

43<sup>rd</sup> International Conference on Sino-Tibetan Languages and Linguistics  
University of Lund  
October 15–18, 2010

***In search of an antidote for scholarly remorse:  
creating a meta-linguistic vocabulary for Lahu\****

James A. Matisoff  
University of California, Berkeley

**I. Abstract**

Twenty-three years of my life (1965–1988) were devoted largely to the compilation of my Lahu-English dictionary.<sup>1</sup> Whatever its merits, this dictionary has so far been of little use to the Lahu people themselves, to my increasing chagrin. In my youthful puristic zeal, I had adopted a rigorously phonemic transcription quite different from the three pre-existing romanizations of Lahu, as well as an alphabetical order modeled on the *devanāgarī*-type systems of Thai and Burmese, rather than on the familiar ABC order of Western languages. This has made the dictionary virtually impenetrable, even for the minority of the Lahu population that is literate in one of the missionary transcriptions or the *pinyin*-based system used in China.

And yet perhaps I should not be too hard on myself. The very concept of alphabetical order—*any* alphabetical order—has been impossible to convey, even to my very best consultants. Realizing this, I attempted back in 1977 to compose a long introduction in Lahu, written in the standard Baptist transcription, justifying my own orthography and alphabetical order on linguistic grounds. This enterprise quickly foundered, mostly due to the lack of a meta-linguistic vocabulary, and this introduction did not make it into my published dictionary.

Fast forward to 2009, when our annual Sino-Tibetan conference was held in Chiang Mai, Thailand. There I was delighted to meet a group of Lahu graduate students in linguistics at Payap University, one of whom expressed great interest in making the dictionary more accessible. We have been in touch ever since, trying out our neologisms on each other, hoping to come up with a technical linguistic vocabulary that will be both easy to understand and esthetically pleasing. This talk will explore the various issues that have arisen in connection with this effort.

---

\* This material is based upon work supported in part by the National Science Foundation under Grant No. 0712570. My thanks to Daniel Bruhn and Liberty Lidz for formatting the present version.

<sup>1</sup> See “History of the Lahu Dictionary project” in Matisoff 1988:1–8.

## II. Phonemes of the M̄np̄ulón subdialect of Black Lahu

Consonants					Vowels			Tones		
p	t	c	k	q	i	ĩ	u	mid [unmarked]	33	○
ph	th	ch	kh	qh	e	ə	o	high rising	35	ó
b	d	j	g		ɛ	a	ɔ	high falling	53	ô
m	n		ŋ					low falling	21	ò
f		š		h				very low	11(2)	ō
v		y	ğ					high checked	4-?	ô?
	l							low checked	2-?	ò?

Syllable canon: (C) V T

### Notes:

- y functions as the voiced homologue of š.
- ğ [ɣ] functions as the voiced homologue of h.
- Lahu lacks a phoneme /s/.

## III. Allophony in the labial and palatal series

- The four labial phonemes /p ph b m/ are affricated before /u/, which is itself unrounded to [ʉ] in this environment:

/pu/ → [pfʉ]      /phu/ → [pfhʉ]      /bu/ → [bvʉ]      /mu/ → [mvʉ]

- The five palatal phonemes /c ch j š y/ become dentals before /i/, which is itself raised to [ɿ] in this environment:

/ci/ → [tɿ]      /chi/ → [tshɿ]      /ji/ → [dzɿ]      /ši/ → [sɿ]      /yi/ → [zɿ]

- The /n/ phoneme has a palatal allophone before /i/:

/ni/ → [ɲi]

#### IV. Alternative transcriptions<sup>2</sup>

At least four writing systems have been devised for Lahu, two of them by Christian missionaries, one by Chinese linguists, and one by the present author. See Tables I–III.

