reasons for the instability of copula forms across the family lies in their syntactic behavior and consequent pragmatic associations. In many languages quational copulas are "optional", meaning present only with contrastive force. See for example Zakaria's (2017, pp. 614-619) meticulous description of the equational "verbless clause" construction for Hyow Chin, which seems to be typical of South Central languages (see e.g. Chhangte, 1993, p. 94; Mroueh, 2019, p. 124), although most available descriptions don't explore the question. For this reason they are easily replaced, since a speaker wanting to speak emphatically can always find a more emphatic substitute ('really', 'it's true', etc.) for the copula. At the same time the original copula is easily reinterpreted as a stance-marking particle (§3). The fact that there is lexical evidence for a substantial number of forms with copular function, as we saw above, shows that this is a recurrent path of grammaticalization in the family and has been since the proto-language.

In this paper I will summarize the comparative evidence for #ni as a PTH copula (§2), and briefly review examples of grammaticalization of #ni connected with polarity and stance (§3). The largest contribution of this paper is to document the further grammaticalization of #ni into an information structure marker; this is the topic of §4.

2. Comparative evidence for copula #ni

Here I will summarize the case presented in DeLancey (to appear) for reconstructing #ni as a copula in Proto-Trans-Himalayan (PTH). We are potentially interested in two of the roots from STEDT mentioned in the previous section: #84 *g-na-s and #5741*nay. The meanings of the words listed under *g-na-s are quite lexically specific, usually 'rest', so any explicitly copular function must be secondary and not reconstructable. But *nay* is presented as the root for several forms that we will consider as evidence for #ni. Matisoff reconstructs *nay*, presumably to account for variation in the vowel, but the commonest reflex all across the family is /ni/. More importantly, the consistent attestion of a monophthongal /i/ rime does not fit any of the correspondence patterns identified by Matisoff (1985). These correspondences are presented in Table 2.

Table 2: Rime correspondences according to Matisoff (1985)

PTB	Written Tibetan	Jinghpaw	Written Burmese	Dimasa	Lushai
*-ay	-e	-ai	-ai	-ai	-ei
*-a:y	-e	-ai	-ai	-ai	-ai
*ey	-e	-i	-i	-ai	-ei
*-əy	-i	-i	-e	-i	-i
*-i	-i	-i	-i	-i	-i
#ni	ni	ni	nei	ni	nii / nei

Not all of the forms presented here as reflexes of #ni are demonstrably so: The Mizo, Dimasa, and Burmese forms are copulas or copula-adjacent verbs, but the Jinghpaw form is a stance marker (§3) and the WT form an information structure marker (§4). We can infer that both of these are reflexes of the copula, but this cannot be proven unequivocally. But even comparing only the Mizo and Dimasa forms we can see that they do not fit Matisoff's correspondence pattern for *-ay or *-a:y. Moreover, there are contrasting forms we find contrasting simple and diphthongal forms in Mizo, there must be two different etyma. For now, then, I will reconstruct #ni, and not consider diphthongal forms such as Burmese *nei*. The remaining forms listed in STEDT under #5741 *nay 'copula' are presented in Table 3.

Table 3: Examples from STEDT #5741 *nay 'copula/be' with monophthongal rimes

Branch	Subbranch	Language	
Eastern	Qiangic	Muya	*ni⁵³*
	Na	Naxi	*ɲi*
	Loloish	Nusu	*ni³¹*
	Burmish	Achang (LX)	*ŋi³¹*
Central	Ao	Yacham	*nyi*
Western	Bodish	Tshona	*ni³⁵*
	Lepcha	Lepcha	*nyí*

To this evidence we can add a clearly reconstructable West Himalayan etymon reflected in Bunan *ni* 'existential copula' (Widmer, 2014, pp. 585-586) and Darma *ni* 'equational copula' (Willis, 2007. pp. 335-336). (In the Western subbranch, of West Himalayan, including Tinan and Kinnauri, this has been replaced by an innovative form #*to*).

