



Writing Tamang

Martine Mazaudon

► To cite this version:

Martine Mazaudon. Writing Tamang. Syomhendo (Kathmandu, Nepal), 1993, July-September, pp.15-26. halshs-00168957

HAL Id: halshs-00168957

<https://shs.hal.science/halshs-00168957>

Submitted on 30 Aug 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Writing Tamang. A brief note on alphabetisation, spelling and transcription in Tamang.¹

Martine Mazaudon

I have been asked to give some advice on the development of a writing system for Tamang and tools to use in the classroom in order to teach reading and writing in Tamang. Further development of the use of Tamang (or for that matter other minority languages of Nepal) in education at a higher level will be outside the scope of my remarks.

Some valuable attempts have already been made to transcribe Tamang. I know of two systems presently and there may be more. I will say from the start that there are several possible solutions, and will outline some of the possibilities and difficulties with each. I will also say that the choice between possible solutions is not entirely technical but also largely political, and on that aspect I will refrain from expressing any opinion at all.

From the technical point of view, we will have to distinguish between what would be a perfect transcription from the point of view of the linguist, and what will be practical for the user. If we can achieve both all the better; if we cannot we will have to compromise. I will first discuss what would constitute completely correct transcriptions of the language, basing my analysis on Eastern Tamang, mainly the dialect of Risiangku, and see if the same system can be used for all dialects, in so far as I am informed on the other dialects.

Transcription and spelling

Ideally, a transcription of a language should represent each individual sound used in that language by a unique letter or symbol, or, less desirably, by a unique combination of letters or symbols. Such a system should be easy for speakers to learn, because in principle the speaker already “knows” the sounds of his language, although it may take him a while to get used to analysing words into individual sounds. All he has to learn is the equation between each individual sound (there are often about 40 of these) and the letter or combination of letters used to write it.

The spelling systems (or “orthographies”) which are actually used for the world’s languages are more or less close to this ideal. Nepali spelling is quite good as a transcription, especially if compared to English or French or Tibetan. In English, for example, there are many ways of writing the same sound, for example the sound [u] (as a linguist would write it), which occurs in the words *to*, *too*, *two*, *through*, *blue*, *you*, *crew*, and the same letter or combination may represent different sounds in different words. One reason for these anomalies is that the pronunciation of languages changes

¹ Published in *Syomhendo*, July-September 1993, 15-26 (Kathmandu, Nepal)

over time, so that the older a spelling system is, the farther it is likely to be from an ideal transcription of the modern pronunciation. Another reason is that foreign spellings may have been kept for borrowed words, like French words in English or Sanskrit words in Nepali. Whatever the reason, learning such a system in school takes up time that could be used to teach mathematics or geography. So why not change it? Here people will object that thousands of books have been published in the old spelling system and that changing it would cut off the younger generations from their traditional literature.

In a language which has not been written before, this question does not arise, and it is advisable to adopt a transcription principle such that each different sound is represented by a different letter, and only one. This is what is done by linguists when they use the International Phonetic Alphabet (IPA). This alphabet, which is based on the Roman alphabet, has been designed on purpose to have a very large number of symbols, or letters, and can provide enough different symbols to write down any language correctly. On the other hand it is not available in most printing presses, and I would not recommend using it for everyday transcription of Tamang, for practical reasons.

So the best thing to do would be to use one of the local alphabets, which are already familiar to a large number of people in Nepal, and which can be printed locally, and adapt it to fit the Tamang language. What I will discuss here is how much adaptation is required and what problems that adaptation may encounter.

What is Tamang linguistically?

Tamang is a language of the Tibeto-Burman family, belonging in the same branch as classical Tibetan, but it is not a descendant of classical Tibetan; it is not a Tibetan dialect. To use a family metaphor, Tamang is a grand-nephew of Classical Tibetan, not a grandchild. So one option is *not* available, that is writing Tamang etymologically by using Tibetan spelling. Let us take an example.

The word for 'eight' in Classical Tibetan is spelled *brgyad*. In Lhasa Tibetan it is *gye*, in Tamang *brat*. These three forms are related, in technical terms they are 'cognates'. But while there are regular rules which allow the transformation of the Tibetan spelling into its Lhasa pronunciation, such rules cannot be made for Tamang. In Lhasa Tibetan we can learn that in a word which begins with *brgy-* only *gy* is pronounced. This is the case in the word 'eight' pronounced *gye*, and also in other words, like '100', written *brgya* and pronounced *gya*. In comparing Tamang to Classical Tibetan, we could observe that in the word 'eight' the initial group *brgy* is pronounced *br*. But the word for '100' in Tamang is *gyarca*, with *gy* instead of *br* corresponding to Written Tibetan *brgy*. So while in Lhasa a single pronunciation rule can be learned, in Tamang we would need two different rules: '*brgy-* is sounded *gy*', for the word '100', but '*brgy-* is sounded *br*' for the word 'eight'. This will not make it easier for Tamang children to learn how to read and write!