*Table I. Consonant Equivalences in Lahu Orthographies*

<i>Catholic</i>	<i>Baptist</i>	<i>Chinese</i>	<i>Matisoff</i>
q	k'	q	q
qh	hk'	qh	qh
k	k	k	k
kh	hk	kh	kh
g	g	g	g
ng	ng	ng	ŋ
c / tc	c / tc	c / z	c
ch / ts	ch / ts	ch / zh	ch
j / dz	j / dz	j / dz	j
t	t	t	t
th	ht	th	th
d	d	d	d
n / gn	n / ny	n	n
p / pf	p / pf	p	p
ph / phf	hp / hpf	ph	ph
b / bv	b / bv	b	b
m / mv	m / mv	m	m
h	h	h	h
gh	g'	x	ğ

---

<sup>2</sup> For a fuller discussion, see DL:20–27.

<i>Catholic</i>	<i>Baptist</i>	<i>Chinese</i>	<i>Matisoff</i>
sh / s	sh / s	sh / s	š
y / z	y / z	y / r	y
f	f	f	f
v	v	v	v
l	l	l	l

*Notes:*

- The missionary transcriptions use 10 phonemically unnecessary consonantal symbols: **pf, hpf, bv mv, ny, tc, ts, tz, s, z**. On the other hand, my transcription of the consonants requires three special symbols: **ŋ, š, ġ**.<sup>3</sup>
- While an accurate phonemic transcription is indispensable for historical/comparative work, it could be claimed that a practical orthography should not require the user to make complex deductions, such as interpreting /š/ as [s], or /yi/ as [z].
- The Catholic transcription “gn” of the palatal allophone [n̠] of /n/ before /i/ reflects Franco-Italian influence.

*Table II. Vowel Equivalences in Lahu Orthographies*

<i>Catholic</i>	<i>Baptist</i>	<i>Chinese</i>	<i>Matisoff</i>
a	a	a	a
i	i	i	i
u / ö	u / uh	u	u
e	e	e	e
o	o	o	o
è	eh	ie	ε

---

<sup>3</sup> I would now have no objection to replacing them by “ng”, “sh” and “gh”, respectively.

<i>Catholic</i>	<i>Baptist</i>	<i>Chinese</i>	<i>Matisoff</i>
ò	aw	aw	ɔ
ü / ö	ui / uh	eu	i
ë	eu	eu	ə

*Notes:*

- The vowels /i/ and /ə/ are correctly distinguished in the missionary orthographies, but are confounded in the Chinese spelling, which writes them both as “eu”.
- Both missionary transcriptions confuse the allophone of /u/ after labials (phonetically [ɯ]), with the phonetically somewhat similar allophone of /i/ after palatals (phonetically [ɨ]), writing them both as “ö” (Catholic) or “uh” (Baptist).
- In the Catholic spelling, the mid vowels /ɛ ɔ/ are written with grave accents “è ò”, while the central vowels /i ə/ are written with diaereses “ü ë”. The use of diacritics to indicate vowel qualities is also characteristic of standard Vietnamese orthography, invented by the French Catholic Alexandre de Rhodes in the 17<sup>th</sup> c. As in Vietnamese, this means that some vowels must carry two diacritics, one for the quality and one for the tone.

***Table III. Tone Equivalences in Lahu Orthographies***

	<i>Catholic</i>	<i>Baptist</i>	<i>Chinese</i>	<i>Matisoff</i>
mid	ca <sub>ˉ</sub>	ca	ca	ca
high rising	ca <sup>ˊ</sup>	ca <sup>ˉ</sup>	caq	cá
high falling	ca <sup>ˋ</sup>	ca <sup>ˋ</sup>	cad	câ
low falling	ca <sub>ˋ</sub>	ca <sub>ˋ</sub>	cal	cà
very low	ca	ca <sub>ˉ</sub>	cal	cā
high checked	ca <sup>ˆ</sup>	ca <sup>ˆ</sup>	cat	câ?
low checked	ca <sub>ˆ</sub>	ca <sub>ˆ</sub>	car	cà?

