ISO/IEC JTC1 **L2/20-259**

2020-08-24

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**
Title: **Proposal to add the Tangsa Script in the SMP of the UCS**
Source: **Stephen Morey (partly based on an earlier document by Anshuman Pandey)**
Action: **For consideration by JTC1/SC2/WG2**
Date: **2020-08-24**

1. Introduction: Tangsa is a script developed since 1990 by Mr Lakhum Mossang, of Namphai Nong, Miao, Arunachal Pradesh, for writing the Tangsa languages (included under ISO639:3 nst) spoken in Arunachal Pradesh, India and across the border in the north of Sagaing Region, Myanmar. It is a recently developed alphabetic script and is genetically unrelated to existing scripts. The script has been revised a couple of times. The 2020 January version consists of 89 characters: 79 letters (48 listed as vowels and 31 listed as consonants) and 10 digits. Tangsa is not a single language (see section 1.1 below) but a network of language varieties some of which are fully mutually intelligible and some of which are fully mutually unintelligible with many in between. The script developed by Lakhum Mossang is able to fully represent all of the sounds, and all phonemic distinctions, in his own variety, the Mossang. The word Mossang is an exonym (outsiders term) for a variety that is spelled in the Roman script as Muishvung[1]. The script is currently used by a number of native speakers of Muishvung (Mossang), several of whom have been learning it and using it for many years. As discussed further below, the Tangsa Script Development Committee, formed in 2019, is now actively promoting the script and its use in Government schools has been approved by the Arunachal Pradesh Pradesh Directorate of Elementary Education.

The name Tangsa was coined in the 1950s by Indian Government officials to cover a range of diverse tribes inhabiting what is now Changlang district of Arunachal Pradesh (including what is now the Tirap Transferred Area, now in Assam state). On the India side of the border there are around 40 'sub-tribes' of Tangsa, each of which has its own speech variety. Some of these varieties are fully mutually intelligible with each other (like the Cholim and Longri, and some are fully mutually unintelligible, like the Hahcheng and Champang). While the script is intended by its creator for all these language varieties, at present it is only in active use for one variety, Muishvung (Mossang).

The script being proposed for inclusion in Unicode includes of 48 symbols for vowels and sounds considered by the script's inventor to be vowel-like (including final -ŋ and syllabic nasal sounds). It also includes 31 consonants and 10 numerals which are fully decimal. The phonetic values (in inventor's own Mossang / Muishvng variety of Tangsa) are discussed in Section 7. A key design feature of the script is that for each vowel there are four symbols, corresponding to four different tones in Tangsa languages. These have been named TANGSA LETTER OZ, TANGSA LETTER OC, TANGSA LETTER OQ and TANGSA LETTER OX, using the Romanized orthography as developed by Rev Gam Win (discussed further below).

[1] Note that the different varieties of Tangsa have multiple names. The form Muishvung is an autonym, the name used by the people themselves, written in the Roman orthography developed by Rev. Gam Win. If written with tone marks it would be written as Muixshvungx. Until September 2020, the community were using the spelling Muishaung of Muixshaungx (with tone marks) but recent analysis of the sound system has decided that the letter v (standing for a short /a/ sound which will will notate as /ə/) should replace the letter a. It would be realised as [mɯ²ʃəuŋ²] in IPA where the superscript 2 stands for the tone category 2. Mossang is a 'general name' used by others to refer to the group. In this paper we will use autonym (general name) in that order. Each group has it's own name for each other group, so the Mossang are also called [mjɔ₂xaŋ₂] by the Cholim, [mjaŋ₂sa₂] by the Lauchäng, [mɯ₂ʃa₂] by the Shecyü̱ &c.

Traditionally there are no native names for the four different tones, however the Rev. Gam Win, who created a Roman based orthography in the late 1990s or early 2000s (see below Appendix 11 for more examples of this orthography, from his 2006 primer *Vphaung Wvnc Juingz Luik Dap Hewa Naga*), gave them names in Muishvng as follows and in the following order:

Symbol for tone in Rev. Gam Win's system	Name of Tone in Rev Gam Win's system	Meaning	Form in Muishvng	Number of Tone (Van Dam 2018, Morey 2015, 2017)
-Q	Thuic htaq	voice-break/cut	short, final glottal stop	TONE-4
-Z	Thuic nyenz	voice-soft	low falling	TONE-1
-X	Thuic hvlz	voice-middle	mid-high falling	TONE-2
-C	Thuic tsanz	voice-hard	mid-high level or rising	TONE-3

In various academic works that compare tones across a range of varieties, a numbering system has been used, as TONE-1, TONE-2, TONE-3 and TONE-4 for example Morey (2015, 2017) and van Dam (2018). TONE 4 corresponds to stop final syllables, whereas the other three are categories with open finals (vowels, nasals, and sometimes -l and -r). In Mossang, TONE 1 is low falling, TONE 2 is mid-high falling and TONE 3 is mid-high level or rising. But the form or realization of these tones differs from variety to variety. In Rera (Ronrang) for example, TONE 1 is high, TONE 2 is mid, TONE 3 is low.

The order in which Lakhum Mossang presented these vowels is the order -Z (TONE-1), -C (TONE-3), -Q (TONE-4) and -X (TONE 2).

As already mentioned, in this proposal, following the encoding used for the Liangshan Yi (https://www.unicode.org/charts/PDF/UA000.pdf) we will refer to the vowel characters as TANGSA LETTER OZ, TANGSA LETTER OC, TANGSA LETTER OQ, TANGSA LETTER OX &c, using the Rev. Gam Win orthography to refer to the letters

The script is a type of alphabet following the Daniels definition (https://en.wikipedia.org/wiki/Writing_system, (under Functional Classification) in the sense that symbols represent either consonants or vowels and each vowel symbol is written independently of consonants, following (to the right) of the consonant that commences a syllable. So the Muishvng word for 'to go' 𖩛𖩶 (TANGSA LETTER KA - U+16AA0) + TANGSA LETTER AZ – U+16A70), consists of the initial consonant /k/ and the vowel /a/ carrying TONE 1 (low tone in Muishvng). This word can be rendered in IPA as [ka¹].

There are some differences between this writing system and other alphabets. The first is that, as already discussed above, vowels and tones are merged into single symbols, so there are four symbols for each vowel, representing each of the four tones. The second is that there is one group of symbols, 𖩄 𖩅 𖩆 𖩇 (TANGSA LETTER UIUZ U+16A98 through to TANGSA LETTER UIUX, U+ 16A9B) which are actually diphthongs, the combination of two vowels together [ɯ] plus [u] , combined with the four tones. Thirdly, there is a different symbol the velar nasal in syllable initial position, where it is written with 𖩃 (TANGSA LETTER NGA, U+16AA3), but in syllable final position it is written with 𖩁 (TANGSA LETTER FINAL NG, U+16A90). The latter symbol is listed with the vowels. Thus the word for 'feel jealous' is written as 𖩃𖩷𖩁 (TANGSA LETTER NGA, U+16AA3, TANGSA LETTER UC, U+16A7F, and TANGSA LETTER FINAL NG, U+16A90) [ŋeŋ²]. By contrast, the other nasal sounds /m/ and /n/ can be written both at the beginning and at the end of a syllable, as with the final part of the word 𖩔𖩈𖩋𖩷𖩋 [kʰun²nen²] (a song language word meaning 'tendency of a person to desire others not to get a benefit'), where in the second syllable, 𖩋𖩷𖩋, TANGSA LETTER NA, U+16AAC is used in both initial and syllable final position.

The script has been taught in a handwritten form by Mr. Lakhum Mossang over the past 30 years. A small number (perhaps around 12) people have learned to use the script fluently, all of them members of the Muishvng (Mossang) community. An example of the kind of work produced by the users of the script is a handwritten text of a traditional song (examples, Figure 2 below).

