Nasal coda loss in Northern Naga: Revising W. T. French's *–Vŋ rimes

Kellen Parker van Dam
kellenparker@gmail.com
Chair for Multilingual Computational Linguistics, Universität Passau, Germany
Department of Languages & Linguistics, La Trobe University, Australia

Abstract
The earliest reconstruction of Patkaian (sometimes known in the literature "Northern Naga") was produced by W. T. French in 1983. While this was a monumental work and of major significance to the field of Tibeto-Burman linguistics, much of the analyses were based on limited data available to the author at the time. As a result, a number of the reconstructed rimes do not hold up today in light of a considerably greater amount of dialectal data to pull from. This paper investigates two recently reconstructed rimes in the language group, previously labeled *#–aŋ*^{I} and *#–aŋ*^{II} in Parker, K. (2019). Additional rimes for which French (1983) reconstructed nasal codas are also investigated. In addressing these rimes, an explanation is given for some notable exceptions. In particular the concept HORSE, which defies any clear placement in a single rime across the modern languages. Borrowing from a close relative is proposed as the source of this discrepancy, and phonetic motivations for the modern-day reflexes are investigated.

Keywords Patkaian, historical linguistics, reconstruction, phonology.

1. Introduction

Two distinct *#–aŋ* rimes are reconstructable in Tangsa-Nocte (Glottolog: kony1249), the most internally diverse branch of Patkaian, formerly known as Northern Naga, (Parker, K. in press). These times previously been referred to as *#–aŋ*^{I} and *#–aŋ*^{II} in Parker, K. (2019). In the majority of language varieties within this subgroup, reflexes of the two rimes developed distinctly from each other, with one retaining the velar coda and resulting in a shape similar to /-aŋ/ and the other undergoing coda loss, typically resulting in an open CV syllable with either an /a/ or /o/.

Complicating reconstruction of the phonological shape of the rime in the proto language is an inconsistency regarding which of the two rimes occur today from one language to another. Specifically, in some variety's reflexes of *#–aŋ*^{I} retain the coda while the other does not, and in other varieties the opposite is seen, where *#–aŋ*^{II} is the one to regain the coda. In still other varieties, the rimes are indistinguishable due to having merged entirely, with both pronounced /-aŋ / today. While there are no attested cases in which a variety will have retained the nasal coda on *#–aŋ*^{II} but not *#–aŋ*^{I}, the exact motivations for the sound change resulting in the loss of nasal coda on *#–aŋ*^{II} requires explanation. What's more, despite the overall regularity of reflexes, a number of etyma create some complication in the reconstruction, most notably that for HORSE *#kVm.raŋ*, suggesting a close borrowing rather than common inheritance.

Previously, this had been assumed to be a case of two distinct rimes rather than an incomplete but otherwise regular sound change for a number of reasons. First, should it be the case that there was a single **–aŋ* rime, phonological motivations for the split would need to be accounted for, but no conditioning factors have been identified. For the most part, reflexes of the two rimes are consistently distinct across Northern Naga, with only a minority of varieties merging them. An explanation for the way in which the two sides of the split would be required which could explain the direction of sound change as well as why it did not occur consistently across varieties.

The issue of reconstructing *#–aŋ* rimes in Patkaian was first pointed out in Parker, K. (2019) but without a suggestion of the possible mechanisms resulting in the current state of the two rimes. Resolving this uncertainty has implications on further historical, typological and phylogenetic research in the region and particularly with members of the purported Sal language group within Tibeto-Burman, as it may help better establish the place of the language group within the Tibeto-Burman genealogy. There is significance as well when taking a more zoomed-in approach to the languages including basic documentation of what is an otherwise largely under-documented group of languages, as the etyma which fall under these two rimes are incredibly common within the basic vocabulary. Such basic concepts covered by these rimes include SKY, BONE, RED, YEAR and ALIVE, as well as the corresponding fourth-tone *#–ak^{I}* rime for concepts such as BLACK and BITE.[1] By resolving historical changes, low-level classification and cognate identification can be more accurately done in the future.

While a reconstruction has previously been published for Northern Naga, being that of French (1983), the provided rime reconstructions have a degree of inconsistency in terms of the modern reflexes which is all the more apparent today as a much greater amount of linguistic data has become available. To give an example of some of the mismatches, ELEPHANT, YEAR and RAPTOR are reconstructed for French's proto Northern Naga (hereafter PNN) as **Ç-glaːŋ *paːŋ* and **laːŋ* respectively, all reconstructed with long vowels and all falling under the *#–aŋ^{II}* rime. LIGHTWEIGHT *#kʲaŋ^{II}* is reconstructed by French as **gyaŋ* with short <a>, while many other frequent *#–aŋ^{II}* stems are not found in French. This is the case across the family, in large part due to the quality of data available to W. T. French at the start of the 1980s.

An additional issue relating to data availability is not just that there were fewer language varieties which had reliable data sets was the fact that much of the data available to French simply did not cover as large a portion of the lexicon of these languages as is available today. Thus, a number of stems which would be informative for a reconstruction in 1983 were simply unattested. For example, INTOXICATED *#paŋ^{II}* and SIT *#ŋaŋ^{I}* are both absent in French, as they were in Needham (1897).

As with *#–aŋ^{II}* rimes, corresponding *#–aŋ^{I}* stems are also reconstructed for PNN again with the resulting proto-language proving less reliable in demonstrating regular sound changes than may have actually been the case. Such examples include alive **Ç-raŋ*, CORPSE **maŋ* and DREAM **maŋ* compared to MACHETE **gləŋ* and SKY **rəŋ*. A pattern already begins to emerge in these cases, being a three-way split combining vowel quality and vowel length, but one which does not consistently map to modern data where more detailed phonemic analyses do not suggest phonemic vowel length.