On the other hand, the Tibetan alphabet, simply as an alphabet, not as a spelling system, is one of the possible choices. In that option, 'eight' in Tamang would be written བྱེ or something like that, as we will see later, and *not* བྱེན་.

Tamang speech sounds and possible writing methods

Two different non-Roman alphabets are more or less familiar to people living in Nepal, and especially to Tamangs: Devanagari, which is used to write Nepali, and the Tibetan script, itself derived from an ancestor of Devanagari many centuries ago. We will discuss transcriptions based on these two alphabets. Our first task, however, will be to study the sounds actually used by Tamang speakers to make up words – in technical terms, the 'phonology'. There are some differences between dialects, which we will discuss later, but they are not too great.

Let us start with the most difficult part, so the rest will seem easy afterwards. Tamang has tones. Nepali and English do not. Tibetan has some tones too, but 1) they are different from Tamang tones and 2) they are not transcribed as such in the Tibetan writing system. So whatever alphabet is used, some addition will be necessary if we want the transcription of Tamang to be completely correct.

Tones

All dialects of Tamang distinguish 4 tones. Their phonetic description here is based on the Risiangku dialect, but, with some reflection, speakers of other dialects can discover them in their own speech. Since there are some words which are not used in one or the other dialect, the examples given here will not work for everybody. You have to try and understand the principle and look for examples in your own dialect. If you do not speak Tamang or do not speak it fluently you should ask a village person to help you find examples. An illiterate grandmother might be a good choice.

One pair of words which seems to be found in all dialects, at least in the Eastern ones, is 'to sell' and 'to catch or tie up an animal'. Both could be transcribed as *cuŋ-ba* in the phonetic alphabet², in Roman letters, or as चुङ-बा in Devanagari, or as རྩུང་བ་ in the Tibetan alphabet. The problem is that these two words are not really homonyms. If a Tamang speaker tells another:

ŋa-i ra cuŋ-ci 'I - goat - sell/tie up - past'

the hearer will be able to understand whether the goat was sold or tied up. There is a difference in the way the syllables *cuŋ-ci* are pronounced. That difference is tone. It consists mostly of differences in the pitch, or level of the voice, in the speed of utterance, in the direction of the pitch change during the pronunciation (starting higher and becoming lower or remaining flat, or rising).

Let us use a new sign to indicate this difference, for instance a number. We could write ¹*cuŋ-ba* for 'to catch and tie up' and ²*cuŋ-ba* for 'to sell'.

When a language has tone, it is not just a few words that use it to distinguish meaning; the whole vocabulary is affected. For instance the words for 'ear of corn or wheat' and 'rain', both *nam* if we do not write tone, should be distinguished. One should be marked with the ¹tone and the other with the ²tone. By listening to yourself saying these two words, and comparing their melody with that of ¹*cuŋ* and of ²*cuŋ* you should be able (after some practice and effort!) to decide which tone to write. The answer in this case is ¹*nam* 'ear of grain' and ²*nam* 'rain'. In the same way 'bile' is ¹*kam* and 'chin'

² I use a hyphen in phonetics as well as in Devanagari to separate a root from its suffixes. In text orthography this can also be done or ignored.

is ²*kam*; 'to stop a river' is ¹*paŋ-ba* and 'to speak' is ²*paŋ-ba*; 'to go away' as a respectful form is ¹*see-ba*, but ²*see-ba* means 'to know'. Words which do not share their consonants and vowels with another word, as in the examples above, still are pronounced with one of the tones of the language. For instance, the word *mar* meaning 'down' or 'downstream' is pronounced with the same tone as ¹*cuy* 'to sell', ¹*nam* 'ear of grain' and ¹*kam* 'bile', and should receive the tone mark ¹: ¹*mar*. I do not know of any word pronounced ²*mar* in Tamang, except the root of the verb ²*mar-ba* 'to extract oil by pressing' (in some dialects only). Even though no confusion in meaning would occur if 'downstream' were written ²*mar* or even simply *mar*, the correct complete transcription is ¹*mar*. The fact that neighbouring languages do not distinguish tone, so that there is no known way to write it, makes it seem as if it were not important. But actually it is just as important to pronounce (and write) tone ¹ and tone ² differently as it is to distinguish between the vowels *a* and *o*. Once you become conscious of them, tones will be as natural to write as consonants and vowels.