Until around 2012, the script was only used in handwriting by Mr Lakhum Mossang and a small band of his devoted students, all of them speakers of Muishvng (Mossang). From around 2012, the development of a font began, firstly by a PhD student, Ms Karen Parker. This font that has subsequently been revised and overhauled and included in the 'Private Use' area of the Unicode by Dr. Kellen Parker van Dam, a former student of Stephen Morey now based at the University of Zürich. An earlier version of the font was used to produce the first printed document in the script, a document prepared by the Tangsa Script Development Committee to distribute at the Pangsau Pass festival in January 2020. Portions of this document, including both English language text and various Tangsa languages in the script, are given as Figures 3-6 below in the 'Examples' section.

As at April 2020, a primer to teach the script, at least for the Muishvng variety is being prepared. The draft version of the first two lessons of the Primer, which introduces the script, is shown in the Examples as Figure 8. The script has been submitted to the State Government of Arunachal Pradesh Directorate of Elementary Education. In August 2020, the members of the Script Development were asked by the Directory of Elementary Education to produce teaching materials, and while these are in draft form, a sample of the current draft is (23rd September 2020) is presented in Figure 10, this consists of the cover and first two pages.

Mr. Lakhum Mossang has devised the script with the intention that it can be used for all of the very diverse Tangsa varieties (See Morey 2015, 2017). The script covers almost all of the consonant sounds found in the various varieties. See Khämlan and Owen 2018 for a list of consonants recorded in a range of Tangsa (termed Tangshang in Myanmar) languages, There are vowels in some other varieties not yet included in the script (for example, a contrast between /e/ and /ɛ/ in the Rera (Ronrang) variety), and for this reason additional glyphs may be needed if the script is to be applied to all Tangsa / Tangshang languages. It is perfectly adequate as it stands for the Muishvng variety, and probably for many of the other Tangsa varieties.

The application of the script to other varieties would involve developing a convention for writing the tones using the existing symbols. As already mentioned, in Muishvng (Mossang), what we are calling TONE 1 is low falling, TONE 2 is mid-high falling and TONE 3 is mid-high level or rising and TONE 4 has a final stop (either glottal stop -ʔ or else -p, -t and -k. These tones areThe cognate tones in Rera (Ronrang) are as follows: TONE 1 high, TONE 2 mid and TONE 3 low. Since in most cases the same group of words carry TONE 1 in both varieties (though realised with a low tone in Muishvng and a high tone in Rera), a meeting held on January 27th at Namphai Nong village suggested that the TONE 1 symbols be used for this group of words in each variety, to be realised differently in each variety. This has not yet been brought into full application for any variety other than Muishvng.

Already, a revision was agreed in late 2019, to add the short /a/ like vowels (the series from U+16A78 to U+16A7B, named as TANGSA LETTER VZ, VC &c), as well as a glottal stop final symbol for the vowel pronounced in Muishvng (Mossang) as [ɔ] (U+16A8A, TANGSA LETTER AWQ) and for the letter [z] (U+16ABE, TANGSA LETTER Z.) These changes were approved following the adoption of the script by the newly set up Tangsa Script Development Committee which met on the 2nd November 2019 to appoint officers and a committee, details of which are outlines in the Examples, Figure 3 and Figure 4 below. The revised set of symbols promoted by that committee is included as Figure 7.

On January 27th 2020, a meeting was held at Namphai Nong village in Assam, attended by members of the script committee and in principle decisions were taken about how the tones of a variety other than Muishvng would be treated. This is discussed in section 9 below:

The script has been used in the preparation of a poster with the message of stopping the Corona Virus (see Appendix 9 below). This was posted on Facebook on 5th May 2020 (https://www.facebook.com/stephen.morey.92/posts/1551215425054338), using the poster prepared using the Translation Commons (https://translationcommons.org/covid-19multilingual/?edit=&official=&uid)

In May 2020, Wanglung Mossang and Stephen Morey have produced a short introductory video to the script, the first of a planned series. In July 2020, following the death of Lakhum Mossang on 11th July, the video was expanded to include an introduction to all the symbols in the current script. This video has been

uploaded to a dedicated YouTube Channel, at https://www.youtube.com/watch?v=2hefxhkGzsE. As at 24th August 2020, the channel has 65 subscribers (up from 6 in May) and five videos, including several with the late Lakhum Mossang introducing his script. Between them the 5 videos have had 1,699 views (up from 115 in May).

1.1 Some background: a note on 'Tangsa', 'Tangshang and 'Naga, Tase', versions of ISO 639-3: nst.

This section is included as background, to place the languages discussed here in the context the ISO 639-3 set. The situation of ISO 639-3 code nst is complex. It arose originally from the division of a range of diverse language varieties into two 'tribes': Tangsa and Nocte, done in India some 60 years ago and largely based on geographical rather than linguistic criteria. Nocte has been given the code ISO 639-3: njb. More recently, on the Myanmar side, both of these have been grouped together as one and called Tangshang. In the current versions of *Ethnologue*, the main entry for nst (https://www.ethnologue.com/language/nst) is headed 'Naga, Tangshang' and this subsumes what would be njb, (and we presume also Tutsa tvt and Wancho nnp) under a single heading on the Myanmar side. The *Ethnologue* entry for Naga, Tangsa in India includes fewer varieties, and does not subsume njb, tvt and nnp. The word *Tase* is the pronunciation of Tangsa in one variety of the language, the Chamchang or Kimsing.

The ISO reference (https://iso639-3.sil.org/code/nst) does not make it clear whether the current ISO coding includes njb, tvt or nnp, in other words follows the Myanmar entry for nst in *Ethnologue*, or whether it does not include them and follows the India entry for nst in *Ethnologue*. As mentioned earlier, these codes are originally based on post-Independence classifications in India that are primarily not linguistic. Whether the India side classification of Tangsa or the broader Myanmar side classification of Tangshang are used, these ISO codes include multiple language varieties, some of which are fully mutually intelligible and some of which are most defintely not mutually intelligible.

While the Script is designed to be dialect-agostic (variety-agnostic) at present it is fully in use for Muishvng.

2. Structure: The characters are all written left to right. Many of the words in the language are monosyllabic, and these syllables take the form of:

INITIAL CONSONANT (a very small number of words have initial vowels)

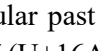
VOWEL + TONE

OPTIONAL SECOND VOWEL + TONE

OPTIONAL FINAL CONSONANT

For example, the word for 'sky', written as [rauŋ^2] phonemically, is given as 𖪲𖩷𖪇𖪐 (TANGSA LETTER RA (U+16AB2), TANGSA LETTER AX (U+16A77), TANGSA LETTER UX (U+16A87), TANGSA LETTER FINAL NG (U+16A90). (Note that a recent re-analysis of the phonology of Muishvung (Mossang) has changed this to [rəuŋ^2] phonemically, is given as 𖪲𖩻𖪇𖪐 (TANGSA LETTER RA (U+16AB2), TANGSA LETTER VX (U+16A7B), TANGSA LETTER UX (U+16A87), TANGSA LETTER FINAL NG (U+16A90).

When a vowel is followed by a final glottal stop, it is possible to write this in two ways. Consider the word /kaʔ/. This can be written as 𖪠𖩶 (TANGSA LETTER KA (U+16AA0), TANGSA LETTER AQ (U+16A76), or it can be written with an additional vowel symbol, as 𖪠𖩴𖩶 (TANGSA LETTER KA (U+16AA0), TANGSA LETTER AZ (U+16A74) TANGSA LETTER AQ (U+16A76). The latter is preferred by the founder of the script, the late Mr Lakhum Mossang, but some users feel that the former is sufficient since U+16A76 is equivalent to /-aʔ/

When diphthongs are written with any one of the the three open tones (ONE, TWO or THREE), the two vowel symbols are written carrying the same tone, as seen above with the word for 'sky'. However with words having final stops, it is usual to write the /a/ vowel as TONE 1 and the /u/ vowel as TONE 4, as in the example of /tauk/ (1st person singular past marker), which is written  (TANGSA LETTER TA (U+16AB0), TANGSA LETTER AZ (U+16A74) TANGSA LETTER UQ (U+16A86), TANGSA LETTER KA (U+16AA0).