Additionally, a handful of stems do not easily reconstruct for one rime or the other. Most significant of these is HORSE *#kVm.raŋ*. In some varieties the second syllable patterns as though a reflex of *#–aŋ^{I}*, and in others of *#–aŋ^{II}*. However, despite some significant examples, the majority of stems do consistently pattern with one rime or the other across modern varieties.

With all of this in mind, this paper presents a resolution for these issues, providing an explanation for the paths of development of the two rimes for the language group as well as an explanation for the irregularity of HORSE in modern Patkaian varieties. This is done in part through comparison to sound change patterns elsewhere in Northern Naga, as well as through comparison to a very similar sound change that has been previously described for the Sūzhōu dialect of Wú. By considering the different reflexes in Northern Naga in relation to attested sound changes in Northern Wú where the historical phonology is much better established, a more accurate transcription of the two rimes can be determined. This results in the reconstructions of **–ɐŋ* for what was previously described as *#–aŋ^{I}*, and **–aŋ* for *#–aŋ^{II}*, bringing the pair more in line with two of French's many reconstructions – **–əŋ* and **–aŋ* respectively – while correcting inconsistencies due to the lower quality of data available at that time. Data are presented in support of the conclusion from a range of Northern Naga

[1] CV/CVN syllables will occur with tones 1-3, and typically have a corresponding 4th tone form with a homo-organic stop coda not unlike the A-D tone system of Tai or the four tones of Middle Chinese (van Dam 2018).

varieties including those belonging to Tangsa-Nocte, Wancho, Konyak, Phom, Chang, Lainong, Makyam, Chang, Pounyiu and Khiamniungan.

While French (1983) provides a much wider range of possible reconstructed rimes for the series of stems which fall under *#–aŋ*^{I} and *#–aŋ*^{II}, this is also the case for other rimes ending with velar codas, as well as a number of non-velar codas are also reconstructed for proto-Northern Naga (PNN) which are less easily justified in light of greater data availability, especially when compared to cognates in the wider Tibeto-Burman area.

This is especially the case for stems for which a final liquid coda is reconstructed, such as COUNT which French reconstructs as **ʔ-we*, breaking from Matisoff's PTB **wyal* and which I reconstruct for Northern Naga as having final **–al* which has undergone a shift to /j/ in many varieties resulting in the <e> vowel posited by French, but which is also widely retained in many varieties not known to French in the early 1980s. In other cases, what I have reconstructed as final **–l* is reconstructed by French as **–r* or **–n* for many stems which have PTB reconstructions with **–l*. Again this is understandable given the data available at the time. One reason for this is that French was operating with a much smaller data set, covering a few at-the-time well-documented language varieties, but also he was limited to varieties which did not so easily retain **–l*. Specifically, French relied on data from Joglei (Yogli), and Ha'wa Nocte (Namsangia) for Tangsa-Nocte varieties, along with Muishaung (Moshang) which we will come back to in a moment, as well as a single variety of Wancho, Konyak, Phom and Chang each. However, a number of varieties – notably Ngaimong, Muklom, *and Muishaung* – prove quite conservative in retaining final **–l* codas. But with the Muishaung data being largely tied to that of Needham (1897) which featured data from a speaker for whom coda /l/ was less stable (Parker, K. 2025) and instead often given as /r/, French would have missed much of the evidence for a retained /l/ coda in Tangsa-Nocte. Additionally, the doculect of Muishaung was of a time when the now nearly-complete process of diphthongisation was only part way complete, resulting in greater variation of **–Vŋ* rime pronunciation than was the case either 100 years earlier or 100 years later.

As for the other varieties mentioned above, French had no access to these data, and so the liquid coda not showing up as prominently in his reconstructions is understandable. This is the case as well for sun, which French reconstructs as **cər* but which I reconstruct with the **–al* rime, corresponding to the same *#sal* stem after which Burling (1983) named the language family. It is less feasible for the higher-order **sal* reconstruction for sun to be correct — which I believe to be the case — and then Northern Naga to have undergone a shift to final **–r*, only for coda /l/ to still be widely attested in the group. The potential outsize influence of Needham's 1897 transcription of a language at the midpoint of a major sound change on French's reconstruction will be discussed further below.

Across the Sal languages, Northern Naga is the most internally diverse today, both in terms of number of distinct varieties and how much they differ from one to another. Tangsa-Nocte has somewhere around 80 different varieties (Morey, 2017) with a number of internal subgroups differing in terms of phonology, morphology, lexicon, and syntax. Muishaung and Joglei, the varieties included in French's reconstruction, are very closely related and both part of the Rangpang/Pangwa group. They share an incredibly high degree of lexical similarity, with only a handful of significant phonological differences. Not included in French's data were highly divergent varieties such as those in the Champhang group, Kon-Nyinshau, Chuyo-Gaqkat, Aasen-Yaqsawa, other Noctean varieties such as Hakhun, Muklom or Phontai (Phong), all three of which now have well documented descriptive grammars published (as Boro 2017, Mulder 2020, and Dutta 2023 respectively). In addition to the wide range of Tangsa-Nocte varieties which were inaccessible to French, other Northern Naga varieties such as Tanhai Konnyak (Jacques and M. T. Konyak 2010), Chen (H. Konyak and Mulder 2022), Karyaw, Khiamniungan, Lainong, Makyam, Lao, Nahen, Makyamic, and Gongvangpounyiu were also unaddressed.

French's proposed correspondences were well motivated and often well described given the data available at the time. However, with access to a much larger set of data for a more representative sample of Patkaian languages, a clearer picture can be painted regarding what correspondences exist, and which cases are likely

exceptions to these patterns due to sound changes which have not diffused across the lexicon. Furthermore, with the data currently available for those varieties which French did analyse, gaps in the lexicon have been able to be filled in, allowing greater confidence in reconstructing rimes with more consistency.

An additional difficulty faced by French was a likelihood of inconsistency in transcription of the data. Sometimes this inconsistency was due to ongoing sound changes, in many cases regarding velar nasal codas, discussed in greater detail below.