We are not quite done with tones yet. All the words quoted until now, whether transcribed with tone ¹ or tone ², are pronounced, in the great majority of Tamang dialects, with a high pitch, something that the Tibetan grammarians have called 'feminine' voice. We are now going to consider words which are pronounced with a low pitch (or 'masculine' voice). Take the words for 'gold' and 'butter'; here again consonants and vowels could be written as *mar*, just as for 'downstream'. It will be clear to all Tamang speakers, that 'gold' and 'butter' are rather pronounced with a 'masculine' voice or low pitch, while 'downstream' is pronounced with a higher voice, more typical of 'feminine' voices. This is a tone difference again. The tone of 'butter' and 'gold' is different from that of 'downstream'; it is also different from the tone of the words that we have written with the mark ², like ²*nam* 'rain'. In addition, the tone of 'butter' is different from the tone of 'gold'. So we need to introduce two more tone marks. In Roman characters, let us use the numbers ³ and ⁴. We will write ³*mar* for 'butter' and ⁴*mar* for 'gold'. Now all words which are pronounced with a 'masculine' voice need to receive one of these two marks. In this way 'to bite' will be written as ³*cii-ba* and 'to remember' will be ⁴*cii-ba*. That is all: there are only 4 tones in Tamang.

In addition to their low pitch, tones ³ and ⁴, in most dialects, have a redundant phonetic feature which is technically called 'breathy voice' or 'breathiness'. To simplify, it makes a word like 'one' ⁴*kik* (or ⁴*ki* depending on the dialect) sound somewhat as if it started with a Nepali ँ (or sometimes ँ). The vowel itself does not quite sound like the *i* in ²*pit-pa* 'to send' for example. The impression of 'blowing' or of heavy breathing — the breathiness — extends over the vowel too. In Tamang pronunciation the difference between voiced consonants (*g, j, ɟ, d, b*) and voiceless consonants (*k, c, t, t, p*) is not used alone to differentiate words. It always comes along with a difference in tone. So that if we write tone we do not need to use the letters *g, j, ɟ, d, b* at all in transcribing Tamang, although we can do it redundantly. (We will see later that conversely, if we write a difference in consonant voicing, we may not have to write some of the tone features.)

Tone numbers

How can we write the tones then? We need 4 symbols which do not exist either in Devanagari or in the Tibetan script. We have seen that in phonetic transcription we can use numbers. Can we do the same in Devanagari? We could incorporate into the Tamang alphabet the four symbols १ २ ३ ४ to write the tones, placed on the line or slightly above the line, in the following way:

¹ kan	१कान	or	१कान	'rice'
² nam	२नाम		२नाम	'rain'
³ mi	३मी		३मी	'man'
⁴ me	४मे		४मे	'cow'

On two or more syllable words, only the beginning of the word needs to be marked: non-initial syllables and suffixes do not have their own tone. So :

³ sajmo	३साङमो	or	३साङमो	'copper'
¹ nana	१नाना		१नाना	'sister'
⁴ meme	४मेमे		४मेमे	'ancestor'
² yoma	२योमा		२योमा	'young female'

With this transcription the sentence 'I ate rice' would look like this: १ङी १कान १चाची

This system may look a little strange, but it has been used successfully for a major language like Thai, and for several minority languages in China. Its advantage, as we will see, is that it would be satisfactory for all dialects. Its disadvantage is that if people find it too strange they will simply omit the tone-number altogether and that will lead to confusions between words.

In the Tibetan script the same principle can be used. We can insert tone numbers at the beginning of words.

We will return to discussing alternative means of transcribing tone in Devanagari or in the Tibetan alphabet after we have introduced the consonants and the vowels. In the meantime I will use tone numbers, or leave the tones unmarked.

Vowels

Tamang vowels are simple. They require only 5 different symbols:

i, e, a, o, u

and a mark for long or short. In the examples above I have written the vowel twice to show length, as in the word ³cii-ba 'to bite'. The usual way in the International Phonetic Alphabet is to use a double dot, like ³cī-ba. This double dot exists also in Devanagari as the *visarga*. In Nepali, the *visarga* is not used to indicate that a vowel is long. But it is so used in Newari and could be used for that purpose in Tamang as in:

short	long		short	long
अ	अः	or in linked form	का	काः
इ	इः	(shown with k)	की	कीः
उ	उः		कु	कुः
ए	एः		के	केः
ओ	ओः		को	कोः

This is the system that has been proposed by the Tamang Language Council and used by Amrit Yoncan in the Adult Education Handbook (published by Action Aid) based on Eastern Tamang. It is also proposed by the Western Tamang journal *Tamang*. Like any

system this has advantages and disadvantages. One advantage is that the same phonetic element, vowel length, is always transcribed in the same way, which is easier to learn.

Another possibility is to use Devanagari in a manner as close as possible to its use in transcribing Nepali. In this option, the columns for short vowels would remain the same, but the columns for long vowels would be modified for those letters for which Devanagari offers a possibility of distinguishing long from short, namely *a*, *i*, *u*.