3. Digits: The ten digits listed as U+16AC0 to U+16AC9. This is a fully decimal system that operates in the same way as the 'Arabic numerals'

4. Punctuation: There are no special punctuation marks and it is intended that the Roman punctuation marks be used if required. The question mark will not be used as there is a question particle as  (TANGSA LETTER HA (U+16AAd), TANGSA LETTER AAZ (U+16A74).

5. Word spacing: The Tangsa script employs spaces between words.

6. Variant Forms: No variant forms have been recorded

7. Character Naming: The suggested character names are descriptive of each character. Vowels are arranged in groups of four, so that the first four symbols are named as TANGSA LETTER OZ (U+16A70) [o¹], TANGSA LETTER OC (U+16A71) [o²], TANGSA LETTER OQ (U+16A72) [oʔ] and TANGSA LETTER OX (U+16A73) [o³]. The ordering of the vowels (-Z, -C, -Q and -X) is that used in Lakhum Mossang's original system. The vowel set UE (TANGSA LETTER UEC U+16A94 to TANGSA LETTER UEX 16A97) has a different order with the -C tone in front of the -Z tone, that is to say TANGSA LETTER UEC U+16A94 then TANGSA LETTER UEZ U+16A95. This also follows the original ordering as established by Lakhum Mossong.

The IPA symbols for these four vowels employ the vowel [o] in combination with the tones as they are realised in the Mossang (Muishvng) variety of Tangsa, where tone 1 (Low tone) [o¹] is low falling, TONE-2 (mid tone) [o²] is mid-high falling and tone 3 (high tone) [o³] is high and sometimes rising. In other Tangsa varieties, the tones are realised differently.

The full list of vowel names and their corresponding approximate phonetic values are given as follows:

TANGSA LETTER O = [o]

TANGSA LETTER A = [a]

TANGSA LETTER V = [ə]

TANGSA LETTER E = [e]

TANGSA LETTER I = [i]

TANGSA LETTER U = [u]

TANGSA LETTER AW = [ɔ]

TANGSA LETTER UI = [ɯ]

TANGSA LETTER UE = [ɤ]

TANGSA LETTER UIU = [ɯu]

The group of sounds listed as TANGSA LETTER SYLLABIC M can be phonetically represented as [m̩]

The consonants are listed after the vowels, and named as TANGSA LETTER KA, TANGSA LETTER KHA &c.

The names of the letters employ the Roman based orthography that was developed by Rev. Gam Win for Muishvng variety. Most of the names are transparent, so that TANGSA LETTER KA refers to [k] and TANGSA LETTER KHA refers to [kha]. However the following consonants have phonetic forms in Muishvng that may not be transparent from the Rev Gam Win system. These are given together with their phonetic equivalent In one case, the voiceless [tɕ] sound, we are suggesting the name TANGSA LETTER CA, rather than using Rev. Gam Win's orthographic <j>.

Number	Symbol	Name of letter	Phonetic equivalent	Notes
16AAF	𖩯	TANGSA LETTER HTA	[tʰ]	
16AB3	𖩳	TANGSA LETTER NHA	[n̪]	dental nasal
16AB5	𖩵	TANGSA LETTER CA	[tɕ]	written with <j> in Gam Win's system
16AB7	𖩷	TANGSA LETTER GHA	[ɣ]	
16AB8	𖩸	TANGSA LETTER HTTA	[ʈʰ]	aspirated voiceless dental stop
16AB9	𖩹	TANGSA LETTER THA	[ʈ]	unaspirated voiceless dental stop
16ABC	𖩼	TANGSA LETTER DHA	[ð]	a voiced dental fricative
16ABD	𖩽	TANGSA LETTER CHA	[tɕʰ]	aspirated

The numerals are listed last after the vowels and the consonants.

8. Sort order: The order is based on the original order as developed by Lakhum Mossang and further amended in January 2020. The only change to this is one required by Unicode. Whereas Lakhum Mossang listed the numerals with TANGSA DIGIT ONE, U=16AC1 first and TANGSA DIGIT ZERO, U=16AC0, Unicode rules require the listing of ZERO first.

9. Issues:

There are a number of issues relating to the script that need to be pointed out

1) There is a special symbol for final /-ŋ/ in the vowel series TANGSA LETTER FINAL NG (U+16A90) rather than using the consonant TANGSA LETTER NGA (U+16AA3). For the other final nasals, /-m/ and /-n/ there is no such special symbol.

2) Four consonant symbols for consonants not used in Muishvng (Mossang) were created using other consonant symbols with 'combining marks' that are like diacritics written above symbols for other LETTERS. These are listed below:

16ABB	𖩻	TANGSA LETTER FA
16ABC	𖩼	TANGSA LETTER DHA

16ABD	𖪽	TANGSA LETTER CHA
16ABE	𖪾	TANGSA LETTER ZA

For example, the TANGSA LETTER FA is made up of th symbol U+16AA9, TANGSA LETTER PHA with a small symbol above it. At this time these 'diacritics' are not proposed for encoding as separate entities as they are not currently productive; however if further consonants need to be added to the script in order to write varieties other than Muishvng, they may become productive.

4) Two symbols that are largely prosodic in nature, and one is used for a toneless prefix. Together with the TANGSA LETTER UE series (U+16A91 to U+16A94), they are termed by Lakhum Mossang as the 'seven sisters', a reference to a well known term for the seven states of Northeast India. As far as we know, these symbols are not used for any words in citation form.

16A91 𖪑 TANGSA LETTER LONG UEX
16A92 𖪒 TANGSA LETTER SHORT UEZ
16A93 𖪓 TANGSA LETTER SHORT AWX

The two prosodic symbols, 16A91 and 16A93 are used as follows. A long falling tone is used in the phrase 𖪬𖪑 [illegible] [nɤː² kəra²] ,'**very very** far' where the first word uses TANGSA LETTER NA, U+16AAC and TANGSA LETTER LONG UEX, U+16A91 rather than being 𖪬𖪗 [nɤ²], a distal deictic ('far') which is TANGSA LETTER NA, U+16AAC and TANGSA LETTER UEX, U+16A97.

The short AW [ɔ] sound is used in the phrase [illegible] 𖪮𖪓 [vrɤ² lɔ²ˀ], a phrase meaning 'let it be only so much' with a short final [ɔ²] vowel (TONE 2, mid falling). The second syllable of this phrase 𖪮𖪓 [lɔ²ˀ], an imperative particle, is written with TANGSA LETTER LA, U+16AAE, and TANGSA LETTER SHORT AWX, U+16A93.

In the Muishvng language there are two quite distinct sounds that are written in Rev. Gam Win's Roman based orthography with the letter v. On the one hand there are tone marked vowels that occur in syllables like 𖪭𖩻𖪮 (TANGSA LETTER HA, U+16AAD, TANGSA LETTER VX, U+16A7B, TANGSA LETTER LA, U+16AAE)[həl²] 'good'. There are also prefixes, consisting of an initial consonant followed by a short tonless vowel, which in the Rev. Gam's system are written as tv-, gv-, shv- &c. Recent analysis shows that the vowel in these prefixes is a phonetically different sounds from the vowels in the TANGSA LETTER V set (U+16A78 to 16A7B), and the symbol U+16A92 (𖪒) (TANGSA LETTER SHORT UEZ), has been adopted for this, as shown in the following examples. Note that this distinction is not marked in the Rev Gam Win's romanized spelling. For an initial vowel prefix (written v- in the Gam Win system), the TANGSA LETTER VZ (U+16A78) is used.