3. Grammaticalization of #ni as a verbal or sentential operator

It is well-known that copulas often grammaticalize into tense/aspect/modality markers in Tibeto-Burman languages, due to the strong tendency across the family to reinterpret nominalized clausal constructions as finite (Noonan, 1997, 2008; Genetti et al., 2008; Genetti 2013; DeLancey 2011; inter alia). For example, several languages have a progressive aspect marker that seems relatable, e.g. Uipo ("Khoibu") *-nei* (Singh, 2014, p. 122), Sümi (Sema) *ani* (Teo, 2019). Below we will see evidence for its reanalysis as a Nonpast marker in Kiranti. Such forms serve only as indirect evidence for #ni as a copula in the proto-language, but that hypothesis offers the most direct explanation for them.

Copulas, especially "optional" copulas that automatically have some contrastive sense when they occur in an affirmative clause, frequently develop into markers of "stance", i.e. evidentiality, speaker's evaluation, mirativity, etc. (Noonan, 1997; DeLancey in press). Both of these paths are attested across Trans-Himalayan, and we have examples of #ni following them in many TB languages. For example Akha (Burmese-Ngwi) has two stance-marking particles which look very familiar: *ni˘*, which is "added to a statement to make it very emphatic, something like "… of

course". Used in both positive and negative declarative statements." (Lewis, 1968, p. 221) and *nga˘*, potentially reflecting the root which we looked at in §1, which marks "a declarative statement about another person, an emphatic statement to a person (esp. concerning what he must do), an emphatic answer to a yes or no question, an emphatic negative statement, and in emphatic questions asking for an explanation" (Lewis, 1968, p. 229). Further examples are discussed in DeLancey (in press).

Finally, in a few TB languages, we see an unusual development of #ni into a marker of negation (Auwera & Vossen, 2017; DeLancey in press). The "optionality" or contrastive sense of the equational copula in most TB languages applies only to affirmative sentences. In most languages negation is marked by a verbal affix, and thus a negated equational sentence requires a copula for the negator to attach to. Thus even when a copula is grammaticalized into a stance marker, it may remain in relict negated constructions (DeLancey, 2022). An example is Dura (Central Himalayan), which has distinct affirmative and negative roots for both equational and existential copulas, as shown in Table 4.

Table 4: Dura affirmative and negative copulas (Schorer, 2016, see also Nagila, 2013)

	AFF	NEG
EQUATIONAL	*le*	*ma-pi*
EXISTENTIAL	*po*	*mu-**ni***

Of interest to us is the form *mu-ni*, where we see an evident reflex of #ni occurring only with the negative affix.

A striking further development of this association is seen in Kiranti, where in a few languages #ni has completely replaced PTH *ma-* as the primary negative marker (van der Auwera & Vossen, 2017; DeLancey in press). We can see the progression of this shift in Table 5.

Table 5: First person singular inflected forms in four Kiranti languages (DeLancey in press)

1SG	Sunuwar	Kulung	Dumi	Athpare
Aff past	Σ-ta	Σ-o	Σ-a	Σ-a-ŋ-e
Aff nonpast	Σ-nu-ŋ	Σ-o:	Σ-t-a	Σ-ŋa-ʔ-a

Neg nonpast	ma-Σ-nu-ŋ	Σ-o:-no	Σ-t-nə	Σ-ni-ŋ-na
Neg past	ma-Σ-du	man-Σ-ŋa	mə-Σ-nə	Σ-nat-ni-ŋ-na

We see that *-ni/-nu/-nə* forms begin as a Nonpast tense marker, as seen in Sunuwar. (I follow van der Auwera and Vossen in linking all three forms with *#ni*; although the phonological correspondence is not perfect, there is no other evidence source for any of the forms, and it seems highly unlikely that they started from different sources and converged on a negative function). In Kulung we see the form acquiring specficially negative sense, limited to the negative Nonpast. Then Dumi presents a further development where *-nə* occurs in all negated forms. Finally in Athpare the innovative negator has completely replaced the older prefix *#ma-,* and become the sole marker of negation.

4. Further grammaticalization of #ni as an information structure marker

Another type of evidence for copular *#ni*, still less direct than the evidence discussed in §3, is found in the form of Information Structure Markers (ISM's). All Tibeto-Burman languages have ISM's which follow all other clitics and postpositions at the end of the noun phrase; see for example Konnerth (2014); Boro (2021). These are an important part of the grammar of any Tibeto-Burman language, but they are often neglected in linguistic descriptions. When they are described they are generally given glosses such as 'topic', 'focus', 'emphatic', and 'contrastive'. There is a well-known grammaticalization path which leads from a copula to an information structure marker, through a cleft construction *It is X that ...* (for examples in Sinitic see Jin, 2020). Widmer (2017) and others have noted the similarity between various reflexes of the *#ni* copula and similarly widespread "topic" or "focus" markers of the same form, for example Tibetan ནི *ni*. Like the copulas, these are found across the family.