*Notes:*

- In both missionary orthographies, the tone-marks are placed one space after the vowel, which makes it impossible to run syllables of the same word together, and unesthetic to join syllables within a word with hyphens, so that there is no way to tell where one word ends and the next one begins.
- In my orthography the tone-marks appear above the vowel. In the case of the two checked tones, part of the tone-mark is a post-vocalic glottal stop, but this offers no impediment to hyphenization. I regard the checked tones as forming a distinct subsystem from the open ones (they derive from proto-syllables with final stops), so that no synchronic or diachronic connection is implied between the high falling and high checked tones, or between the low falling and low checked tones.
- The missionary systems use both superscript and subscript tone-marks, which makes them hard to type and easy to confuse with each other.
- The tone-marks in the Catholic system are particularly cumbersome, since several of them are not standard diacritics at all, but rather curved lines or hooks meant to suggest mnemonically the actual contours of the tones.
- The Chinese system of marking tones by arbitrary post-vocalic consonants has much to recommend it. Such transcriptions were popular in the U.S. in the 1940's and 1950's for languages like Thai and Burmese, and similar romanizations are now in use for Hmong and Mien. It makes it possible to write polysyllabic words with no spaces between the syllables, usually without ambiguity as to the syllable boundary. It is eminently printable, typeable, and word-processible. In its earlier version it was seriously flawed by writing the low falling and very low tones with the same letter -l. This has now been rectified by using -f for the latter, as urged in Matisoff 1984.

Here is a sample sentence in all four orthographies:

*Catholic*

ya<sub>˘</sub> gni<sub>˘</sub> ò      nga<sub>˘</sub> hü<sub>˘</sub> te<sub>˘</sub> qha<sup>ˆ</sup>      qhò<sub>˘</sub> phò<sub>˘</sub>      sha      ca<sub>˘</sub> gha<sup>ˆ</sup> e tu<sub>˘</sub> le

*Baptist*

ya<sub>˘</sub> nyi aw<sub>˘</sub>      nga<sub>˘</sub> hui te<sup>ˇ</sup> hk'a<sup>ˆ</sup>      hk'aw<sub>˘</sub> hpaw<sup>ˇ</sup>      sha<sub>˘</sub>      ca g'a<sub>˘</sub> e tu<sub>˘</sub> le

*Chinese*

yar ni awl      ngal heu ted qhat      qhawl phawd      shal      ca xar e tul le

## *Matisoff*

yà?-ni ̄	ḡà-hi tē qhâ?	qhò-phô	šā	ca ḡà? e tū le
today	our village	where	game	go and hunt (IRR., QST)
NP <sup>1</sup>	NP <sup>2</sup>	NP <sup>3</sup>	NP <sup>4</sup>	VP

‘Where shall our village go to hunt today?’

## **V. Alphabetical order**

In my aforementioned youthful zeal, I thought it would be helpful to the Thai Lahu whose language I was studying to use for my dictionary an alphabetical order similar to that of Thai,<sup>4</sup> that is, an Indic-inspired *devanāgarī* order determined by the position of articulation of the initial consonant, proceeding from the back of the mouth to the front: i.e., first the stops and nasals in the order uvulars, velars, palatals, dentals, labials; then the fricatives (from back to front), and finally the liquid. Across the bottom of each printed page were “running feet”, listing the phonemes in alphabetical order, first the tones (using the carrying-vowel a), then the vowels, then the consonants:<sup>5</sup>

a á â à ā â? à? a i u e o ε ɔ i ə q qh k kh g ḡ c ch j t th d n p ph b m h ḡ š y f v l

However, relatively few Lahu are literate in Burmese or Thai. On the other hand, most literate Lahu in Burma and Thailand have learned the roman alphabet in missionary schools,<sup>6</sup> while Lahu children in China learn their ABC’s during their first year of schooling, since the roman alphabet is necessary for the *pinyin* transcription of Chinese.

In fact the concept of “alphabetical order” itself turns out to be extremely difficult to convey in Lahu.

## **VI. Metalinguistic explanations**

By the late 1970’s I began to realize that the unfamiliar transcription and alphabetical order would make my dictionary unusable even by the most literate Lahu. Since I felt it was too late to change my already voluminous manuscript, I decided to compose a long introduction in Lahu, written in the standard Baptist orthography, justifying my own transcription and alphabetical order on linguistic grounds. But how could one create the metalinguistic terminology necessary for such explanations?

---

<sup>4</sup> As well as to the other “Indospheric” literary languages of Southeast Asia, including Cham, Javanese, Mon, Khmer, Burmese, and Tibetan, and even Japanese *kana*.