English	Gam Win Spelling	IPA transcription	Tangsa
blood	tvghuiyz	təɣɯi¹	[illegible]
creator	shvkex	ʃəke²	[illegible]
likewise	kvrulc mvrenc	kərul³ məren³	[illegible]
others	wvghanc	βəɣan³	[illegible]
lastly	vtvsuip	ətəsɯp	[illegible]

This symbol can be used for a sound that is phonetically different from the short A sound that is written by U+16A78 through to U+16A7B

4) There are three syllabic nasals that are largely used in Muishvng (Mossang) as exclamations and confirmation particles (yes). Note that the sound made by the syllabic nasal tends to be a syllable [m] sound, hence the name used in this proposal.

16A9d 𖪝 TANGSA LETTER MC
16A9e 𖪞 TANGSA LETTER MQ
16A9f 𖪟 TANGSA LETTER MX

(5) Potential additional characters

In some other Tangsa / Tangshang varieties there is a vowel that could be written as [y] or [ø] (see Khämlan Binkhäm and Owen 2018: 23). It will be necessary to add a series of four symbols to cover this set of vowels

Recent work by Deepjyoti Goswami on the Rera (Ronrang) variety strongly suggests that there is a distinction between /e/ and /ɛ/. This distinction is not mentioned in Khämlan Binkhäm and Owen (2018: 23), but to write Rera using the script, a further set of four symbols would be needed.

Consonant phonemes discussed by Khämlan Binkhäm and Owen (2018: 23) that are not so far included in the script would be [tsʰ], [j] (where it is a distinct phoneme from [ʤ] / [ʣ] and possibly [ɬ]. Note that this latter is included by Khämlan Binkhäm and Owen (2018: 23) because the variety in which it is found is included under Tangshang in Myanmar. But that variety (sometimes termed Chuyo) would likely be listed as a Wancho variety in India.

Another possibly additional symbol would be a toneless prefixal [ɪ] that is found in varieties like Cholim (Tonglum)

6) Application of the script to varieties other than Muishvng (Mossang)

As mentioned earlier, the vowel symbols combine vowel and tone, but the tones of other Tangsa varieties are different from Muishvng (Mossang). At a meeting on January 27th, it was agreed that the vowel symbols would be used according to the tone categories.

Thus, in Rera, the group of words which belong to TONE 1 (-Z series) are realised with a high tone, while that same group of words in Muishvng (Mossang) are realised with a low tone (see van Dam 2018). It was agreed that these words would be written with the TONE 1 group of symbols, and interpreted differently in the different varieties (i.e. pronounced with a low tone in Muishvng and a high tone in Rera). Annotations to the names list will be provided to indicate this.

This is similar to the use of <ch> in Roman script which in the word *chat* is pronounced [ʧ] in English and [ʃ] in French

Note that there are some Tangsa varieties, like Champang, where tone has such a functional load that it is not clear that tones are still a phonemic feature of the language. How a toneless language would be written using this script is not clear.

7) Use of the script.

Several communities in India (representing differentsub-tribes) have expressed that they wish to use a Roman-based orthography in conjunction with the Tangsa script for the time being, as they begin the development of literature. It should be noted that on the Myanmar side a number of Roman based orthographies are currently being promoted by different sub-tribes. One of these orthographies is the Tangshang Naga Unified Orthography (Khämlan Binkhäm and Owen), but there are others which are not currently documented in published sources.

	16A7	16A8	16A9	16AA	16AB	16AC
0						
1						
2						
3						
4						
5						
6						
7						
8						
9						
A						
B						
C						
D						
E						
F						

Vowels

16A70 TANGSA LETTER OZ
16A71 TANGSA LETTER OC
16A72 TANGSA LETTER OQ
16A73 TANGSA LETTER OX
16A74 TANGSA LETTER AZ
16A75 TANGSA LETTER AC
16A76 TANGSA LETTER AQ
16A77 TANGSA LETTER AX
16A78 TANGSA LETTER VZ
16A79 TANGSA LETTER VC
16A7A TANGSA LETTER VQ
16A7B TANGSA LETTER VX
16A7C TANGSA LETTER EZ
16A7D TANGSA LETTER EC
16A7E TANGSA LETTER EQ
16A7F TANGSA LETTER EX
16A80 TANGSA LETTER IZ
16A81 TANGSA LETTER IC
16A82 TANGSA LETTER IQ
16A83 TANGSA LETTER IX
16A84 TANGSA LETTER UZ
16A85 TANGSA LETTER UC
16A86 TANGSA LETTER UQ
16A87 TANGSA LETTER UX
16A88 TANGSA LETTER AWZ
16A89 TANGSA LETTER AWC
16A8A TANGSA LETTER AWQ
16A8B TANGSA LETTER AWX
16A8C TANGSA LETTER UIZ
16A8D TANGSA LETTER UIC
16A8E TANGSA LETTER UIQ
16A8F TANGSA LETTER UIX
16A90 TANGSA LETTER FINAL NG
16A91 TANGSA LETTER LONG UEX
16A92 TANGSA LETTER SHORT UEZ
16A93 TANGSA LETTER SHORT AWX
16A94 TANGSA LETTER UEC
16A95 TANGSA LETTER UEZ
16A96 TANGSA LETTER UEQ
16A97 TANGSA LETTER UEX
16A98 TANGSA LETTER UIUZ
16A99 TANGSA LETTER UIUC
16A9A TANGSA LETTER UIUQ
16A9B TANGSA LETTER UIUX
16A9C TANGSA LETTER MZ
16A9D TANGSA LETTER MC
16A9E TANGSA LETTER MQ
16A9F TANGSA LETTER MX

Consonants

16AA0 TANGSA LETTER KA
16AA1 TANGSA LETTER KHA
16AA2 TANGSA LETTER GA
16AA3 TANGSA LETTER NGA
16AA4 TANGSA LETTER SA
16AA5 TANGSA LETTER YA
16AA6 TANGSA LETTER WA
16AA7 TANGSA LETTER PA
16AA8 TANGSA LETTER NYA
16AA9 TANGSA LETTER PHA
16AAA TANGSA LETTER BA
16AAB TANGSA LETTER MA
16AAC TANGSA LETTER NA
16AAD TANGSA LETTER HA
16AAE TANGSA LETTER LA
16AAF TANGSA LETTER HTA
16AB0 TANGSA LETTER TA
16AB1 TANGSA LETTER DA
16AB2 TANGSA LETTER RA
16AB3 TANGSA LETTER NHA
16AB4 TANGSA LETTER SHA
16AB5 TANGSA LETTER CA
16AB6 TANGSA LETTER TSA
16AB7 TANGSA LETTER GHA
16AB8 TANGSA LETTER HTTA
16AB9 TANGSA LETTER THA
16ABA TANGSA LETTER XA
16ABB TANGSA LETTER FA
16ABC TANGSA LETTER DHA
16ABD TANGSA LETTER CHA

16ABE 𖫾 TANGSA LETTER ZA

Digits

16AC0 𖬀 TANGSA DIGIT ZERO
16AC1 𖬁 TANGSA DIGIT ONE
16AC2 𖬂 TANGSA DIGIT TWO
16AC3 𖬃 TANGSA DIGIT THREE
16AC4 𖬄 TANGSA DIGIT FOUR
16AC5 𖬅 TANGSA DIGIT FIVE
16AC6 𖬆 TANGSA DIGIT SIX
16AC7 𖬇 TANGSA DIGIT SEVEN
16AC8 𖬈 TANGSA DIGIT EIGHT
16AC9 𖬉 TANGSA DIGIT NINE

Unicode Properties

16A70;TANGSA LETTER OZ;Lo;0;L;;;;;N;;;;;
16071;TANGSA LETTER OC;Lo;0;L;;;;;N;;;;;
etc.

