

Nominalization, Relativization, Genitivization in Thulung Rai

Author(s): Aimée Lahaussois

Proceedings of the Twenty-Eighth Annual Meeting of the Berkeley Linguistics Society: Special Session on Tibeto-Burman and Southeast Asian Linguistics (2002), pp. 87-98

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

Nominalization, relativization, genitivation in Thulung Rai

AIMÉE LAHAUSSOIS
LACITO/CNRS, Paris, France

Matisoff showed in his seminal 1972 article that the functions of nominalization, relativization and genitivation are intimately related in Lahu: they are marked by the same particle *ve*. This is shown with a few illustrative examples.

Genitive

- (1) *ŋà ve mí-cho*
I shoulder-bag
'my shoulder-bag'

Relativization

- (2) *và? qhe chu ve Píchə-pā ô te yâ*
pig as fat Shan that one person
'That Shan over there who's fat as a pig'

Nominalization

- (3) *ə-ši tə? la ve thà? nə mâ ɣa mə lâ*
blood emerge come ACC you NEG get see Q
'Didn't you see that blood was coming out?'

In Thulung, there is a set of markers which cover these same functions and which also look etymologically related, although they are not phonologically identical: the class of markers is *-m*, *-mu*, *-mim*, *-ma*.

It is my goal to describe the use of these markers, as they relate to the three functions of nominalization, relativization, and genitivation in Thulung. While the pattern is not as neat as in languages like Lahu, the Thulung data shows participation in what has been called the Standard Sino-Tibetan Nominalization pattern ("SSTN", Bickel 1999).

Glossing these markers is problematic, because they are a set, rather than a single marker, and also because they often participate in more than one function. I do not have evidence of grammaticalization in any direction¹, nor can I tell at this point whether this pattern is converging towards or diverging from the SSTN. In light of this, I have decided to label the markers according to what appears to be their dominant usage. The marker *-mu* is only used to nominalize, and as such I label it NOM. *-m* appears most commonly as a relativizer, thus it merits the label REL. *-mim* is an alternative to *-m*, with a more restricted distribution, and REL2 is an appropriate label (for simplicity, this label is kept even in cases where it is used for nominalization); *-ma* is discussed, under genitivization, as part of this set of markers, because of its presence on an alternative set of possessive pronouns, but it is not synchronically an independent morpheme, so it receives no label.²

Relativization

Thulung has externally-headed relative clauses, most often preposed to head. There are two relativizers, *-m* and *-mim*, with *-m* having a more general distribution (used for both past and non-past), while *-mim* can only be used for non-past clauses.³

- (4) [*go khok-to-m/*-mim*] dzam brəpa bai-ra
 1s cook-1s/3s.PST-REL/REL2 rice good be-3s.PST
 ‘The food I cooked was good.’
- (5) *wa-lwak-mim-ka* [*makai py-ry-m/*-mim*]
 o.brother-y.brother-PLU-ERG corn eat-3s/3s.PST-REL/REL2
 bwa bre-mri.
 pig buy-3p/3s.PST
 ‘My brothers bought a pig that ate corn.’

¹ This is unlike the situation in Chantyal, where Noonan is able to convincingly claim grammaticalization of the nominalizer into the marker for the other functions.

² Abbreviations used in the examples are as follows: verb agreement encodes both agent and patient, and I represent the combinations separated by a /, so that verbal agreement showing 1s acting on 3s is abbreviated as 1s/3s; POL preceding a pronoun indicates that it is the polite form; Other abbreviations are GEN=genitive, INSTR=instrumental, ERG=ergative, DAT=dative, ABL=ablative, LOC=locative, TOP=topic, CONTR=contrastive, OBL=obligation, NEG=negation, IRR=irrealis, SEQ=sequencer, CONV=converb, PROG=progressive, CAU=cause, PURP=purposive, HS=hearsay, DU=dual, PLU=plural.

³ There is an alternative relativization strategy, with participles *-pa* and *-ma* (for non-past and past clauses respectively). One difference from finite-relativized clauses is that the participles are non-finite and therefore do not encode participant information, unless separately specified by an independent pronoun

- (6) [go dwak-pu-**m/-mim**] kitap gani-lai gwa-gwa dwa.
 1s like-1s/3s-REL/REL2 book POL.2p-DAT give-give like.3s/3s
 ‘I want to give you a book I like.’⁴
- (7) oram je [go phəŋtasi-ŋu-**m/-mim**] hapa dzupa bu.
 this clothes 1s wear-1s-REL/REL2 very nice be.3s
 ‘These clothes I am wearing are very nice.’