<sup>5</sup> Morphemes beginning with a vowel are alphabetized before all those that begin with consonants.

<sup>6</sup> Lewis’ valuable dictionary (1986) lists the Lahu entries in the roman alphabetical order, each accompanied by a Thai gloss as well as an English one.

One possible solution would be to borrow morphemes from a major literary language. After all, most technical linguistic terms in English are borrowings from Latin or Greek, while Burmese, Khmer, and Thai linguistic terminologies are largely of Pali/Sanskrit origin. Yet the geographical dispersal of the Lahu people makes the choice of donor language difficult. Should one use Burmese, Thai, or Chinese borrowings?

All in all it seemed to me that it would be better to use native Lahu lexical material as far as possible. Clearly the way to do this would be to coin new *compounds* or *collocations*, since monosyllabic morphemes tend to be highly polysemous. As an extreme example, consider the morpheme **khô**, for which eight subsenses are distinguishable, depending on the collocations in which it occurs:<sup>7</sup>

1. ‘sound; noise’ **bôʔ-khô** ‘sound of shooting’ (**bôʔ** ‘shoot’); **phî-lò-khô** ‘sound of dogs barking’ (**phî** ‘dog’, **lò** ‘bark’)
2. ‘voice’ **khô ší ve** ‘be hoarse’
3. ‘words; something spoken; a stretch of speech’ **qa-mî-khô** ‘song / poem’; **tô-pháy-khô** ‘riddle’ (**tô** ‘word / speech’, **pháy** ‘hide’)
4. ‘language; variety of speech’ **Lâhū-khô** ‘Lahu language’; **šu-khô** ‘foreign language’
5. ‘word; turn of phrase; sentence’ **khô tê khô** ‘one word / one phrase / one sentence’; **khô-ŋe** ‘short form (of a compound word)’; **khô-yî** ‘long form (of a compound)’
6. ‘words of advice; words of good sense’ **khô na ve** ‘listen to advice’ (**na** ‘listen’); **chə-mô-khô** ‘wisdom of the elders / proverb / wise saying’ (**chə-mô** ‘elder’)
7. ‘probable reason’ **yô mâ šī ve khô** ‘It’s probably because he doesn’t understand.’ (**yô** ‘3<sup>rd</sup> person’, **mâ** ‘not’, **šī** ‘know / understand’)
8. [*neolog.*] ‘tone (of a tone language)’. [See Section VII.]

In the end, the task of writing this explanatory Lahu introduction proved to be too daunting, even with the patient help of Yâ-pā-é,<sup>8</sup> my best consultant during my 1977 fieldtrip, and the introduction never made it into the published dictionary.

---

<sup>7</sup> See Matisoff 1988:380–383.

<sup>8</sup> See Matisoff 1988:1257.

## VII. *New hope*

The eighth sense of **khô** above is actually a good example of a successful neologism introduced by a missionary.<sup>9</sup> The general word for ‘tone’ is now the widely accepted four-syllable expression **khô-mu-khô-nè** (lit. “high-sounds-low-sounds”), which follows the ABAC pattern typical of hundreds of other Lahu “elaborate expressions”, containing both a repeated syllable and an “elaborate couplet”, i.e. a pair of semantically correlative morphemes, usually antonymous (as here) or synonymous.<sup>10</sup> The seven Lahu tones are all conveniently nameable, using the simple morphemes for ‘middle’ (**ô-qô-ji**), ‘high’ (**mu**), ‘low’ (**nè**), ‘long’ (**yî**), and ‘short’ (**ŋe**):

<b>khô ô-qô-ji</b>	mid tone	33	(“middle tone”)
<b>khô-mu-yî</b>	high rising tone	35	(“long high tone”)
<b>khô-mu</b>	high falling tone	53	(“high tone”)
<b>khô-nè</b>	low falling tone	21	(“low tone”)
<b>khô-nè-yî</b>	very low tone	11(2)	(“long low tone”)
<b>khô-mu-ŋe</b>	high checked tone	4-?	(“short high tone”)
<b>khô-nè-ŋe</b>	low checked tone	2-?	(“short low tone”)