16AC0;TANGSA DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
16AC1;TANGSA DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
16AC2;TANGSA DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
16AC3;TANGSA DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
16AC4;TANGSA DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
16AC5;TANGSA DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
16AC6;TANGSA DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
16AC7;TANGSA DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
16AC8;TANGSA DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
16AC9;TANGSA DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;

Examples

(Note: As far as we know the only printed document in the script was part of the document produced by the Script Committee in 2020, samples of which appear in Figures 3-6)

Figure 1: Full list of Lakhum Mossang's script (2003 version)

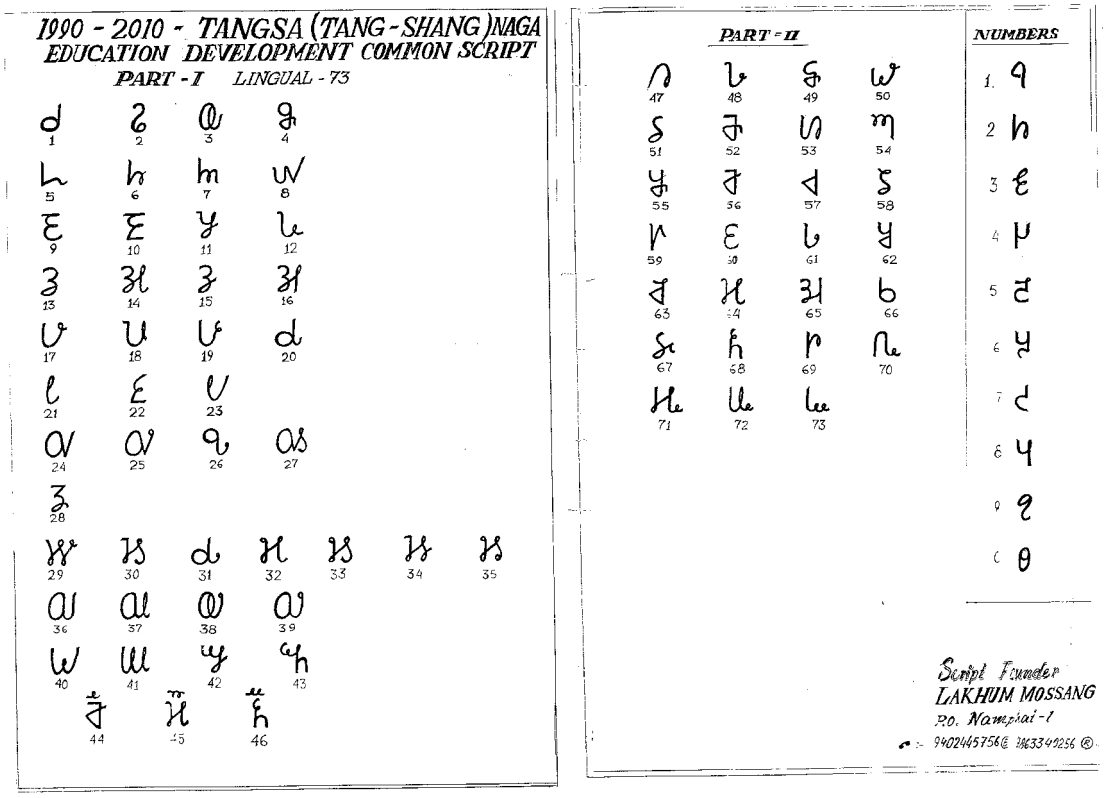


Figure 2: Wihu Song Manuscript written by Lakhum Mossang

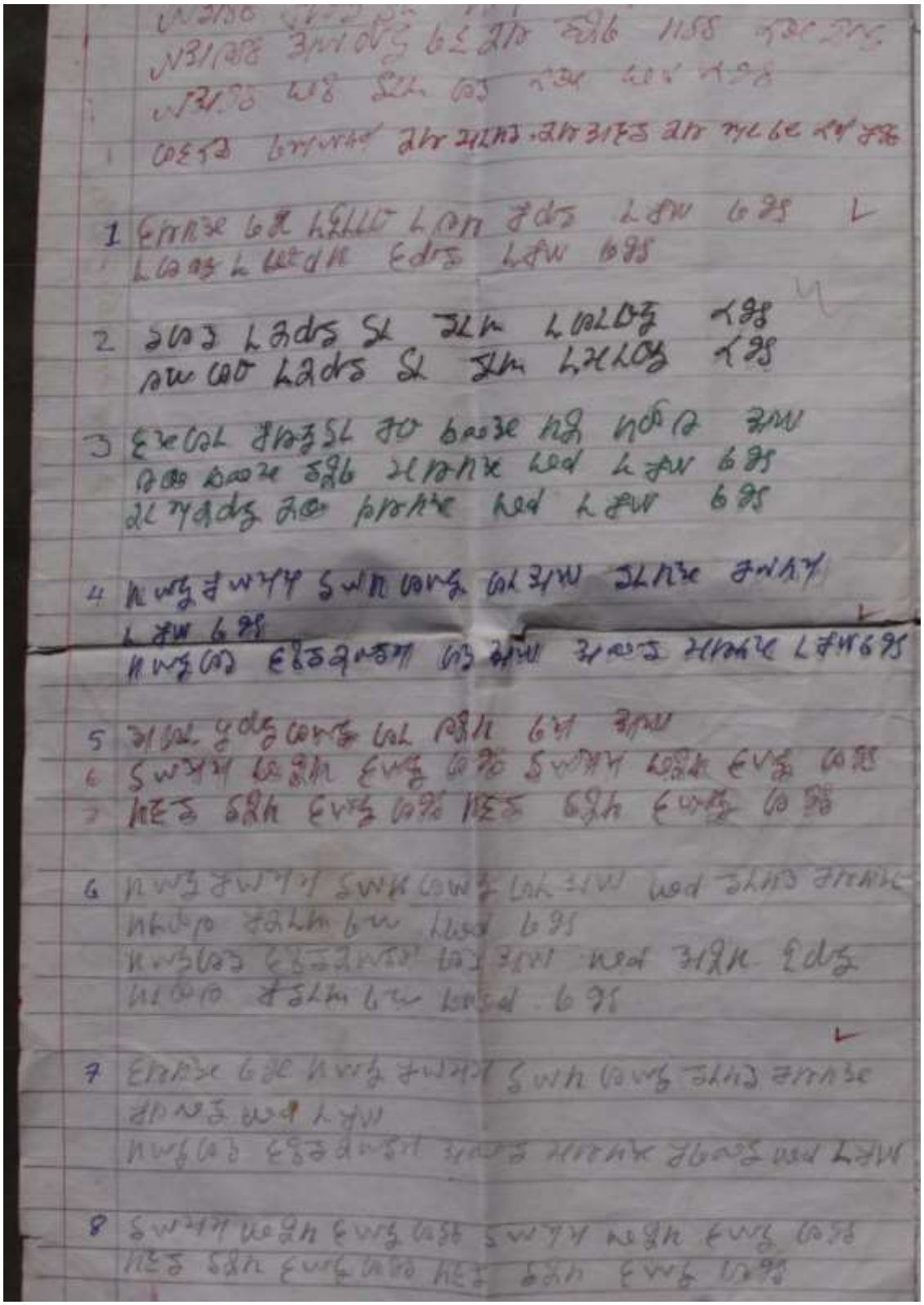
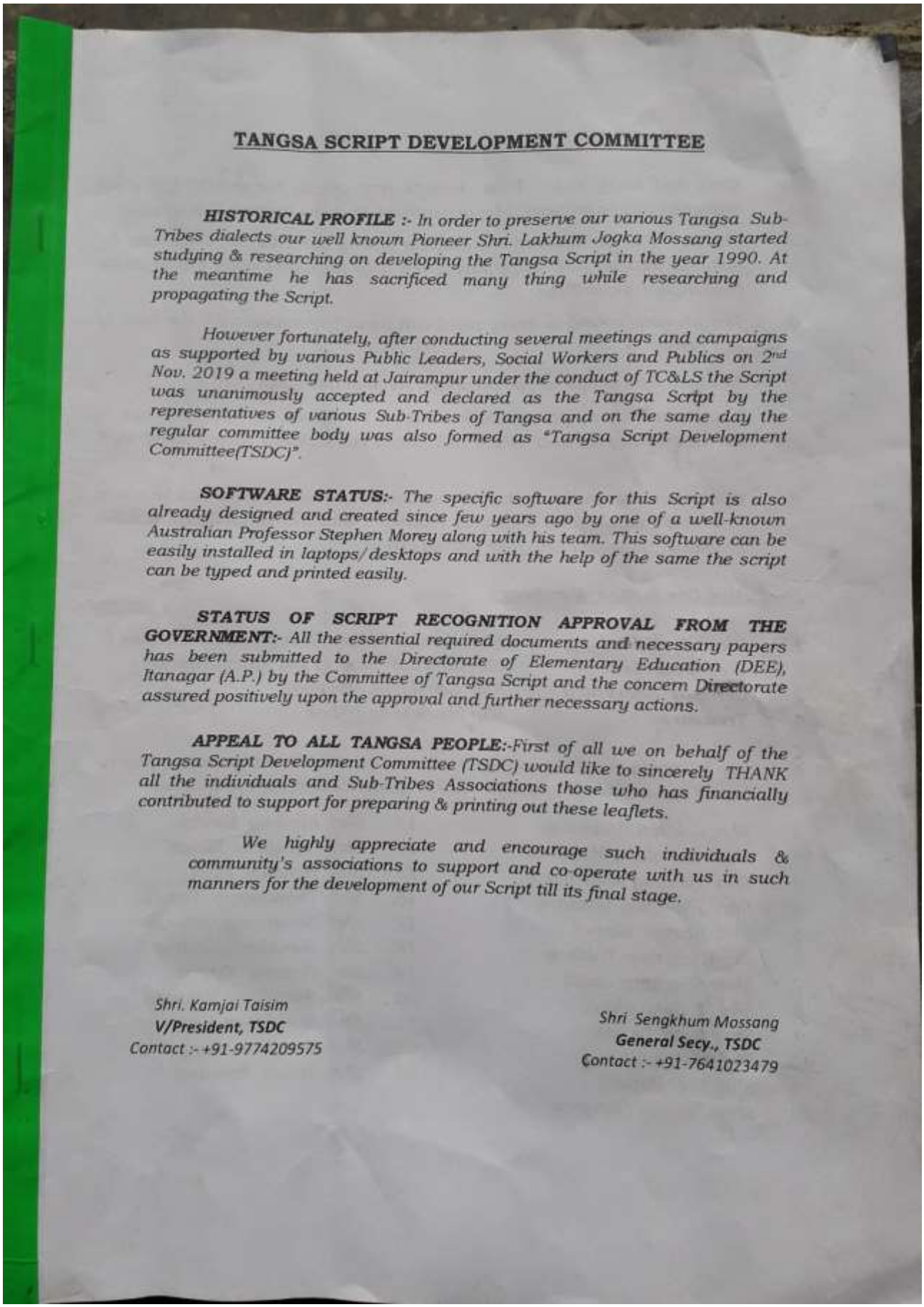