अ	आ	or in linked form	क	का
इ	ई		कि	की
उ	ऊ		कु	कू
ए	एः		के	केः
ओ	ओः		को	कोः

There is no long *e* or long *o* in Devanagari. So for these two vowels there is no other choice than to use the *visarga*. This is the system used by Doreen Taylor, and in the primers presently in use in the Trisuli area.

The advantages of this option would be that a written text in Tamang would look more like a text in Nepali, and I suspect that some people might think it more pleasing from an aesthetic point of view because it is more compact. From the pedagogical point of view, since the students are going to learn to read and write Nepali as well as Tamang, one might think that the less difference there is between the two systems, the easier it will be for the students. But the fact is that *dirgha* and *hrasva* are only orthographic headaches in modern Nepali, since they are pronounced exactly the same. So learning how to use them in Tamang on the basis of pronunciation would be no help when learning Nepali spelling.

The disadvantages of this option are as follows. First the system is not homogenous, since length is written differently for different vowels, and there is no typographical economy since *visarga* is needed anyway for *e* and *o*. Secondly, writing a vowel before a consonant as for Devanagari short *i*, may be pedagogically more difficult. Thirdly, the vowel *a*, as all Tamangs know, is different in Nepali and in Tamang. The difference in Nepali between अ and आ is not only a difference in length, but also and mainly a difference in quality. In the International Phonetic Alphabet, Nepali short *a* would be written [ə] or [ʌ], a central vowel, not as open as the Nepali long *a*, which would be transcribed as [a:]. The two *a* vowels of Tamang both have the quality of the Nepali long *a*, but one is short and the other is long. So it could be rather confusing to a Tamang student to see his own short *a* written with no symbol as in Nepali, and it could reinforce his usual mistake of pronouncing Nepali 'to die' [məɾnu] with a Tamang short *a*, making it sound like Nepali 'to kill' [ma:ɾnu]. Lastly, if we decide to have one vowel unwritten, the inherent *a*, the use of *halant* will be necessary at the end of each syllable which ends in a consonant, and to indicate consonant clusters. Since the use of *halant* in Nepali printed books is very sporadic, it is likely that Tamangs would also omit it, leading to confusions. So whatever we do for long and short *i* and *u*, I would recommend always writing the vowel *a*. Hence का for *ka* and काः for *ka:*.

Vowels in the Tibetan script

If we want to use the Tibetan alphabet, ཀ ཁ ག ཁྱ ཀྱ will transcribe ki, ke, ka, ko, ku correctly. But I do not know of a symbol equivalent to *visarga* which we could use for length. Some syllable-final letters of Classical Tibetan produce a long vowel in Lhasa Tibetan, -r for example. But since Tamang has a real final *r*, pronounced as *r*, we cannot use that symbol to show vowel length. Final -g in Tibetan also lengthens the vowel, as in *dmag* 'army', which is pronounced [ma:ʃ]. But this -g has 'side effects' in Tibetan; it produces a falling pitch (or tone), and we have seen that Tamang needs to distinguish between level and falling tones. The best choice might be to reserve the Tibetan *a-chung* for marking length, as follows:

short	long
ཀ	ཀྱ
ཁ	ཁྱ
ག	གྱ
ཁྱ	ཁྱྱ
ཀྱ	ཀྱྱ

Nasal vowels

Tamang has nasal vowels in a small number of words. They can be written easily using *anunasika* in Devanagari, as in the word for 'worms' རྩོམ་པ་ *kōtai* in phonetic script. Since nasal vowels are always long it is not necessary to write the *visarga* with them, as length will be understood from *anunasika*.

If we use the Tibetan alphabet, I believe there exists also a very rarely used dot (only in a few loan words from Sanskrit) which is drawn over the line to mean about the same thing as *anunasika*.

Diphthongs

Diphthongs are found in all dialects of Tamang. I believe that the maximum number of vowel combinations or diphthongs found in any dialect is the following:

ai au oi ui

In Devanagari such combinations can be written by using the linked form of the vowel for the first member, followed by the full form of the vowel for the second. Other systems would be possible, but since this is usual, it is best to keep it. So 'to fall' *'tai-ba* would be ཁྱེད་པ་, 'to belch' *'kau-ba* རྩོམ་པ་, and 'to plough' *'moi-ba* མཐོང་པ་. In Risiangku, the diphthong *oi* can sometimes sound more like *ue* or *we*. This is a free variation.

The vowel sequence *ui* can also be considered as a diphthong, རྩོམ་པ་ although it could also be written as a glide *w* followed by *i*, thus རྩོམ་པ་. For instance 'to carry', *'pui-ba* or *'pwi-ba* in phonetics, could be written either as རྩོམ་པ་ or as རྩོམ་པ་ (We will talk later about ligatures). If we write it as རྩོམ་པ་ the form of the verb root will be less changed in the different forms of the conjugation, for example in the imperative རྩོམ་པ་ *'puyo*. On the other hand, if most Tamang speakers feel that it sounds more like *wi* than like *ui*, there is no obstacle to writing རྩོམ་པ་. A decision on such points should only be made after experimentation, and variation in spelling can also be tolerated.