In settling these uncertainties, this paper contributes to ongoing efforts at reconstruction and typology in North East India and Myanmar, while also providing a first look at a new reconstruction of Northern Naga (Patkaian), revising that done published in French (1983).

2. Conventions used in this paper

In order to distinguish the revised reconstruction from proto-Northern Naga as reconstructed by French (1983), the branch will hereafter be referred to as Patkaian, and the new reconstruction as proto-Patkaian (PPk). Further discussion on the name can be found in Parker, K.(in press), with the main motivation being that there is no link between Northern Naga and the other so-labelled 'Naga' languages spoken further south which would otherwise separate Northern Naga from the rest of Sal. Post and Burling (2017:225) offer the alternative label "Northern Sal" for the same reasons of recognising a break between it and the Naga languages.

Unless marked otherwise, reconstructions that follow are my own. Those of French will be marked with PNN and – where deemed beneficial to the reader – the revised reconstructions with PPk . Tone information is included when useful for illustrating different otherwise homophonous rimes, but has been left out of tables in order to have more room for additional segmental word forms. When present, subscript numerals indicate tone categories, while superscript numerals indicate pitch contours following the five-level system established in Chao (1930).

Reconstructed etyma given with an asterisk are ones for which the correspondences are clear. The reconstruction should be taken at least in part as symbolic, and not necessarily a proposed recreation of the pronunciation of the proto-language. Forms which are not fully reconstructed but point to likely cognate classes are marked with a hash (#), as is the case with *#–aŋ*^{I} and *#–aŋ*^{II}, which have been transcribed as such elsewhere, but which should be more realistically transcribed **–ɐŋ* and **–aŋ* respectively. Concepts are given in small caps and glosses in single quotes.

3. Data sources

In addition to recent contributions to the literature, considerable as-yet unpublished data has been collected across Northern Naga varieties. Much of this has been based upon the *Culturally Appropriate Lexicostatistical Model for Southeast Asia* word list, containing 235 concepts (Matisoff 1978) as well as another list of 1,467 concepts of undetermined provenance. Much of this data was collected by the author or others operating in the region around the same time, and remains largely unpublished. A few historical sources are also to be found. One of the earliest – and perhaps most significant in terms of illustrating sound changes – is the Muishaung (Moshang) word list collected by Needham (1897) already mentioned above, much of which was reprinted in Grierson 1930). Much of French's reconstruction thus indirectly relies on data from Needham, as interpreted by Grierson and later Shafer (1953). These are also the primary sources of Tangsa-Nocte data for Marrison (1967), and it is not until later when researchers such as Dutta (1969) and Walker (1948) elicited newer data. In many cases data were provided by speakers directly.

Due to the lack of additional documentation happening for much of the early 20th century, Needham remains one of the most valuable sources on Tangsa-Nocte as spoken at the time both for the extensive account of Muishaung as well as additional data for Shecyü in the text. By comparing between sources, we are able to see a number of rimes have remained quite stable throughout the past century and a half. Table 1 shows the **–an*

rime across three time periods. Many of the early word lists and descriptions were sparse, resulting in many holes in the table. However it is clear even with these gaps that there has been little if any change in *–*an* across recent generations.

Table 1: *–*an* in Muishaung Tangsa

gloss	Needham	Marrison	Das Gupta	modern
IRON	yân	yan	–	ʒan$_2$
PLATE		–	–	pan$_2$
COW	mân	–	man	man$_2$
BRACELET	sân	san	–	san$_2$
BREAD	–	–	–	βan$_2$

4. Possible errors in French's sources: Examples from Needham (1897)

Despite the value of Needham's contribution, there are still many inconsistencies. Some of those are themselves valuable as they provide insight into changes that must have been going on at that time. Others, however, are likely to be errors in transcription. Muishaung consultants have on multiple occasions told me that this is the case, believing an Assamese interpreter to be the one to blame. From Table 1 Needham's <ân> may seem to reliably reflect *–*an*, but many exceptions are also to be found. Table 2 shows a number of such cases, with <ân> clearly corresponding to *#–aŋ*^{II}*. This is significant as *#–aŋ*^{II} lost the nasal coda and went through a stage of nasalisation on the vowel before this too was lost, and Needham used underlined <n̲> as the typographic convention for nasal vowels.

While modern Muishaung shows much greater uniformity than what is described in Needham's account, this apparent inconsistency is also of value, however. According to Needham's own descriptions of the orthography employed in his text, <â> likely corresponds to /ɑ/ and <o> to /ɒ/. This /ɑ/ value for <â> provides the first clue for the likely phonetic values of the vowels in our competing *#–aŋ* forms, discussed in more detail in Section 6 below. Note the similarity of transcriptions from Marrison (1967) with those of Needham, the underline being lost entirely by the time of Marrison.

Table 2: Shift from **aŋ*^{II} to /ɔ/ in Muishaung

gloss	Needham 1897	Marrison 1967	Das Gupta 1980	modern
ELEPHANT	jân	jan	co:	tɕɔ$_1$
RED	shân	shan	–	ʃɔ$_2$
LIGHTWEIGHT	chân	chan	–	tɕɔ$_1$
LOOK	lân	lan	–	lɔ$_3$
BONE	rân̲	–	–	rɔ$_1$
BAMBOO	wân̲	wan	–	βɔ$_2$
OLD.MAN	–	likang	–	li$_2$kɔ$_2$
OLD.MAN	vî-kân̲ wâ	vikanwa	–	–
BEAR	cha'bâ	chaba	–	tɕəp$_4$bɔ$_2$

It's also entirely possible that Needham would have interpreted nasalisation as segmental in some cases. The data provided in the original 1897 publication shows considerable variety in a small number of pages.