Figure 3: Page 1, document produced by Tangsa Script Development Committee, January 2020



TANGSA SCRIPT DEVELOPMENT COMMITTEE

HISTORICAL PROFILE *:- In order to preserve our various Tangsa Sub-Tribes dialects our well known Pioneer Shri. Lakhum Jogka Mossang started studying & researching on developing the Tangsa Script in the year 1990. At the meantime he has sacrificed many thing while researching and propagating the Script.*

However fortunately, after conducting several meetings and campaigns as supported by various Public Leaders, Social Workers and Publics on 2nd Nov. 2019 a meeting held at Jairampur under the conduct of TC&LS the Script was unanimously accepted and declared as the Tangsa Script by the representatives of various Sub-Tribes of Tangsa and on the same day the regular committee body was also formed as "Tangsa Script Development Committee(TSDC)".

SOFTWARE STATUS:- *The specific software for this Script is also already designed and created since few years ago by one of a well-known Australian Professor Stephen Morey along with his team. This software can be easily installed in laptops/desktops and with the help of the same the script can be typed and printed easily.*

STATUS OF SCRIPT RECOGNITION APPROVAL FROM THE GOVERNMENT:- *All the essential required documents and necessary papers has been submitted to the Directorate of Elementary Education (DEE), Itanagar (A.P.) by the Committee of Tangsa Script and the concern Directorate assured positively upon the approval and further necessary actions.*

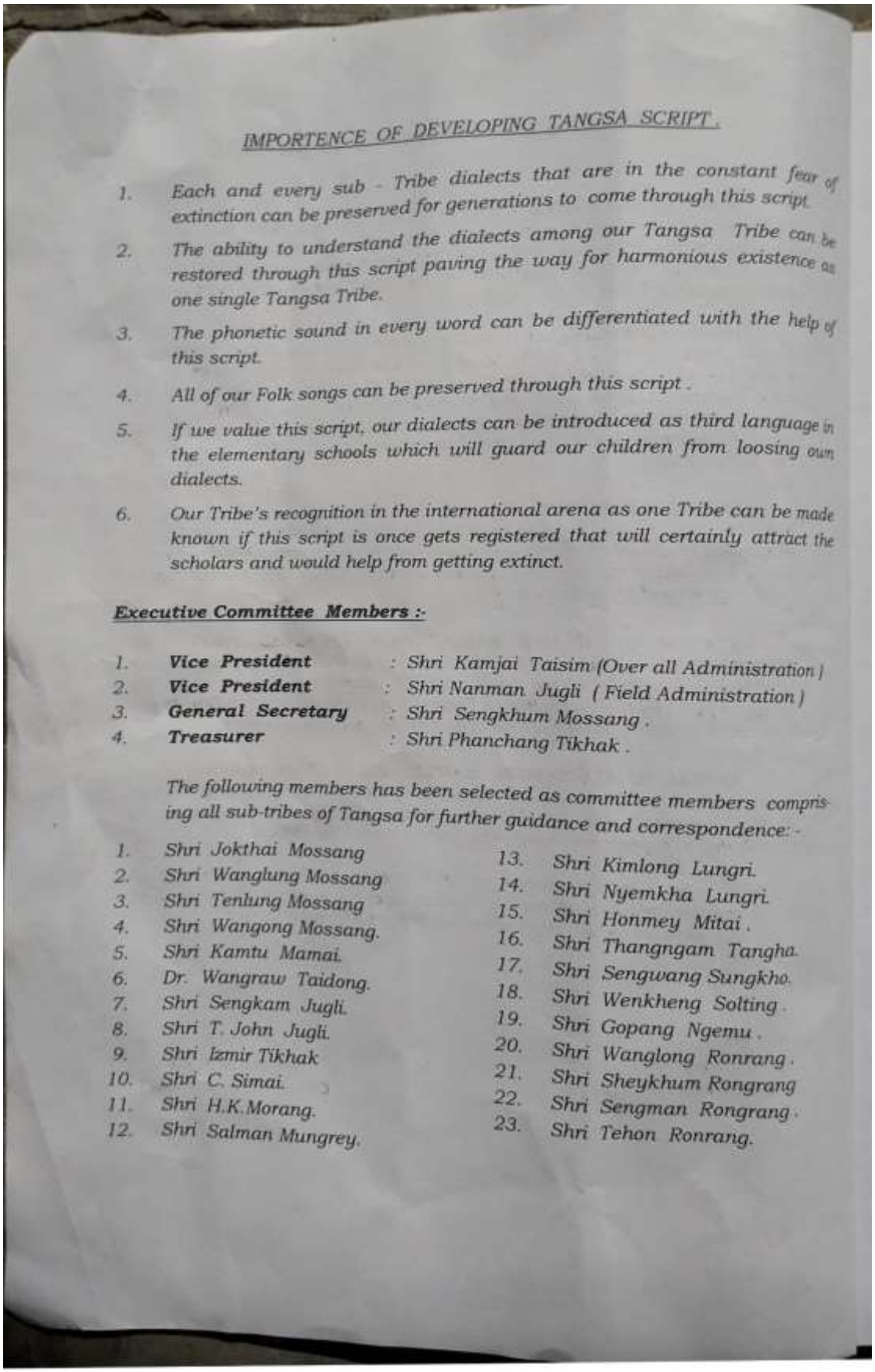
APPEAL TO ALL TANGSA PEOPLE:- *First of all we on behalf of the Tangsa Script Development Committee (TSDC) would like to sincerely THANK all the individuals and Sub-Tribes Associations those who has financially contributed to support for preparing & printing out these leaflets.*

We highly appreciate and encourage such individuals & community's associations to support and co-operate with us in such manners for the development of our Script till its final stage.