What ISM's do is to point to some relation between the marked NP and the discourse context. The commonest function, contrast, is illustrated in examples (1) through (5) from several different languages. The best-known example is probably Written Tibetan ནི *ni*.

(1) a. *rdo.bdag-gis rta ma-btang zhes*
horseman-ERG horse NEG-send QUOT

smras.pa-s lce chod=cig
said-because tongue cut.off=IMPER
'Because the horseman said "Don't let the horse get away", cut off his tongue.'

b. *dbyug.pa.can=**ni** rdo 'phangs.pa-s*
Yugpacan=FOC stone threw-because

lag.pa chod=cig
hand cut.off=IMPER
'As for Yugpacan, because he threw a stone, cut off his hand.' (Jäschke, 1954, p. 96)

Here the ISM *=ni* in (1b) serves to indicate the change of topic from the previous sentence (1a); the speaker (the King) is adjucating a dispute between the horseman and Yugpacan; having dealt with the former in (1a), and he now turns his attention to Yugpacan in (1b), and uses *=ni* to make sure the hearer shifts attention. The form is retained in the same function in Modern Standard Tibetan, as in example (2).

(2) *nga=**ni** yong-gi-min*
I=**TOP** come-IMPF-NEG.EGO
'For my part, I'm not coming.' (Denwood, 1999, p. 103)

A similar-looking 'focus clitic' in Hyow (Southwestern South Central/Kuki-Chin) indicates "'only' or 'just' or of the cleft construction 'it is the X', restricting the reference of the noun phrase" (Zakaria, 2017, p. 172), as in example (3).

(3) *èydɔ̂ shɔ́=**ní** ká-á-ní-hyúl-dú ɔ́lɔ̂*
then wild.pig=**FOC** 1A-DIR-follow2-ITER again
'Then we followed *the wild pig* again.' (Zakaria, 2017, p. 172)

Here the ISM reminds the hearer that the wild pig is being set apart in the discourse from other referents which "we" might have been following.

In Poumai Naga (Angami-Pochuri) a "particular" marker *=ni* indicates that the marked noun is unexpected here, or contradicts some prior expectation.

(4) a. *pwu-nì   swi-jè   swuthà-ré*
    PN-ERG   dog-PAT kill-PERF
    'Pwu killed the dog.' (Veikho, 2021. p. 203)

  b. *pwu-nì   swi-ni   swuthà-ré*
    PN-ERG   dog-PART kill-PERF
    'Pwu killed the dog (given two or more options of something to kill).' (Veikho, 2021, p. 203)

    OR

    'Pwu killed the dog (contrary to the presupposition that he did not).' (Veikho, 2021, p. 204)

For one more example, consider Bodo (Bodo-Garo) =*nw* "corrective", which likewise indicates that the marked NP is to be considered instead of or unexpectedly in addition to some presupposed referent.

(5) *bé belà-ao   boró raù-a=lo*
    this time-LOC Bodo language-SU-REST

    nòŋ-a bʰarot-ni gasìbu raù-a=nuu COP-
    NEG India-GEN all   language-SU-COR

    pʰelèŋ-ni   lamà-ao
    die-GEN   path-LOC
    'It is not just the Bodo language, but all languages of India are on the path of extinction.' (Boro, 2021, p. 90)

Here the marking of *bʰarot-ni gasìbu raù-a* 'all languages of India' with =*nw* serves to contrast that reference with the previous reference specifically to Bodo.

Information structure markers relatable to #ni occur across the family; a sampling is presented in Table 6.