Such successes have served as models for my own sporadic neologistic efforts over the years, but it was not until very recently that I saw real hope that there could be rapid progress in creating metalinguistic terminology for Lahu. At ICSTLL #42 in Chiang Mai (Oct. 2009), it was my great pleasure to meet a group of Lahu graduate students in linguistics at Payap University. One of them expressed great interest in making my dictionary more accessible. We have been in touch ever since, trying out neologisms on each other, hoping to come up with a technical linguistic vocabulary that is both easy to understand and esthetically pleasing.<sup>11</sup>

## VIII. Examples of suggested phonetic terminology<sup>12</sup>

It is not problematic to use native Lahu *nouns* for the various vocal organs, e.g.:

1. **ô-tô?-lô?** ‘palate’ (**ô-tô?-lô? he ve** ‘hard palate’, **ô-tô?-lô? nû ve** ‘soft palate/velum’)

<sup>9</sup> Probably Paul Lewis, to whom also belongs the credit for coining Lahu equivalents of such musical terms as *note*, *rest*, *soprano*, etc., which have enabled the Christian Lahu to sing hymns in four-part harmony.

<sup>10</sup> Other elaborate expressions are of the form ABCB, where it is the 2<sup>nd</sup> and 4<sup>th</sup> syllables that are identical. For a list of nearly a thousand of these elaborate couplets, see Appendix III of Matisoff 2006:403–448.

<sup>11</sup> There is a good chance that he will be able to spend some time at Berkeley next spring so that we can work together toward this goal.

<sup>12</sup> Listed in my Lahu alphabetical order as an exercise for the reader!

2. ð-šá ‘air/breath’ (< PTB \*sak [HPTB 317]; ð-šá m̄ə? ve ‘blow breath; aspirate’, šá ġə ve ‘breathe’, šá-ġə-kh̄ə ‘laryngeal; h’)
3. qá-qə-lə ‘larynx/glottis’; cì ‘tooth/teeth’; nā-qh̄ə ‘nose’; m̄ə?-qə ‘mouth’
4. m̄ə(-gì) ‘lip(s)’ (m̄ə-gì-kh̄ə ‘a labial sound’; m̄ə-gì-m̄ə?-kh̄ə ‘labial stop’)
5. ha-tē ‘tongue’ (ha-tē-m̄ə ‘tip of the tongue/apex’, ha-tē cí-n̄ə? k̄i ‘tongue-root/dorsum’)

Several *spatial nouns* can easily be used in articulatory descriptions, e.g. for vowel frontness and backness: (ð-)ġū-šī ‘front’, (ð-)qó-ji ‘central’, (ð-)qh̄ə?-n̄ə ‘back’.<sup>13</sup>

- (ð-)ġū-šī ve bù-kh̄ə, ‘front vowel’, etc.

Most important is the recruitment of *verbs* for articulatory phonetic duty:

1. k̄i ‘melt’ (< PTB \*s/m-grəy [HPTB 189–90]) – bù-kh̄ə n̄i mà k̄i dà? á ve ‘fused vowels’
2. cú è ‘tightly closed; puckered up’ – m̄ə?-qə cú è te ve ‘pucker the lips’; bù-kh̄ə m̄ə?-qə cú è ve ‘rounded vowel’
3. te ‘put down / set on / establish’ (< PLB \*?d̄iŋ<sup>1</sup> < PTB \*m/s-d̄iŋ [HPTB 307–8]) – t̄e-kh̄ə ‘consonant’
4. t̄ə? ‘emerge’ (< PTB \*s-twak [HPTB 321]) – nā-qh̄ə c̄e t̄ə? kh̄ə ‘nasal sound’; ð-k̄é-l̄é t̄ə? kh̄ə ‘lateral sound’
5. d̄i? ‘vibrate’ (< PLB \*N-tit<sup>H 14</sup>) – qá-qə-lə d̄i? ve ‘larynx vibrates’
6. ní-tī ‘immobilize by pressing / squeezing’ – ní-tī k̄i ‘point of occlusion’ (SYN. m̄ə? dà? k̄i)

ní ‘squeeze / press’ (< PLB \*s-nyit [HPTB 349])

tī restrain / hold back / fix in place’ (perhaps ≈ tī ‘to plant’)