Shri. Kamjai Taisim
V/President, TSDC
Contact :- +91-9774209575

Shri Sengkhum Mossang
General Secy., TSDC
Contact :- +91-7641023479

Figure 4: Page 2, document produced by Tangsa Script Development Committee, January 2020



IMPORTENCE OF DEVELOPING TANGSA SCRIPT.

1. Each and every sub - Tribe dialects that are in the constant fear of extinction can be preserved for generations to come through this script.
2. The ability to understand the dialects among our Tangsa Tribe can be restored through this script paving the way for harmonious existence as one single Tangsa Tribe.
3. The phonetic sound in every word can be differentiated with the help of this script.
4. All of our Folk songs can be preserved through this script.
5. If we value this script, our dialects can be introduced as third language in the elementary schools which will guard our children from loosing own dialects.
6. Our Tribe's recognition in the international arena as one Tribe can be made known if this script is once gets registered that will certainly attract the scholars and would help from getting extinct.

Executive Committee Members :-

1. **Vice President** : Shri Kamjai Taisim (Over all Administration)
2. **Vice President** : Shri Nanman Jugli (Field Administration)
3. **General Secretary** : Shri Sengkhum Mossang.
4. **Treasurer** : Shri Phanchang Tikhak.

The following members has been selected as committee members comprising all sub-tribes of Tangsa for further guidance and correspondence: -

1. Shri Jokthai Mossang
2. Shri Wanglung Mossang
3. Shri Tenlung Mossang
4. Shri Wangong Mossang.
5. Shri Kamtu Mamai.
6. Dr. Wangraw Taidong.
7. Shri Sengkam Jugli.
8. Shri T. John Jugli.
9. Shri Izmir Tikhak
10. Shri C. Simai.
11. Shri H.K.Morang.
12. Shri Salman Mungrey.
13. Shri Kimlong Lungri.
14. Shri Nyemkha Lungri.
15. Shri Honmey Mitai.
16. Shri Thangngam Tangha.
17. Shri Sengwang Sungkho.
18. Shri Wenkheng Solting.
19. Shri Gopang Ngemu.
20. Shri Wanglong Ronrang.
21. Shri Sheykhum Rongrang
22. Shri Sengman Rongrang.
23. Shri Tehon Ronrang.

Figure 5: Page 3, document produced by Tangsa Script Development Committee, January 2020

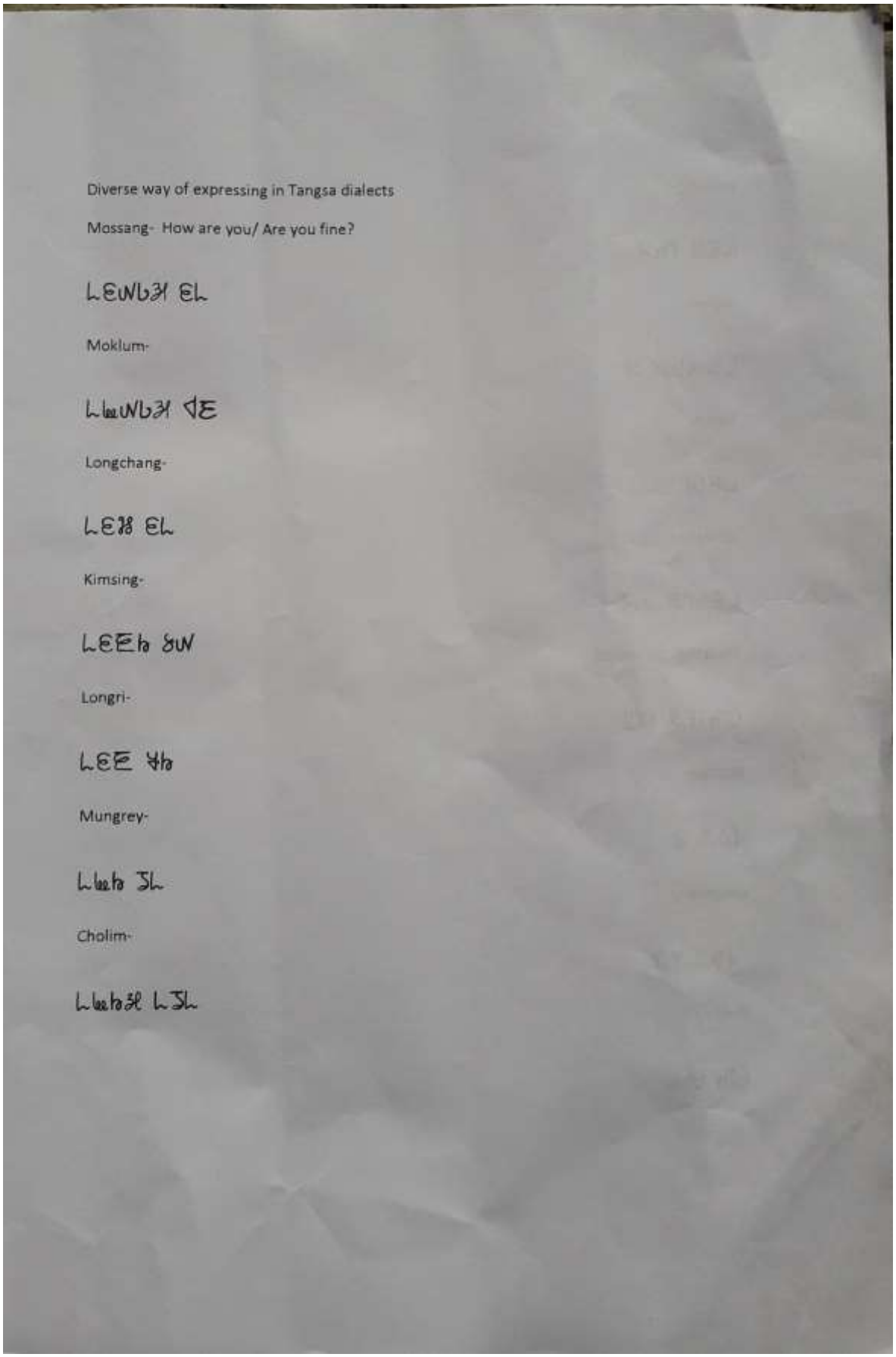


Figure 6: Page 9, document produced by Tangsa Script Development Committee, January 2020