Another significant point is that of nasal/non-nasal alternation. Tangsa-Nocte varieties have a four-way tonal distinction, likely derived from phonemic phonation distinctions. Three of the four tones are found on open syllables, including those with nasal or liquid codas, and the fourth tone is found only on closed syllables (van Dam 2018). This pattern is similar to that of the traditional four-tone system of Middle Chinese, with *ping, shang* and *qu* versus checked *ru*, or Tai *A B* and *C* versus checked *D* tone.

These 'rime families' are common throughout Tangsa-Nocte. Most often these will involve an alternation between a nasal coda on one hand and a non-nasal stop coda with the same place of articulation. One example of such an alternation would be between MACHETE **ʒˈaŋ^{I}* and HAND **ʒˈak^{I}*. With incredibly few exceptions, rime families undergo the same sound changes within a given variety. Thus, while those two concepts differ on the segmental level, we may also say they differ simply by tone. While perhaps a step too far for most readers, a single **ʒˈaŊ* form could be proposed. We see this same pattern across the lexicon when looking at verbal stem suppletion correspondences across dialects (Morey 2017:364f), but also in splits which have occurred between two tonal variants for a given lexeme. As an example, the verb HEAR/LISTEN corresponding to French's PNN **Ç-ta:t* has two possible etyma in Northern Naga varieties, namely **tat* and **tal$_{I}$*. In most Rangpang Tangsa varieties it is the latter, while among Heimi varieties it tends to be the former. Alternations also occur for other phonemes, notably liquids /l/ and /r/ where a fourth tone equivalent coda becomes /t/. Fully open CV syllables shift to a glottal stop, and final bilabials likewise retain place of articulation when undergoing denasalisation.

Thus, one finds *#–aŋ^{I}* finals corresponding to *#ak^{I}* final words such as BLACK *#njak* and BITE **kak*, among others. No non-nasal equivalents of *#–aŋ^{II}* are to be found. All vocalic changes which occurred on *#–aŋ^{I}* also occurred in the same way on *#–ak^{I}* finals.

Despite potential concerns with some of the earliest data, Needham (1897) provides considerable assistance in terms of dating sound changes. This will be discussed in Section 6 below. First, further explanation of the confusion regarding Patkaian *#–aŋ* rimes is required.

5. Competing *#–aŋ* rimes in Patkaian

Table 3 shows examples of *#–aŋ^{I}* and *#–aŋ^{II}* rimes from none Tangsa-Nocte varieties and Wancho. In order, these are Cholim, Muklom, Haidley, Joglei, Khalak, Chamchang (Kimsing), Lochhang, Lungkhi, Muishaung, Tutsa, Wancho and Phom. PNN indicates French's Proto-Northern Naga reconstruction, with the exception of sit **ŋaŋ*, which French did not reconstruct. Instead, the proto-Tibeto-Burman (PTB) form is given (Matisoff 2003). PPk refers to the revised reconstructions, but corresponding to the same proto-language as referred to in French as PNN.

Dream, sky and sit are in the *#–aŋ^{I}* rime while lightweight, raptor and bone have *#–aŋ^{II}*, as does elephant, although in Gaqlun and some other varieties elephant is showing a reflex of *#–aŋ^{I}*.

Table 3: Examples of **–aŋ^{I}* & **–aŋ^{II}* rimes

	PNN	PPk	Ch	Ml	Hl	Jg	Kl	Km	Lo	Lk	Ms	Tu	UW	Pm
DREAM	**maŋ*	*#maŋ^{I}*	maŋ	mʌŋ	mo	məŋ	maŋ	ma	ma	məŋ	mɐuŋ	maŋ	məŋ	maŋ
SKY	**C-raŋ*	*#graŋ^{I}*	raŋ	ɹʌŋ	ro	rəŋ	raŋ	ra	ra	rəŋ	rɐuŋ	raŋ	gəŋ	ɣaŋ
SIT	**ŋaŋ*	*#naŋ^{I}*	ŋaŋ	-	ŋəŋ	ŋaŋ	ŋa	ŋa	ŋəŋ	ŋɐuŋ3	-	-	-	
LIGHT	**gyaŋ*	*kˈaŋ^{II}*	tɕa	ʧaŋ	tʃa	tɕʰaŋ	tʃa	ʒɯ	cʰəu	tɕʰa	tɕɔ	ʧaŋ	-	ʃaŋ
HAWK	**la:ŋ*	*laŋ^{II}*	la	laŋ	la	laŋ	la	lɤ	ləu	la	lɔ	laŋ	lã	laŋ
BONE	**ra:ŋ*	*graŋ^{II}*	ra	ɹaŋ	ra	raŋ	ra	rɤ	rəu	ra	rɔ	raŋ	gã	ɣaŋ
ELEPH.	—	*glaŋ^{II}*	tɕʰa	-	tɕʰaŋ	tʃʰa	tɕɤ	cʰəu	tɕʰa	tɕɔ	-	-	-	

6. Resolving the uncertainty

As mentioned in Section 4, Needham's description of his orthographic conventions gives us an idea of where to begin in explaining the split and apparent flip-flop, wherein some varieties have *#–aŋ*^{I} as closer to /aŋ/, and others instead *#–aŋ*^{II} as more similar to /aŋ/ today. As Needham states, <â> is long as in 'father', and <a> short as 'comp*a*ny' or the u in 'but'. Thus, we can assume <â> to be close to /ɑ/, and <a> as something more mid-central.

Additionally, as in Table 3, a number of languages show a clear difference in vowel quality, where the *#–aŋ*^{I} form is higher as is the case with Muklom /ʌ/ vs /a/, Joglei and Wancho /ə/ vs /a/, and Muishaung /ɐu/ vs /ɔ/. If a difference of vowel quality is what separates *#–aŋ*^{I} and *#–aŋ*^{II}, then an explanation can be given with evidence from a very distant relative in the Sino-Tibetan family: The variety of Wú Chinese spoken in Sūzhōu in eastern China.

As the historical prestige variety within Northern Wú, Sūzhōu dialect has retained a number of older features lost elsewhere in the Yangzte River Delta, and those changes which have occurred have been the subject of much attention over many years. One such change is the loss of velar coda **ŋ*, but only on certain lexemes.