Table 6: Information structure markers which could be related to #*ni*

Branch	Subbranch	Language	ISM
Eastern	Loloish	Khatso	*ni* 'topic'
	Qiangic	Pumi	*ni* 'additive'
		Guiqiong	*ni* 'emphatic'
Central	SoCentral	Hyow	*ni* 'focus'
	Angami	Poumai	*ni* 'particular marker'
	Bodo-Garo	Bodo	*nu*

			'contrastive'
Western	Bodic	Tibetan	*ni* 'topic'
	Kiranti	Thulung	*ne* 'topic'

Distributively speaking there is ample evidence here to reconstruct an ISM #ni for PTH, but this can slightly more economically be interpreted as evidence that #ni was the primary equational copula in the proto-language, and occurred in cleft constructions. On either interpretation the widespread attestation of /ni/ as an ISM across the family is additional evidence for the antiquity of the copular root.

5. Summary

The evidence for copula #ni at the PTH level, like that for other proposed copula roots, is scattered across the family. There is no root, including #ni, which is robustly attested as a copula in each branch. Additional support for reconstructing it as a copula comes from resemblant grammatical forms, attested throughout the family, which have functions that are known to sometimes develop from further grammaticalization of copulas, particularly sentential stance markers and information structure markers which cliticize to NP's. Combining these various types of evidence allows us to construct a more convincing case for the PTH provenance of the copula #ni.

Sources for data in tables: Achang, Dai & Cui (1985); Bodo, author's notes; Geshiza, Honkasalo (2019); Guiqiong Li (2015); Hyow, Zakaria (2017); Khatso, Donlay (2019); Khroskyabs, Lai (2017); Kyomkyo Situ, Prins (2016); Lepcha, Plaisir (2007); Munya, Bai (2019); Naxi, Lidz (2010); Northern Qiang, Huang (2004); Mizo, Chhangte (1993); Poumai, Veikho (2019); Pumi, Daudey (2014); Thulung, Lahaussois (2004); Tibetan, Jäschke (1954), Denwood (1999); Tshona, Lu (1986); Youle Jino, Hayashi (2014); Zaiwa, Lustig (2010); Zbu, Gong (2018).

References

Auwera, J. van der, & Vossen, F. (2017). Kiranti double negation: A copula conjecture. *Linguistics of the Tibeto-Burman Area,* 40(1), 40-58.

Bai, J. (2019). *A grammar of Munya* [PhD dissertation, James Cook University].

Boro, K. (2021). Focus enclitics in Bodo. *Linguistics of the Tibeto-Burman Area* 44(1), 76-112.

Bradley, D. (1979). *Proto-Loloish*. Curzon.

Chhangte, L. (1993). *Mizo syntax* [Doctoral dissertation, University of Oregon].

Dai, Q., & Zhichao, C. (1985). *Achang yu jianzhi* [Outline of Achang]. Minzu Chubanshe.

Daudey, H. (2014). *A grammar of Wadu Pumi*. [Doctoral dissertation, LaTrobe University].

DeLancey, S. (2011). Finite structures from clausal nominalization in Tibeto-Burman languages. In F. H. Yap, K. Grunow-Hårsta, & J. Wrona (Eds.), *Nominalization in Asian languages: Diachronic and typological perspectives* (pp. 343-62). John Benjamins.

DeLancey, S. (in press). Copula to negator: grammaticalization of the copula #ni in Tibeto-Burman languages. *Studies in Language*.

DeLancey, S. (to appear). The Tibeto-Burman copula #ni. *Bulletin of Chinese Linguistics*.

Denwood, P. (1999). *Tibetan*. Benjamins.

Donlay, C. (2019). *A grammar of Khatso*. De Gruyter Mouton.

Genetti, C. (2013). Tense-aspect morphology from nominalizers in Newar. In T. Thornes, E. Andvik, G. Hyslop, & J. Jansen (Eds.), *Functional-historical approaches to explanation* (pp. 195-220). Benjamins.

Genetti, C., Bartee, E., Coupe, A., Hildebrandt, K., & Lin, Y.-J. (2008). Syntactic aspects of nominalization in five Tibeto-Burman languages of the Himalayan area. *Linguistics of the Tibeto-Burman Area,* 31(2), 97-143.

Gong, X. (2018). *Le rgyalrong zbu, une langue tibéto-birmane de Chine du Sud-ouest. Une étude descriptive, typologique et comparative* [PhD dissertation, Université de Paris].

Hayashi, N. (2014). Youle Jino adjectives and their semantic mapping. *Journal of Foreign Studies,* 64(3), 9-22.

Honkasalo, S. (2019). *A grammar of Eastern Geshiza* [Doctoral dissertation, University of Helsinki.