All of these diphthongs (including *ui*) can be preceded by a palatal vowel or glide (*i* or *y*) and by a labio-velar vowel or glide (*u* or *w*) in the following way:

<i>yai</i>	याइ	as in	^१ याइ-बा	'to return'	or in IPA ³	^१ <i>yai-ba</i>
<i>yoi</i>	योइ		^२ गयोइ	'pond'		^३ <i>kyoi</i>
<i>yui</i>	युइ		^२ कयुइ	'water'		^२ <i>kyui</i>
<i>yau</i>	याउ		^१ सयाउरी	'ant'		^१ <i>syauri</i>
<i>wai</i>	वाइ		^२ गवाइ-बा	'to brew beer'		^३ <i>kwai-ba</i>

y and *w* also occur in front of long vowels: *ya:*, *yo:*, *yu:*, *wa:*, and (in one word only) *ywa:*. This last combination will be a little clumsy to write. It occurs in Risiangku in the word ^२*sywa:* 'to sprinkle', which is pronounced as ^२*sya:* in some other dialects. This, depending on the village, will thus be written ^२सयवा:-बा or ^२सया:-बा

Diphthongs and triphthongs in the Tibetan script

Although Classical Tibetan does not have many groups of vowels, it has a way to write them by adding a vowel symbol to the *a-chung* used as support. We have proposed to use the *a-chung* as a mark of vowel length, but since Tamang diphthongs are always short there would be no confusion. The examples above could be transcribed as follows:

'to return'	^१ <i>yai-ba</i>	ཡའི་བ་
'pond'	^३ <i>kyoi</i>	ཁྱོའི་
'water'	^२ <i>kyui</i>	ཀྱའི་
'ant'	^१ <i>syauri</i>	ཤརྱའི་རེ་
'to brew beer'	^३ <i>kwai-ba</i>	ཁྱའི་བ་

Consonants

1. Plain consonants

Simple consonants can be written easily either with Devanagari letters or with Tibetan letters. There is no need for illustration in initial position. In syllable-final position, it should be remembered that among stopped consonants only *k*, *t*, *p* (the voiceless stops) are found in Tamang, never the voiced stops (*b*, *d*, *g*). So when transcribing Tamang in Devanagari we should use क त प rather than ग द ब at the end of syllables. In using the Tibetan alphabet, however, it is traditional to write the voiced consonants in syllable-final position, and never the voiceless ones. This usage can be kept as it does not lead to any confusion. So Tamang ^१*kik* 'one' would be written in Devanagari as ^४गीक *gik* and in Tibetan script as གིག *gig*.

³ I use *y* instead of the IPA symbol *j* for the palatal glide (the consonantal form of *i*) because this is more usual in Nepal and India.

Final consonants and halant

The use of *halant* with final consonants is not necessary if we decide to write all the vowels, including short *a*, and to have no inherent vowel. In this case, when reading मान 'man' 'medicine' we will know that it cannot be *mana* (the measure word) because the latter would be written माना.

The same is true of two consonants which come to be in contact at the boundary of two syllables, as in ³sajmo 'copper' रेसाङ्मो. Since we write all the vowels we will know not to pronounce it *sajamo*.

But if we chose to use the Nepali system with the short *a* unwritten we have to write *halant* every time as in रेसाङ्मो or मान्.

Intervocalic consonants

Consonants which occur inside a word, especially between two vowels, sound voiced in Tamang. For instance the verbal suffix *-pa*, most of the time sounds like *-ba*. This is just the opposite of what happens at the end of syllables: inside a word we do not find *k, t, p*, but only *g, d, b*, or else people can vary from one to the other or pronounce a sound in between, without creating any confusion. After we talk about tones we will see that if we wish we can write Tamang perfectly without ever using *g, j, ɟ, d, b*. Consonants inside a word could always be written with a symbol of the voiceless series *k, c, t, t, p*. For example चा-पा 'to eat', or, equally well, चा-बा.

In Eastern Tamang, in particular in Risiangku, it is necessary to use both aspirated and non-aspirated consonants inside words. For example 'dog' needs a *kh*: नाखी, while 'chicken' needs a *k* (or *g*) नाका. In most other dialects, only *k* is required, and 'dog' is नाकी.

2. Initial consonants and tones

The observant reader must have felt uneasy with many of the examples given, because I keep writing *k, c, p* in phonetics, and I often use *b, d, g* in Devanagari or Tibetan. We are coming to this problem now. You will remember that there is a connection between tone and initial consonant and we suggested that this could be used to propose alternative transcription systems.