Additionally, *-m* is blocked in certain instances, and what is relevant seems to be the syllable length of the finite verb. Mono-syllabic verb forms block the use of relativizer *-m*.

- (8) gui si-**mim**/***-m** din
 1pi die-1pi-REL2/REL day
 ‘the day we die’
- (9) gu-ka ra-**mim**/***-m** bela
 3s-ERG say-3s/3s-REL2/REL time
 ‘At the time when she says...’

The synchronic distribution of the relativizers thus appears to be as follows:
-m is the general relativizer, available for both past and non-past clauses (except for mono-syllabic verb forms.)
-mim must be used with mono-syllabic verb forms and can be used anytime a non-past form is being relativized.

This distribution may appear to be fairly unusual, and earlier data on Thulung shows that this was not always the case. Allen, who worked on Thulung in the 1970’s, noted that tense was relevant for the choice of relativizer: “It would seem that *mim* is to present tense forms what *-m* is to past tense ones”. (1975: 88). So the earlier distribution of relativizers seems to have been exclusively based on tense. Allen does however give one example of a non-past clause marked with *-m*⁵, suggesting that the seeds of change were already present in the 1970’s.

The current distribution whereby mono-syllabic verb forms must take the relativizer *-mim* seems to be a shift in the interpretation of the relevant factor for relativizer choice. The only mono-syllabic verb forms in the language are non-past⁶, and it is possible that speakers reinterpreted the relevant factor as being not tense but syllable-length.

⁵ mi theŋpum loa koNNa reakpu
 not I-understand+m words only I-write

⁶ This is the case for verbs whose agent is a first plural exclusive or a third person singular.

Nominalization

There are a number of constructions which are counted as examples of nominalization, and they are discussed below. Interestingly it is not the same marker which appears on all of them, which may suggest different paths of development (in other words some of these nominalized forms may have developed from relativized constructions while others may have been nominalized from the start).

The first two types of nominalization are typical nominalized constructions throughout Tibeto-Burman, and they both take the same nominalizer, *-mu*.

Citation form of verbs

Matisoff states that “as a general rule of thumb applicable throughout the Tibeto-Burman family, whenever one discovers the particle used in verb citation, one can be sure of having discovered the most important nominalizer of the language.” (1972: 248). Thulung uses the marker *-mu* for this function.

- (10) *on-mu-lai tsapa bəne-mu basi*
 run-NOM-DAT strong make-NOM OBL
 ‘To run, one must make oneself strong.’
- (11) *kho-mu-kam lagi...*
 cook-NOM-GEN sake
 ‘In order to cook...’
- (12) *lamdi-mu bhanda-ne plen-ra lo-mu thik ra-ɣro*
 walk-NOM than-TOP plane-LOC go-NOM fine say-1s/3s.PST
 ‘I said that going by plane is better than walking.’

It is the nominal status of these roots which allows the verb to take on case marking.

Verb complementation

Similarly, verb complementation, where the nominalized clause is the complement of the verb, is marked with *-mu*.

- (13) *go dika [mukli lo-mu] tsahebe-u.*
 1s tomorrow Mukli go-NOM need-1s/3s
 ‘I must go to Mukli tomorrow.’
- (14) *go [dzudzuuluj ho-mu] dwak-pu.*
 1s mountain climb-NOM like-1s/3s
 ‘I like to climb mountains.’

There are also some instances of less prototypically “nominal” constructions which are nevertheless examples of nominalization. All of these use *-m* as a nominalizer (sometimes in complementary distribution with *-mim*, when monosyllabic forms are possible.) In other words, the nominalizer in these cases is what we have seen to be the relativizer in Thulung, indicating that perhaps these constructions have different origins from standard nominalization.

Clause nominalization

Two types of clauses are nominalized, and these are causal and temporal clauses. The causal construction consists of a nominalized clause (which expresses the cause), followed by a grammaticalized case marker (the instrumental), and finally the main clause. As for the temporal construction, it is a nominalized clause followed by a temporal expression (borrowed from Nepali), which can be either *patshi* ‘after’ or *səmma* ‘until’.

Some examples of both types are shown below.