In the case of Sūzhōu dialect, the loss of **ŋ* only occurred when preceded by a low vowel, while in cases where the preceding vowel was not low, it was instead fronted and nasalised upon coda deletion (Zhu 2019). These nasalised vowels still exist throughout Wú dialects today. This intermediate step of a CṼ syllable likely occurred for Northern Naga as well among those varieties which today lack a segmental coda among these rimes. In many Patkaian, however, nasalisation of vowels is not generally phonemic, and so the nasalisation on the vowel was likely lost quite quickly in such cases. There is evidence of phonemic vowel nasalisation a century ago in Needham (1897), and Wancho retains it through to today, but it is otherwise uncommon in the group. Retention in Wancho lends some support to this explanation, however, as *#–aŋ*^{II} rime is today realised as /ɑ̃/.

With the case of Wú dialects such as Sūzhōu as a reference, *#–aŋ*^{II} is thus reconstructed as **–aŋ* and *#–aŋ*^{I} as **–ɐŋ*. The apparent flip-flop found in varieties such as Lungri where /–aŋ/ corresponds to **–ɐŋ* can therefore be explained through coda deletion on **–aŋ* followed by loss of nasalisation on the vowel, followed by a lowering of the vowel in **–ɐŋ*. Unlike Sūzhōu Wú, the vowels in these cases did not typically undergo fronting, although this did occur with Jiingu (Dunghi) and the Hahcheng-Ngaimong subgroup as shown in Table 8.

Additionally, **–aŋ* (*#–aŋ*^{II}) has lost the final velar nasal in the majority of Tangsa-Nocte varieties and much of Patkaian / Northern Naga as whole. Some of the few exceptions are Chodok, Hawi, Muklom, Tutsa, Phom, the Tikhak-Yongkuk group and the Ngaimong-Hahcheng group. Conversely, the rime **–ɐŋ* (*#–aŋ*^{I}) has retained the nasal coda in the majority of not just of Tangsa-Nocte but Patkaian more generally. Lochhang, Lama, Mungre, Shecyü and Bote are some of the rare exceptions.

Needham's account provides another useful indicator of the timing of one notable sound change in Muishaung, namely the diphthongisation or vowel fracture of a number of rimes. In the case of Table 4, this is the **–ɐŋ* (formerly *#–aŋ*^{I}) rime which has become /auŋ/ in Muishaung. Muishaung shows a number of similar phonological innovations, many of them likely to be recent. In this case in particular, we can see precisely how recent due to the apparently inconsistencies in how **–ɐŋ* words were transcribed in Needham.

Table 4: Shift from **aŋ*^{I} to /ɐuŋ/ in Muishaung

gloss	Needham 1897	Marrison 1967	Das Gupta 1980	modern
MACHETE	yaung	yaung	ja[u]ng	ʒɐuŋ$_1$
GO UPRIVER	–	–	wang	βɐuŋ$_2$
FOREHEAD	khâng	khang	–	kʰɐuŋ$_2$

BEGIN	–	–	phaung	pʰauŋ₃
MANY	pâng	pang	–	pɐuŋ₃
PLACE	–	–	baung	bɐuŋ₁
WING	rong	rong	–	rɐuŋ₂
SIT	nong	nong	ngaung	ŋɐuŋ₃
SKY	–	rong	rang	rɐuŋ₂
FUT.	khâung	–	–	kʰɐuŋ₂
FUT.TR	–	–	shaung	ʃɐuŋ₂
FUT.INT	–	–	–	tʰɐuŋ₂
SAND	phân̲	phan	–	pʰɐuŋ₂
2PL	m'nang	–	mnang	m₁nɐuŋ₁

Das Gupta (1980), a much more recent source, shows at least some evidence of this change not having yet been complete at the time of writing. MACHETE is transcribed in two different ways in the text, once with the diphthong and once without, and both SKY and 2SG are given without diphthongisation. Needham (1897), echoed in Marrison (1967), also gives a different vowel for WING and SIT, however French (1983) reconstructs the former as **Ç-rəŋ* and the reconstructed PTB form of the latter also shows a lower vowel than /o/, so we may disregard Neeham's <o> here as either a mistake or something characteristic of his consultants but which did not survive into the present day, where all **–ɐŋ* finals in Muishaung are /ɐuŋ/ and all **–ək* (formerly *#–ak*^{I}) finals are /ɐuŋ/. In the present day the sound change is complete, and no phonemic /ɐŋ/ or /ɑŋ/ finals occur in Muishaung.

6.1 Timing of coda loss in Muishaung

Based on Needham's writing it can also be determined that in the late 1800s the loss of the velar coda on **–aŋ* (*#–aŋ*^{II}) was also not yet complete, as a number of these terms are given with a less-clear <an̲> with an underlined <n̲>. This is not explicitly explained in the text, although an oblique <*n*> is described as a "nasal sound like the French n in 'enfant"'. We may assume this is another typographical error and the underlined <n̲> is intended to mark nasalisation of the preceding vowel. We can be even more confident of this interpretation based on the sound changes described for Sūzhōu Wú discussed in Section 6 – wherein the loss of a coda resulted in a nasalised vowel which has since been lost – as well as the presence of a nasalised vowel in Wancho and other varieties which have phonemic vowel nasalisation, for example bone **graŋ* as /gã₂/ in the Upper Wancho dialect of Kamhua Noknu, or RAPTOR **laŋ* as /lã₁/ in the same variety. Burling (1998) in writing about a different variety of Wancho did not include vowel nasalisation, giving RAPTOR as /o⁴⁴la⁴⁴/. Whether this is indicative of a loss of nasalisation in that case or simply an oversight in transcription is unclear. This is the case as well for YEAR **grəŋ.uɸaŋ* which Burling gives as /zaŋ⁴⁴pwa⁴⁴/ but which in Kamhua Noknu is /gəŋ₁põ₃/, the shift of the final vowel /a/ to nasalised /o/ being a result of influence from the labialised onset.