Until now we have been using numbers to transcribe the tones in our phonetic and Devanagari transcriptions. This is the most straightforward way to indicate the tones, though it is quite arbitrary: ¹ simply represents the category of words that have the same tone as 'kan' 'rice', whatever that is phonetically. Here again are some examples, with two different Devanagari equivalents:

chest	¹ ku	कु	कु
chin	² kam	काम	काम
back	³ ko	को	को
thatch	⁴ ki	की	गी

The first Devanagari column above uses क with all tones, and this is perfectly adequate. We mentioned above, however, that with tones ³ and ⁴ this initial consonant usually sounds more like ग or घ. So we have used ग in the right-hand column above, although it is not necessary to do so. Similarly, with tones ³ and ⁴ we could write ज ड द ब instead

of च ट त प. This does not make the numbers unnecessary, however: they are still needed to distinguish रेग from र्ग, and they are also needed to distinguish words beginning with other consonants such as *m, y, r, l, s* etc., for example in ¹*nam* 'ear of grain' ¹नाम and ²*nam* 'rain' ²नाम; or in ¹*mar* 'downstream' ¹मार, ³*mar* 'butter' ³मार and ⁴*mar* 'gold' ⁴मार. Some more examples:

god	¹ <i>la</i>	¹ ला
eye	² <i>mi:</i>	² मी:
hill	³ <i>ri</i>	³ री
thief	⁴ <i>jo:</i>	⁴ यो:

Using consonants to write tone

The systems that have actually been used to write Tamang in Nepal so far usually ignore the difference between tones ¹ and ², and between ³ and ⁴. That is, they put the two 'high' tones together and the two 'low' tones together. Instead of using numbers, they write the initial consonants differently for the high and the low tones – just as we have shown above for क and ग.

In her forthcoming dictionary of Western Tamang, however, Doreen Taylor intends to add an extra sign, the apostrophe, to her system, to mark words with tones ² and ⁴, making them different from words with tones ¹ and ³, and thus making her transcription complete. We can add this sign to the Action Aid transcription as well. Here, then, are the two systems, with the apostrophe added:

	<i>phonetics</i>	<i>tone-numbers</i>	<i>continuants + dot stops in voiced form + apostrophe</i>	<i>continuants + h stops in voiced form + apostrophe</i>
god	¹ <i>la</i>	¹ ला	ला	ला
eye	² <i>mi:</i>	² मी:	'मी:	'मी:
hill	³ <i>ri</i>	³ री	री	रही
thief	⁴ <i>jo:</i>	⁴ यो:	'यो:	'चहो: or 'यहो:
chest	¹ <i>ku</i>	¹ कु	कु	कु
chin	² <i>kam</i>	² काम	'काम	'काम
back	³ <i>ko</i>	³ को	गो ⁴	गो
thatch	⁴ <i>ki</i>	⁴ की	'गी	'गी
earth	¹ <i>sa</i>	¹ सा	सा	सा
three	² <i>som</i>	² सोम	'सोम	'सोम
to make	³ <i>so-pa</i>	³ सो-पा	सो-पा	सहो-पा
mushroom	⁴ <i>syamo</i>	⁴ सयामो	'सयामो	'सहयामो or 'सह्यामो

⁴ The use of a dot to indicate low-breathy tone could be extended to stops consonants too, thus 'back' could be written को and 'thatch' 'की.

Both systems use ग ज ड द ब for tones ³ and ⁴. For other initial consonants, Doreen Taylor writes an *h* after the consonant (because of the “breathiness”) and Amrit Yoncan writes a dot under the consonant. Thus instead of marking the difference between high (¹, ³) and low (³, ⁴) tones directly, by a tone symbol, they mark it indirectly, by using two sets of initial consonants. Neither of them marks the difference between ¹ and ², or between ³ and ⁴, when writing ordinary Tamang text, although the apostrophe is available for this purpose when more accuracy is needed, as in a dictionary. Thus both of these systems make a compromise. This leads of course to some homographs (words which sound different but are written the same), but any language can tolerate a reasonable amount of homographs and homonyms.

Continuant initials

In the transcription of continuant initials on the low tones, the advantage of using the consonant *h* rather than the dot is that a dot, like a tone number is a small unusual sign that people may be tempted to omit when they write. The disadvantage of using *h* is that, although it is no problem with *m*, *n*, *ŋ*, *s*, it interferes with the transcription of some clusters. For example, the word for 'hill', ³*ri* in phonetics, has breathy voice like all tone ³ words and would be written रही ⁵. The word ³*la* 'hillside' would be written लहा or ल्हा in that system. The word ³*wapa* 'rooster, male' would be written बहा-पा. In Eastern Tamang, there are a few words which begin with an aspirated *r* (or *hr*), aspirated *l* (or *hl*), aspirated *w* (or *hw*) and so on. We probably want to use the *h* symbol for those. Consider the word ¹*hri* 'the spirit of a dead child'. To keep it distinct from low tone ³*ri* 'mountain', we could decide to write the *h* first: हरी (or ही). In some dialects, like Taglung in the North of the Kathmandu Valley, there are numerous words with initial aspirated *w*, or *hw*, which need to be distinguished from low tone breathy *w*. For example 'to cry' in that dialect is ¹*hwa-pa* (corresponding to ¹*kra-pa* in most other dialects). This would have to be written हवा-पा. If we use the dot for breathiness, the *h* remains free for aspiration. Otherwise we shall have to make a rule about the order of the *h* as compared to the initial.