- (15) *go pomuʈhok mi-peu-wa-m-ka krym si-ŋro*
 1s food NEG-eat.1s/3s-IRR-REL⁷-CAU hunger feel-1s.PST
 ‘Because I didn’t eat, I felt hungry.’
- (16) *meram tsuʈtsu krym si-ra-m-ka khrap-saŋa bu*
 that child hunger feel-3s.PST-REL-CAU cry-PROG be.3s
 ‘Because that child is hungry, he is crying.’
- (17) *bju-ka l-ry-m patshi muu-tsu-tsip-ka*
 eagle-ERG carry-3s/3sPST-REL after thather-child-DU-ERG
mal-to mal-to lək-tsi-ʔe
 search-CONV search-CONV go-3d.PST-HS
 ‘After the eagle carried her off, her two children went searching.’
- (18) *meno roypha-ŋu ma dhol kuk-pu-m səmma...*
 there arrive-1s SEQ drum beat-1s/3s-REL until...
 ‘Until I arrive there and beat the drum...’

The same distribution of markers is seen as for relativization, whereby a monosyllabic verb form must be nominalized with *-mim* (and *-m* is blocked).

⁷ As indicated in the introduction, I have assigned labels to these markers based on the function they mark most frequently, and these labels appear in the glosses even when I am giving evidence of participation in a different construction. Because of the complexity of the phenomenon, and the overlap, the labels therefore become slightly misleading in certain cases.

- (19) *meram si-mim patshi*
 that die.3S-REL2 after
 ‘After she dies...’

Perhaps the presence of what we saw to be a relativizer in these constructions which are based on nominalized clauses involves a stage where they were indeed relativized clauses, with some relevant but perhaps semantically weak head noun, which was followed by either the instrumental case marker (in the expression of cause) or the temporal expression (for temporal clauses)⁸. These head nouns could then have been dropped, with the result that the construction now looks like a nominalized one.

Verbal periphrasis

This is the combination of the nominalized past-tense finite verb and of the copula. The resulting construction conveys perfect aspect. This construction also appears in a number of other languages of Nepal, such as Hayu, Chantyal, Limbu, Yamphu, among others.

- (20) *dzetha-mim tsahi waythuu babante l̥-mri-m bu.*
 o.brother-PLU CONTR other where go-3p.PST-REL be.3s
 ‘The older ones went somewhere else.’
- (21) *hu grenem-ra l̥s-ta-m bu*
 there nettle-LOC go-3s.PST-REL be.3s
 ‘She went out to the nettles.’

In this construction, nominalization is accomplished by means of *-m* (and because the construction calls for a past form of the verb as the input, the verb is automatically disyllabic and *-mim* does not occur), which we saw was the relativizer elsewhere.

Sentence nominalization

Matisoff refers to the marking of entire sentences as nominalized as their “reification”, and says that they can often be translated as beginning with “It’s a fact that ...”. While it is reported for a number of languages of the area (Belhare and Chantyal among others) that the result is pragmatically marked, I believe this construction to be pragmatically unmarked in Thulung. The pattern seen in

⁸ In fact there is another temporal construction in Thulung with the loan word *bela* ‘time’ in place of *patshi* or *somma*. Because *bela* is a noun, this construction looks more like relativization, with the temporal clause being relativized to the head noun *bela*. Perhaps what has really happened is that the *patshi* and *somma* clauses are built on the same model as the *bela* type, and the same markers, in other words those which are really relativizers in form, were applied to the clause, even though there was a functional difference due to the lack of head noun.

conversations is that the answers take the same form as the questions (that is, the same presence or absence of a nominalizer.)

- (22) A: *bante lamdi-nni-m?*
 where walk-POL.2p.PST-REL
 B: *go ηopsu grum-ḍa lamdi-ηro-m.*
 1s friend visit-PURP walk-1s.PST-REL

A: 'Where did you go?'
 B: 'I went to see a friend.'

- (23) A: *gumi bante bu-mi?*
 POL.3p where live-POL.3p
 B: *gumi basbari-ra bu-mi.*
 POL.3p Basbari-LOC live-POL.3p
 A: *gumi ba-lajka ro-mri-m?*
 POL.3p where-ABL come-POL.3p.PST-REL
 B: *gumi mukli-lajka ro-mri-m.*
 POL.3p mukli-ABL come-POL.3p.PST-REL

A: 'Where does she live?'
 B: 'She lives in Basbari.'
 A: 'Where is she from?'
 B: 'She is from Mukli.'