Table 5: **aŋ*^{II} rime

	PPk	Ch	Jg	Kl	Km	Lo	Lk	Ms	Mr	Ng	Pt	Rk	Sc
ELEPHANT	**tɕʰaŋ*	tɕʰa	tɕʰaŋ	tʃʰa	tɕɤ	cʰəu	tɕʰa	tɕɔ	tɕʰɤ	tɕeŋ		tɕʰa	tɕɤ
YEAR	**βaŋ*	βa	paŋ	va	βɤ	vo	va	βɔ	ve			wa	vɤ
BE.BORN	**tɕʰaŋ*	tɕʰa		tʃa		cʰəu	tʰa	tɕɔ	tɕʰɤ	tɕeŋ	tʃaŋ		tɕɤ
LIGHTWEIGHT	**tɕʰaŋ*	tɕa	tɕʰaŋ	tʃa	ʒɯ	cʰəu	tɕʰa	tɕɔ	tɕʰɤ	tɕeŋ	tʃaŋ	tɕʰa	tɕɤ
RED	**ʃaŋ*	xa	ʃaŋ	sa	ʃɯ	səu	ʃa	ʃɔ	ʃɤ	seŋ	saŋ	aʃa	ʃɤ
RAPTOR	**laŋ*	la	laŋ	la	lɤ	ləu	la	lɔ	laŋ			la	lɤ
LOOK	**laŋ*	la	laŋ		lɤ	ləu	–	lɔ	le	leŋ		la	lɤ

BONE	ra1	raŋ1	ra	rɤ1	rəu1	ra1	rɔ1	əre1	reŋ1	raŋ1	ra1	rɤ1
HORNBILL		-raŋ2				-ra2	-rɔ2	-joŋ2		-raŋ	-ra2	-jɔŋ2

The inconsistencies between PNN and PPk are also not limited just to the *#–aŋ* rimes. Another rime having a velar nasal coda which is widespread throughout the basic lexicon has been tentatively reconstructed as *#–juŋ*. As with the *#–aŋ* rimes, French's reconstructions show a considerable range in shape. Examples of such terms include FLLOWER and BREAST (likely a case of polysemous stems), both PNN **puːŋ*, NOSE **C/V-guːŋ*, BACK **kʰeːŋ*, WHITE **luaŋ* and FLAME **Ç̊-loːŋ*. These are another case where having considerably more data provides insight into a fairly consistent rime, despite a number of nuclei reconstructed for PNN. Not included in French (1983) but widely attested is the stem for MOUTH, *#njʊŋ*. As with examples above, # here indicates that the correspondences are clear, but the underlying phonological values are still being formalised. These reconstructions may be considered accurate as correspondences to the proto-language but not as pronunciations. This is distinct from another rime, reconstructed for PPk as *#–ʊŋ*, which includes etyma such as STONE *#lʊŋ* (PNN *Ç-luŋ*), INSECT *#ʒʊŋ* (PNN **gluŋ)* and HEART *#mʊŋ* (PNN **moŋ*), although in this case there is perhaps something to be said for influence of the /w/ glide as reconstructed for PTB). These too show impressive consistency across Tangsa-Nocte, Khiamniungan, Lainong, Makyam, Konyak and Wancho. Table 6 shows reflexes of these two rimes for a sample of Patkaian languages.

In order from left to right those are Wakching Konyak, Upper Wancho of (Losu & Morey 2003), Phom, Ngaimong-Shangwan, Mungre, Shangthi, Raqnu, Pinku, Muishaung, Joglei and Yaqsawa.

Table 6: **–ʊŋ* and **–jʊŋ* rimes in PPk

	WK	**Ln**	**UW**	**Pm**	**Ng**	**Mr**	**St**	**Rn**	**Pi**	**Ms**	**Jg**	**Ys**
–ʊŋ	oŋ	uaŋ	oŋ	oŋ	uŋ	oŋ	uŋ	uŋ	oŋ	uŋ	oŋ	uŋ
–jʊŋ	eŋ	an	õ	oŋ	oŋ	uŋ	y	o	o	ɨ	oŋ	ɤ

While semantic shifts have resulted in many of these etyma no longer in use among some varieties, and poor data availability for some Patkaian varieties means there will be some holes regardless, these two rimes prove quite regular across Northern Naga.

6.2 Irregular correspondences on HORSE

As for the **–aŋ* and **–ɐŋ* rimes, we find considerable regularity across Patkaian although a few notable exceptions exist. The most significant of these is for HORSE, PTB **k-m-raŋ* (Matisoff 2003), which in Tangsa-Nocte varieties does not consistently pattern with one of the two rimes. Instead, some varieties have reflexes of **raŋ* while in others it appears to derive from **rɐŋ*. It is here suggested that this word had previously been lost in Tangsa-Nocte to be replaced later by a loan from a close relative, most likely Singpho or another Jinghpawic variety with which Tangsa-Nocte speakers were in close contact. Evidence of such a replacement can be seen in Muishaung, for which the word today is /kum₂rɔ₂/, with a reflex of **raŋ* on the second syllable. Despite considerable phonological innovation in Muishaung compared to other Rangpang Tangsa varieties, this is often in the form of splits rather than mergers. Thus, much can be gleaned from Muishaung in terms of rime categories. Much of Tangsa-Nocte otherwise shows reflexes of **rɐŋ*, even those varieties which are closest to Muishaung. It is therefore possible that a handful of varieties — notably Muishaung, Ringkhu and Hakhun — saw a shift of horse to the **raŋ* stem during some intermediate stage while vowel nasalisation was still present. Marrison (1967) gives <gimrang>, while Das Gupta (1980) transcribes <kumro>. With an onset voicing merger having occurred in Tangsa-Nocte, the <g> and <k> onsets transcribed here should be taken as equivalent. Marrison's transcription supports interpretation as a reflex of **raŋ* in this case, since we would expect to see something more like <raung> for **rɐŋ* at this stage, as it is clear from Needham (1897) that this

shift was well underway nearly 80 years prior. This is further supported by Das Gupta's account (1980). Meanwhile, while alternation between [i~ɨ~u] is common in Tangsa-Nocte, there is no clear indication that this should have been /i/ at this stage, and consulting older Muishaung speakers who were alive at the time of Marrison do not support an older /kim/ pronunciation of the first syllable.