<sup>13</sup> For the dimension of vowel *height* it is better to use adjectival verbs (as with the names of the tones): mu ve bù-kh̄ə ‘high vowel’, ð-kā-lé ve bù-kh̄ə ‘mid vowel’. Note that (ð-)qó-ji and (ð-)kā-lé both mean ‘middle’, but are here artificially repartitioned so that one refers to the front/back dimension and the other to the high/low one, just as the English terms *mid* and *central* are arbitrarily assigned to a particular vocalic dimension.

<sup>14</sup> This is an allofam of PLB \*?-dit<sup>L</sup> ‘whistle/trill’ [HPTB 349] > Lahu tí-šī? m̄i ve ‘whistle’, i.e. Lahu tí- ≈ d̄i?.

7. **bù** ‘make a sound / resound / be loud’ – **bù-khô** ‘vowel’; **bù-khô-kè** ‘pure vowel / monophthong’ (**kè** ‘clean’); **bù-khô-tì** ‘diphthong’ (**tì** ‘muddy’)
8. **mè?** ‘close abruptly’ (< PTB \*s-mi:t [HPTB 350]) – **mò?-qə mè? khô** ‘a stop’; **qá-qə-lə mè? khô** ‘glottal stop’; **mê-gì mè? khô** ‘labial stop’; **mè? dà? ve** ‘occlude’<sup>15</sup>
9. **mê?** ‘blow’ (< PTB \*s-mut [HPTB 364]) – **ə-šá mê? ve khô** ‘an aspirate’; **ə-šá mâ mê?** unaspirated’
10. **šô?** ‘rub / stroke’ (< PLB \*sap<sup>H</sup> [HPTB 337]) – **šô? bù ve khô** ‘fricative’

#### *Other morphemes*

- dà?** (Verb particle) ‘reciprocal action’ – **mè? dà? ve** ‘occlude’; **kì dà?** ‘fuse; melt together’
- cə** (Classifier) ‘kind’

This classifier (actually an old loan from Tai) can be used to refer to *manner of articulation*, e.g.: **Lâhū-khô mè?-khô šê? cə cə ve** ‘Lahu has three kinds (sc. manners) of stops.’

#### **IX. Desiderata for the future**

Much remains to be done, even with respect to phonetic terminology. Once a satisfactory set of phonetic terms has been developed, we will have to devise ways to discuss such phonological concepts as contrast, complementary distribution, free variation, phoneme, allophone, etc.

The next step will be to proceed to morphological and syntactic terminology—no small task in view of the fact that my grammar and dictionary use quite a large number of sometimes idiosyncratic form-class and grammatical concepts. The Index of Abbreviations for these runs to some seven pages (GL:xxxix–xxxvii).

It is my fond hope that the new generation of Lahu linguists will soon feel comfortable discussing fine points of their language’s structure in Lahu itself. If so, perhaps my academic approach to Lahu will have had a positive aspect after all!

---

<sup>15</sup> A possible alternative to this verb is the compound **ní-tī** [q.v.].

## References

- Lewis, Paul. 1986. *Lahu-English-Thai Dictionary*. 556 pp. Thailand Lahu Baptist Convention. Bangkok: Darnsutha Press.
- Matisoff, James A. 1973/1982. *The Grammar of Lahu*. Berkeley, Los Angeles, London: University of California Press. xxxvii + 693 pp. (“GL”)
- . 1984. “Some problems in the orthography of Lahu.” Translated into Chinese as “拉祜语文字方案中的若干问题 *Lāhù-yǔ wénzì fāng'àn zhōng de ruògān wèntí*” by Zhao Yansun 赵衍荪. *Minzu Yuwen* 1984.3:27–38.
- . 1988. *The Dictionary of Lahu*. Berkeley and Los Angeles: University of California Press. xxv + 1436 pp. (“DL”)
- . 2006. *English-Lahu Lexicon*. Berkeley, Los Angeles, London: University of California Press. xxxvi + 472 pp. (“ELL”)