Nominalized sentences also occur in narrative, but with less frequency than in conversation, and even then, the frequency depends on the speaker and the context. Even the item marked seems to vary somewhat: while the nominalizer is generally marked on the final verb, I have also seen the marker follow the final hearsay marker (as seen in (25) below). From looking at these sentences within narratives, I do not think that pragmatic marking is the trigger. There does not seem to be anything that distinguishes the sentences which are marked, compared to those which are not.

- (24) *mu-ku u-tshoktso-ka mem bet-tsi-m-ʔe*
 that-GEN his-anger-INSTR like.that do-3d/3s.PST-REL-HS
 'They acted that way out of anger with him.'
- (25) *memma meram badzi-lajka iki-beppap-mim glwa-mri*
 after.that that bet-ABL our-ancestor-PLU win-3p.PST
ma tsahi gui thuluj dys-ti-ʔe-m.
 SEQ CONTR 1pi Thulung become-1pi.PST-HS-REL
 'After our ancestors won that bet we became Thulung.'

One interesting fact is that the nominalized sentences are very often past in form. This leads me to believe that nominalized sentences are sentences having perfect aspect (expressed by verbal periphrasis, as seen above) from which the final copula has been omitted. This is an idea suggested by Noonan as the origin of nominalized sentences in Chantyal (although he makes it clear that nominalized sentences in Chantyal have evolved separately into pragmatically marked situations). This would explain the past tense constraint on nominalized sentences: Thulung has not yet grammaticalized the difference between verbal periphrasis with the copula omitted and nominalized sentences.

One sentence I elicited reinforces this, as the consultant told me that the final copula could perfectly well be left out with no change in meaning, implying that speakers themselves consider nominalized sentences to be ‘the same’ as sentences with verbal periphrasis missing the copula.

- (26) *ama-neb-ra* *gani* *rok-ni-m* (*bu-ni*)
 my-house-LOC POL.2s come-POL.2s.PST-REL (come-POL.2s)
 ‘You came to my house.’

In sum, sentence nominalization occurs, as it does a great many Tibeto-Burman languages. However, the role of this construction in Thulung does not seem to correspond to the parallel in other languages, as it is pragmatically unmarked. There is some evidence that it is a result of truncation of the copula for a perfect aspect-marking periphrastic construction, so perhaps what we see is an intermediate stage, and that grammaticalization will result in pragmatic differences in nominalized sentences versus non-nominalized.

Genitivization

A possessive relationship between two nouns is usually expressed with genitive markers *-ku* or *-kam*, thus the pattern of identically (or similarly for Thulung) marked relativization, nominalization and genitivization would seem to break down here. But there are two situations in which the genitive is in fact expressed with a marker from the group we saw above: one is when the possessed noun is a time word, in which case it is marked with *-m* or *-mim*. The other case is with possessive pronouns, for which there are several variant forms, some of which also show the presence of one of these markers. These two slightly marginal cases of genitivization marked with the same class of markers seen for nominalization and relativization are perhaps remnants of an earlier stage of the language, before the genitive markers *-ku* and *-kam* came to be used⁹.

⁹ In addition to which, the final *-m* on the genitive case marker *-kam* is a subject for further study.

Possession of time words

For both native and borrowed time words, the genitive is expressed with *-m/-mim* rather than the standard genitive markers. The distribution of these two markers is as follows: *-m* appears post-vocally, while *-mim* appears post-consonantly. While this pairing of *-m* and *-mim* is familiar from relativization, the distribution is different, and for this I have no explanation for the time being.

- (27) *nemtha-m/*mim* *dzam*
evening-REL/REL2 rice
'the evening meal'
- (28) *dika-m/*-mim* *lagi*
tomorrow-REL/REL2 sake
'tomorrow's sake, ie. for tomorrow'
- (29) *aneb-mim¹⁰/*-m* *din*
today-REL2/REL day
'today's day, ie these days'

When the construction has no overtly marked head, the result is an NP referring to an individual by his day of birth (it is surprisingly common for people to talk about themselves or their children this way.)