Aspirated consonants

Aspirated consonants (*k^h*, *c^h*, *t^h*, *th^h*, *p^h*) occur only with the two high tones, that is without breathy voice. So they will never be written with the dot. Words which start with such initials only need to be differentiated for tone ¹ and tone ². We can use the apostrophe for this, as follows (or use tone-numbers):

¹ <i>p^hum</i>	'clod of earth'	फुम
² <i>p^hum</i>	'egg'	'फुम
¹ <i>c^ha-pa</i>	'to graze'	छा-पा
² <i>c^ha-pa</i>	'to hurt'	'छा-पा

⁵ I do not think it advisable to use the conjunct ह्र here as this means hr rather than rh, and as we will see presently we need to make that distinction.

Consonants and tones in the Tibetan alphabet

There are in Tibetan a large number of rarely used letters or symbols which are used mostly in loan words from Sanskrit, and with which I am not familiar enough to present them here. It is very likely that some such symbol could be found that could be used in the same way as we proposed to use the apostrophe in Devanagari. So I will leave this problem open for some further research. The interested reader can consult a learned lama on this matter. It should be remembered though that until such a symbol is added to the Tamang alphabet in Tibetan script, the transcription will miss half of the tonal features.

In order to transcribe the breathiness of low tones, we could use the series of voiced consonants for the stopped consonants, exactly as in devanagari. Ex. ཁོ for 'ko 'back'. The pronunciation of the series of voiced consonants in many dialects of Central Tibetan is very similar to what it is in Tamang, a devoiced consonant, with breathiness and low tone. So that would work very well. Tibetan also has a high-tone *s* ས and a low-tone *s* ན, a high-tone *sy* སྐ and a low-tone *sy* སྐ which can be used to distinguish Tamang high tones from low tones with these initials.

For continuants on the other hand there is a difficulty. The Tibetan *h* is voiceless, and when associated with other consonants it leads to high tone. The Nepali *h* is voiced, which made it more suitable for use as a notation of Tamang breathiness. Using the Tibetan *h* to indicate a low tone would be very counter intuitive for anyone knowing Tibetan. We would have to look for something else.

3. Consonant clusters

In syllable initial position, aside from the words in *hr*, *hl*, *hy*, *hw* which we have mentioned above, and which are not as numerous in all dialects, Tamang has frequent consonant clusters formed of a consonant followed by *y*, *r*, *l*, or *w*. There exist in Devanagari conjunct forms for all the clusters found in Tamang, and they can be used if we wish. On the other hand, most alphabets in the world keep the same form for the same letter in all contexts, and this solution is certainly pedagogically easier. As long as we always write the vowel *a*, there is no need to use conjunct letters, or to use *halant*. Example: 'kra-pa 'to cry' can be simply करा-पा. If short *a* is left unwritten, then *halant* could be used on the first member of a cluster: करा-पा.

When two consonants come into contact at the juncture of two syllables, the Nepali usage in Devanagari is to use conjuncts. This can be done in Tamang, but is not necessary. ²*tarṇa* 'fish' can be written 'तारडा (or रतारडा) and ⁴*palpa* 'toad' 'बालपा (or ढपालपा or 'पालपा according to the system chosen to write tones). The choice here is probably between aesthetic and pedagogical considerations. In a literacy program pedagogy may be more important. But here again experience in the classroom is the real teacher.

In the Tibetan script, syllables are separated by a dot, *tseg*, so there is no need for conjunct letters at the juncture of two syllables. Only in syllable initial position are clusters written in linked form. They are simpler than the Devanagari conjuncts in that the shape of the letters is less changed. Since we have used the *a-chung* to write length, we don't have another letter available to write short *a*. So in using the Tibetan alphabet, we will leave short *a* unwritten, as the inherent vowel. In this case it becomes

necessary to use the conjunct form of the letters in initial clusters. Example: *kra-pa* 'to cry' ཀླཤ་. It would not be possible to write ཀླཤ་ because that would read [*karba*].

Dialectal variation and regular correspondences, the problem of standardisation

Standardisation

It is a well-established fact that any language spoken over a significant area has different dialects, and Tamang is no exception. It is also true that most written languages have a standard written form that is used by speakers of different dialects. For many Tamangs, these observations immediately raise the question of which dialect of Tamang should serve as the basis for written Tamang.