Huang, C. 黄成龍. (2004). *A reference grammar of the Puxi variety of Qiang* [Doctoral dissertation, City University of Hong Kong].

Jacques, G. (2014). *Esquisse de phonologie et de morphologie historique du tangoute*. Brill.

Jacques, G., & Pellard, T. (2020). Phylogenies based on lexical innovations refute the Rung hypothesis. *Diachronica,* 38, 1-24. ff10.1075/dia.19058.jacff. ffhal-03084829f.

Jäschke, H. A. (1954). *Tibetan grammar* (2nd ed.). Frederick Ungar.

Jin D. (2020). Copula functions in a cross-Sinitic perspective. *Folia Linguistica,* 54(1), 89-132.

Kansakar, T. R., Yadava, Y., Chalise, K. P., Prasain, B., Dhakal, D. N. & Paudel, K. (2011). *A grammar of Baram*. Central Department of Linguistics, Tribhuvan University.

Konnerth, L. (2014). Additive focus and additional functions in Karbi (Tibeto-Burman) =tā. In Carpenter, K., David, O., Lionnet, F., Sheil, C., Stark, T., Wauters, V. (Eds.), *Proceedings of the 38th Annual Meeting of the Berkeley Linguistics Society* (pp. 206-222).

Kurabe, K. (2016). *A grammar of Jinghpaw, from northern Burma* [Doctoral dissertation, Kyoto University].

Lahaussois, A. (2004). Thulung Rai. *Himalayan Linguistics Archive* No. 1.

Lai, Y. (2017). *Grammaire du khrosyabs de Wobzi* [Doctoral dissertation, Université Sorbonne Nouvelle Paris 3].

Lewis, P. (1968). *Akha-English dictionary*. Department of Asian Studies, Cornell University.

Li, J. (2015). *A grammar of Guìqióng*. Brill.

Lidz, L. (2010). *A descriptive grammar of Yongning Na (Mosuo)* [Doctoral dissertation, University of Texas].

Lu, S. (1986). *Cuona Menba yu jianzhi (Outline of Cuona Monpa)*. Minzu Chubanshe.

Lustig, A. (2010). *A grammar and dictionary of Zaiwa*. Brill.

Matisoff, J. (1985). God and the Sino-Tibetan copula, with some good news concerning selected Tibeto-Burman rhymes. *Journal of Asian and African Studies*, 29, 1-81.

Mroueh, J. (2019). Tedim verb person marking. *Himalayan Linguistics,* 18(1), 119-133.

Nagila, K. B. (2013). Negation: Evidence from the Dura language. *Nepalese Linguistics* 28, 95-99.

Noonan, M. (1997). Versatile nominalizations. In *Essays on language function and language type.*

In J. Bybee, J. Haiman & S. A. Thompson (Eds.), *In Honor of T. Givón*, 373-94. John Benjamins.

Noonan, M. (2008). Nominalizations in Bodic languages. In M. J. López-Couso & E. Seoane (Eds), *Rethinking grammaticalization: New perspectives* (pp. 219-37). John Benjamins.

Plaisir, H. (2007). *A grammar of Lepcha*. Brill.

Prins, M. (2016). *A grammar of rGyalrong Jiǎomùzú (Kyom-kyo) dialects*. Brill.

Sangdong, D. (2012). *A grammar of the Kadu (Asak) language* [Doctoral dissertation, University of Melbourne]

Schorer, N. (2016). *The Dura language: Grammar and phylogeny.* Brill.

Singh, L. B. (2014). *A descriptive grammar of Khoibu* [Doctoral dissertation, Assam University].

Teo, A. (2019). *Investigating differential case marking in Sümi, a language of Nagaland, using language documentation and experimental methods*. [Doctoral dissertation, University of Oregon].

Veikho, S. L. (2021). *Grammar of Poumai Naga (Poula)*. Brill.

Thurgood, G. (1982). The Sino-Tibetan copula *wəy[(o)]. *Cahiers de linguistique Asie Orientale* 11(1), 65-81.

Yliniemi, J. (2016). Attention marker =ɛo in Denjongke (Sikkimese Bhutia). *Linguistics of the Tibeto-Burman Area* 39(1), 105-160.

Zakaria, M. (2017). *A grammar of Hyow* [Doctoral dissertation, Nanyang Technological University].