Table 7 shows the distribution of HORSE across varieties of Tangsa Nocte. Note that in Mungre and Shecyü, the second syllable in horse does not appear to be a reflex of either of the two stems, suggesting the recentness of borrowing (Parker, K. 2020).

Table 7: Second morpheme of HORSE across Tangsa-Nocte varieties

variety	*–ɐŋ	HORSE σ2	*–aŋ
Bote	-aŋ	raŋ	-a
Champang	-əŋ	rəŋ	-ə
Cholim	-aŋ	raŋ	-a
Gaqlun	-aŋ	raŋ	-a
Hahcheng	-aŋ	raŋ	-eŋ
Hakhun	-ɤ	ra	-a
Jiingi	-əŋ	rəŋ	-ɛ
Joglei	-əŋ	raŋ	-aŋ
Khâlak	-aŋ	raŋ	-a
Longri	-aŋ	raŋ	-a
Maitai	-aŋ	raŋ	-a
Muishaung	-auŋ	rɔ	-ɔ
Mungre	-ɔ	raŋ	-ɛ
Ngaimong	-aŋ	raŋ	-eŋ
Rinkhu	-aŋ	ra	-a
Shecyü	ɑ	raŋ	ɨ
Tikhak	-aŋ	raŋ	-aŋ

Another exceptional stem is ELEPHANT, which, like HORSE, is not consistent in terms of which rime it patterns with. It has been suggested that this is a loan from Southwestern Tai (van Dam 2018; Morey & van Dam 2019) along with a number of other stems found in Tangsa-Nocte relating to wet rice cultivation.

Table 6 shows the reflexes of the two rimes discussed here as they appear across Patkaian. Columns represent regular correspondences for Tutsa, Chuyo, Haqchum, Jiingi, Hahcheng, Shecyü, Muishaung, Pinku, Phom, Wancho and Lainong in that order.

Table 8: Reflexes of *–ɐŋ and *–ɑŋ in some Patkaian varieties

		Tu	Cy	Hc	Ji	Hh	Sc	Ms	Pi	Pm	UW	Ln
*–ɐŋ	#–aŋ^{I}	aŋ	ɑŋ	ɑŋ	əŋ	ɑŋ	ɑ	ɐuŋ	ɑŋ	æŋ	aŋ	ɑi
*–aŋ	#–aŋ^{II}	aŋ	ɑŋ	ɑ	e	iŋ~eŋ	ɨ	ɔ	ɑ	æŋ	ã	o

7. Conclusion

From the above analysis a few conclusions can be made. First, with considerable consistency of reflexes, two distinct rimes can be reconstructed for stems reconstructed as *–*aːŋ*, *–*aŋ* and *–*ɐŋ* in French (1983). These are reconstructed here as *–*ɐŋ* and *–*aŋ*. Very loosely, French's *–*aŋ* and *–*ɐŋ* tend to correspond to PPk *–*ɐŋ*, while French's PNN *–*aːŋ* roughly aligns with PPk *–*aŋ*. However, many exceptions are to be found, almost certainly due to a paucity of data available for the PNN reconstruction.

Second, similar to sound changes seen with the Sūzhōu dialect of Northern Wú, in many Patkaian varieties, the *–*aŋ* rime saw a loss of the velar nasal coda, resulting in an intermediate stage involving nasalisation of the vowel. This nasalisation was then lost in nearly all varieties outside of Wancho which has retained phonemic vowel nasalisation. Meanwhile, *–*ɐŋ* rimes generally retained the coda across Patkaian varieties with only a few exceptions. In many cases the vowel was lowered resulting in an /aŋ/ reflex. In the case of Muishaung, the centrality of the vowel instead underwent vowel fracture, a process that was still in its early stages at the time of Needham (1897). A parallel change occurred to the PPk *–*ɐk* rime, which consistently patterns with *–*ɐŋ* across much of Patkaian.

Additionally, the lack of a reconstructable *–*ak* rime may be explained as the result of an early merger with *–*ɐk* during the development of Patkaian tone systems, where in nearly all varieties but Khiamniungan CVC syllables formed their own 'checked' tone category. For many of the Patkaian tone systems, duration is a secondary feature, and such checked syllables are almost always significantly shorter in duration. Mulder (2020) provides one possible counter-example in Muklom, with a length distinction on checked tones described for /puk/ for stomach and /puːk/ for ELEPHANT. For this to be conclusive, more investigation is needed. Gogoi (2025 p.c.) has suggested a similar length distinction on checked syllables for Tutsa, although voicing may play a role in this and it too requires further investigation.

The work undertaken by W. T. French in his 1983 reconstruction of Northern Naga is still of considerable importance. It was no small task, especially given the overall lack of data for what is otherwise an incredibly diverse and vibrant branch within Tibeto-Burman. its significance cannot be understated, not just in terms of showing relatedness of stems across the languages included in the study, but also in solidifying the name of the family for decades to come, with its only shortcoming being that it was a product of a time where data was much more difficult to come by, as would be the case for the following two and a half decades after its publication. Only in the past decade has the situation improved substantially in terms of how much data access there is, in large part due to the increasing number of community member scholars contributing to the collection and publication of linguistic works, boding well for an ever-improving understanding of the linguistic and cultural history of the Patkai mountains.