If we look at the history of well-known languages, however, we often find that the establishment of a standard came long after the languages began to be written. When Europeans started writing in their native languages instead of writing in Classical Latin, there was a long period of great variability in spelling and usage before standards were established, and these standards often did not correspond precisely to any one spoken dialect. Similarly, in the beginning, when Tamangs first write their language, it is natural that each will write in his own dialect, since that is the only one he can be sure of; he has not had the experience of reading dozens of books in a literary standard. Imposing an artificial standard at this point might well discourage Tamangs from writing at all, which would be most undesirable. Later, as literary culture develops, writers will naturally gravitate toward the systems that seem most satisfactory or that have the authority of the best writers behind them, and in the end it should be possible to compromise on a standard which has already been well-tested by usage.

Dialect variation

In a number of cases it may be possible to use transcriptions which will match several dialects in spite of some differences in the pronunciation. Whenever that is possible, these should be preferred, but it may not always be practical. A good example is tone.

The phonetic pronunciation of the tones is not exactly the same in all dialects. Even among Eastern dialects some misunderstandings can occur. For example, tone 3 words like ³*mar* 'butter' pronounced by a speaker from Ramechhap were always recognized by a speaker from Risiangku in a test. But tone 4 words like ⁴*mar* 'gold' were often a problem, because the Risiangku speaker expected them to be lower in pitch. In other dialects, like that of Taglung, the difference may be much more extreme. Tone 4 in Taglung is not at all low in pitch, it has no breathiness, and the initial stop consonant (e.g. *p*, *t*, *k*), if any, is not voiced (i.e. not pronounced *b*, *d*, *g*). The word 'one' sounds like *ki* (की) not *gi* (गी or घी). In these circumstances, if we write arbitrary tone numbers (as we have done) they work for all dialects because the numbers only indicate categories of words which have the same pitch, melody, and voice quality, without saying what these are precisely. The spelling ལྷ་की is just as correct for Taglung, where the tone is high, as it is for Sahu, where the tone is low and where the *k*- is more like *g*-. This is because it is true in all dialects that 'gold', 'one', 'cow', and 'to

remember' are pronounced with the same tone, and differently from words which have tones 3, 2, or 1.

Vowels and consonants also vary between dialects. In many cases the variation is regular. For example, all words which end in *-at* in Risiangku end in *-e* in Taglung, Temal, or Sahu. Words which start with *kr-* in Risiangku, Sahu, and Temal, have *hw-* in Taglung. Words in *kw-* in Risiangku begin with *w-* in Sahu. And so on. We could try to invent an orthography such that words would be spelt the same for all dialects, and the pronunciation derived by regular rules. For instance, we could write *⁴kyat* for 'work', like the Risiangku pronunciation, and have Taglung speakers learn that, for them, *-at* is pronounced *-e*. I believe that this would be unnecessarily difficult and that it would be better, in the beginning at least, to write either *⁴kyat* or *⁴ke* depending on the dialect of the writer.

Classroom tools

It is always useful for a language to have a dictionary and a grammar for classroom use. There is a difference, however, between mother-tongue teaching and second language teaching. In the case of Tamang, most of the teaching will be mother-tongue teaching. In this case, particularly in the case of adult education, the students already know the language, including the grammar and the vocabulary, and only need to learn to put it on paper and to read it.

The main incentive to reading is to have interesting reading material. Since Tamang is a language whose traditions are mostly oral, Tamang "texts" exist in the form of stories, songs, religious formulas of lamas or jhankris, etc., but they are not available in written form. These texts could be written down and constitute reading material for the classroom. To write down traditional texts in a readable form two principles should be borne in mind: 1) In as much as these texts constitute Tamang culture and tradition, you do not want to distort them, and 2) since you want to use them for teaching you want them to be grammatically correct. On the first point, the danger is to try to remake the texts, to rewrite them with a personal interpretation which may not be correct. On the second point, observe that when a child or an inexperienced adult starts writing his own language, he writes grammatically incorrect sentences, which, if you think about it, is amazing, because when he speaks, he does not make such mistakes. The time taken by the process of writing itself is enough to distort the language: the writer loses track of his sentence as he writes it down. For traditional texts, a good way to proceed is to go to people who know the texts, your grandfather or grandmother for instance, and ask them to tell the story, song etc. into a tape recorder. Some people will be better than others at this, and all will get better after they get used to the tape recorder. After that, someone who may not know the texts, but knows how to write, can transcribe them, possibly checking difficult points with the original speaker. If you think you know the texts and can transcribe them too, you should still try to distinguish the two phases, the telling and the transcription. Of course, Tamang literature will not be limited to traditional texts, and writers of short stories, novels, poems, essays, biography, history, and books on practical subjects, should be encouraged as well.