- (30) *buddhabar-mim/*-m*
Wednesday-REL2/REL
'[The child] born on Wednesday'
- (31) *basta-m/*-mim*
yesterday-REL/REL2
'the one born yesterday'
- (32) *bam-din-mim*
which-day-REL2
'the one from which day (N. *kun din-ko*)'

It seems significant that the markers used for this sub-set of genitivization are the same ones used for relativization (with a different distribution though: that for relativization is based on syllable-length, whereas that for genitivization is dependent on the final phoneme of the marked word). Noonan suggests that the genitive function derives from the relative: "Once the attributive function becomes established in relative clauses, it may be extended to other sorts of attributives."(1995)

¹⁰ Perhaps there is a possibility that the reading 'these days' is a result of *-mim* as pluralizer, rather than genitive marker, but that still wouldn't explain how it can be followed by *din*.

Possessive pronouns

Thulung has four sets of possessive pronouns, clearly related, and interchangeable when used prenominally. These show a borderline case between nominalization and genitivation.

Possessive pronouns

1st person		2nd person		3rd person	
<i>a</i>	<i>aki</i>	<i>i</i>	<i>ini</i>	<i>u</i>	<i>uni</i>
<i>ama</i>	<i>akima</i>	<i>ima</i>	<i>inima</i>	<i>uma</i>	<i>unima</i>

What the table above shows is the possibilities for each of the three persons: the first row of possessive pronouns are just that, and can only be used prenominally. The second row can be used prenominally, but can also be used independently, standing in for the omitted noun.

Thus we have:

- (33) *a-khlea*
aki-khlea
ama-khlea
akima-khlea
 my-dog
 'My dog'

However, only those forms ending in *-ma* can stand in for the noun.

- (34) *akima/aki-khlea* *ɖokpu bu*
 my-dog big be.3s
 'My dog is big.'
- (35) *akima/*aki* *ɖokpu bu*.
 Mine is big.

This *-ma* is not synchronically a morpheme, yet it certainly seems to represent some earlier nominalizing suffix which turned possessive pronouns into substantives. I chose to discuss this non-productive nominalizer under genitivation because synchronically the pronouns where it appears are both genitive and nominative forms, judging from their ability to perform both functions.

DeLancey (1989) gives an interpretation of the origin of a Newari genitivizer, which proves useful for the case of Thulung.

‘Ram’s thing, a dog’

-where an empty noun is apposed to another noun, becomes reinterpreted as

‘Ram’s dog’

-and the empty noun becomes a genitive marker instead.

Perhaps this is also the case here:

akima khlea

‘mine, a dog’

gets reinterpreted as a genitive, ‘my dog.’

So if the original scenario is *akima* being only nominal (we have no such data, but it seems fairly likely that such was the case at some point), then *-ma* is reinterpreted as being a genitivizer.

I have described various constructions covering the three functions of relativization, nominalization and genitivization, all of which are expressed using markers from the same set: *-m*, *-mim*, *-ma* and *-mu*. Noonan suggests for Chantyal the following line of development: nominalization (through eventual erosion of genitive marker which is initially used to link the attributive to the head noun) turns into relativization which turns into genitive (by analogy to other attributives of use of nominalizer in relativization). The fact that we have four different markers in Thulung complicates the picture, making it difficult for the time being to trace a direction of development. Nevertheless I believe these markers to be related etymologically, and that it is not a coincidence that these three classically related functions in Tibeto-Burman languages also align in Thulung.

References

- Allen, N.A. 1975. *A sketch of Thulung grammar*. Ithaca: Cornell University.
- Bickel, Balthasar. 1999. Nominalization and focus constructions in some Kiranti languages. In Y. Yadava & W. Glover (eds.) *Topics in Nepalese Linguistics*. Kathmandu: Royal Nepal Academy
- DeLancey, Scott. 1989. Relativization and nominalization in Tibeto-Burman. Ms., University of Oregon.
- Genetti, Carol. 1992. Semantic and grammatical categories of relative clause morphology in the languages of Nepal. *Studies in Language* 16: 405-427.
- Matisoff, J.A. 1972. Lahu nominalization, relativization, and genitivization. In J. Kimball, ed. *Syntax and Semantics*, Volume I. 237-258
- Noonan, Michael. 1997. Versatile Nominalizations. In J. Bybee, J. Haiman & S.A. Thompson (eds). *Essays on language function and language type*. Amsterdam/Philadelphia: John Benjamins.

Aimée Lahaussais

Aimée Lahaussais
6 square Grangé Paris 75013 France

aimee@vjf.cnrs.fr