Abbreviations

PPk	Proto-Patkaian
PNN	Proto–Northern Naga of French (1983)
σ	syllable within a lexeme

Language abbreviations

Ch	Cholim	Mr	Mungre
Cy	Chuyo	Ng	Ngaimong
Hh	Hahcheng	Pm	Phom
Hl	Haidle	Pt	Phong
Hc	Haqchum	Pi	Pinku

Ji	Jiingi	Rn	Raqnu
Jg	Joglei	Rk	Ringku
Kl	Khalak	St	Shangthi
Km	Kimsing	Sc	Shecyü
Ln	Lainong	TK	Tanhai Konyak
Lo	Lochhang	Tu	Tutsa
LW	Lower Wancho	UW	Upper Wancho
Lk	Lunkhi	WK	Wakching Konyak
Ms	Muishaung	Ys	Yaqsaq
Ml	Muklom		

References

Bano, Atiqua (2008). "A descriptive study of Phom language". PhD thesis. North Eastern Hill University, Shillong.

Boro, Krishna (2017). "A grammar of Hakhun Tangsa". PhD thesis. University of Oregon.

Burling, Robbins (1983). "The Sal Languages". In: *Linguistics of the Tibeto-Burman Area* 7.2.

Burling, Robbins and L. Amon Phom (1998). "Phom phonology and word list". In: *Linguistics of the Tibeto-Burman Area* 21.2, pp. 13–71.

Burling, Robbins and Mankai Wangsu (1998). "Wancho phonology and word list". In: *Linguistics of the Tibeto-Burman Area* 21.2, pp. 43–71.

Chao, Yuen Ren (1930). "ə Sistim əv Toun-letəz". In: *Le Maître Phonétique* 30, pp. 24–27.

Dam, Kellen Parker van (2018). "The Tone System of Tangsa-Nocte and Related Northern Naga Varieties". PhD thesis. Melbourne, Australia: La Trobe University.

Dam, Kellen Parker van and Syed Iftiqar Rahman (2022). "Tangsa-Nocte as a Continuum: A diagnostic feature list for classification of varieties". In: *Himalayan Linguistics: Languages and Peoples of the Eastern Himalayan Region* 21.2.

Das Gupta, Kamalesh (1980). *The Tangsa Language: A Synopsis*.

Dutta, Niharika (2023). "A descriptive grammar of Phong". PhD thesis. Gauhati University.

Dutta, Parul (1969). *The Tangsas of the Namchik and Tirap Valleys*. North East Frontier Agency.

French, Walter Thomas (1983). *Northern Naga: A Tibeto-Burman Mesolanguage*. City University of New York.

Gogoi, Sayanjyoti (2025). personal correspondence.

Grierson, George A (1903). *Linguistic Survey of India*. Calcutta: Office of the Superintendent of Government Printing.

Jacques, Guillaume and Methna Tanhai Konyak (2010). "A Preliminary Study of Tanhai Konyak Historical Phonology". In: *La linguistique* 46.2, pp. 41–57.

Konyak, Hoipo and Mijke Mulder (2022). "A Brief Outline of Chen Phonology". In: *Proceedings of the Payap University Research Symposium*. Chiang Mai: Payap University.

Losu, Banwang and Stephen Morey (2023). "The Wancho language of Kamhua Noknu village". In: *Linguistics of the Tibeto-Burman Area* 46.2, pp. 201–234.

Marrison, Geoffrey Edward (1967). "The classification of the Naga languages of north-east India." PhD thesis. School of Oriental and African Studies (University of London).

Matisoff, James A (1978). A culturally appropriate lexicostatistical model for Southeast Asia languages.

— (2003). *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. Berkeley, Los Angeles: University of California Press.

Morey, Stephen (2017a). "Tangsa". In: *Sino-Tibetan Languages*. Ed. by Graham Thurgood and Randy LaPolla. London and New York: Routledge.

— (2017b). “The sociolinguistic context of the Tangsa languages”. In: *Sociohistorical Linguistics in Southeast Asia*. Brill, pp. 169–187.

Morey, Stephen and Kellen Parker van Dam (2019). “Material culture and Agriculture in Tangsa (Naga) languages – evidence from Tai borrowings and traditional song texts”. 24th International Conference on Historical Linguistics. Australian National University, Canberra, Australia.

Mulder, Mijke (2020). “A descriptive grammar of Muklom Tangsa”. PhD thesis. La Trobe.

Needham, Jack F (1897). *A Collection of A Few Môshâng Naga Words*. Shillong: Assam Secretariat Printing Office. Parker, K. (forthcoming). “Patkaian (Northern Naga)”. In: *Oxford Handbook of the Tibeto-Burman Languages.* Oxford University Press

— (2025). “‘A Classified Account of J. F. Needham’s A Collection of A Few Môshâng Naga Words’”. In: *Journal of Asian and African Studies*.

— (2020) “Borrowing and re-borrowing in Southeast Asia: Complicating factors in historical reconstruction”. *Workshop on Tracing contact in closely related languages*. Zürich

— (2019). “Horsing around: The Complicated State of Tangsa-Nocte *aŋ Rhymes”. 52nd International Conference on Sino-Tibetan Languages and Linguistics. University of Sydney, Sydney, Australia.

Post, Mark W and Robbins Burling (2017). “The Tibeto-Burman languages of Northeast India”. In: *The Sino-Tibetan Languages*, pp. 213–242.

Shafer, Robert (1953). *Classification of the northernmost Naga languages*. Vol. 39, pp. 225–264.

Walker, G.E.D. (1948). *Outline grammar of Yogli Naga*. Shillong.

Wayesha, Ahsi James (2010). “A phonological description of Leinong Naga”. MA thesis. Chiang Mai: Payap University.

Zhu, Yuhong (2019). “Diachrony of Coda **ŋ* in Suzhou, Wu Chinese: Regular Change and Differing Readings”. In: Proceedings of the 30th North American Conference on Chinese Linguistics (NACCL-30). ed. by Yuhong Zhu et al. Vol. 2. Columbus, pp. 444–461.