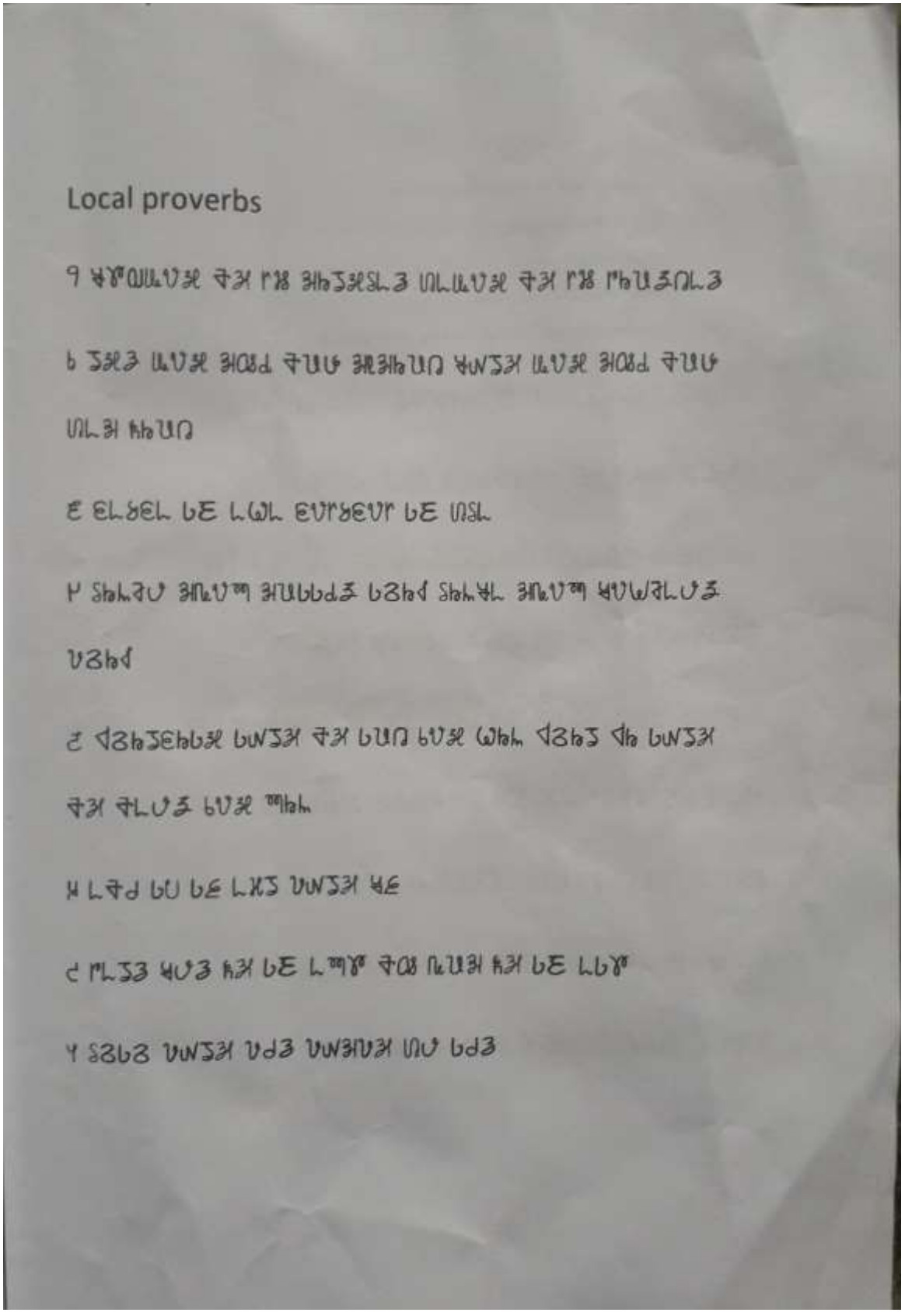


Figure 7: Revised list of Lakhum Mossang Characters January 2020

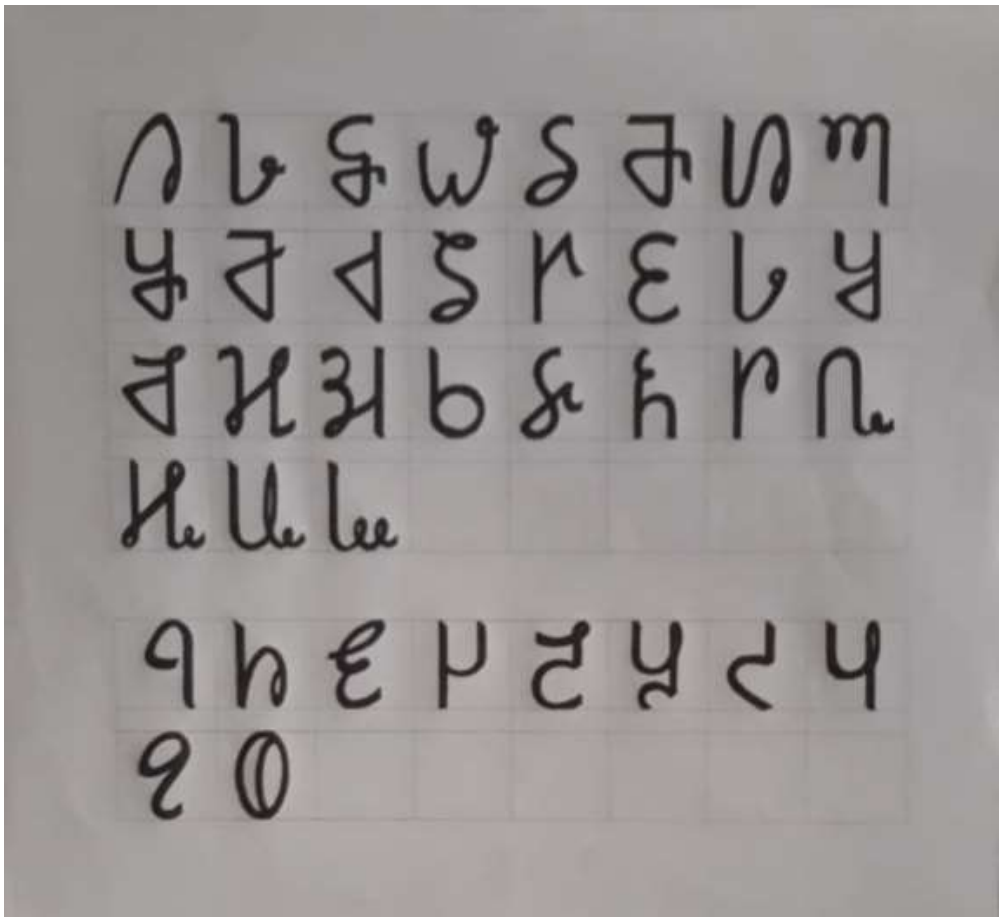


Figure 8: Samples of first two 'lessons' of the draft Primer, April 2020

D

𖪰𖪔𖪮𖪍𖪐 𖫁 𖫁

(Lesson 1)

𖪮𖪌𖪢𖪠𖪛 (Vowels)

𖩰 oz	𖩱 oc	𖩲 oq	𖩳 ox	𖩴 az	𖩵 ac	𖩶 aq	𖩷 ax
𖩸 vz	𖩹 vc	𖩺 vq	𖩻 vx	𖩼 ez	𖩽 ec	𖩾 eq	𖩿 ex
𖪀 iz	𖪁 ic	𖪂 iq	𖪃 ix	𖪄 uz	𖪅 uc	𖪆 uq	𖪇 ux
𖪈 uiz	𖪉 uic	𖪊 uiq	𖪋 uix	𖪌 awz	𖪍 awc	𖪎 awq	𖪏 awx
𖪐 eng	𖪑 uex	𖪒 uezz	𖪓 awx	𖪔 uec	𖪕 uec	𖪖 ueq	𖪗 uex
𖪘 uiuz	𖪙 uiuc	𖪚 uiuq	𖪛 uiux	𖪜 mz	𖪝 mc	𖪞 mq	𖪟 mx

𖪮𖪌𖪢𖪮𖩴 (Consonants)

𖪠 k	𖪡 kh	𖪢 g	𖪣 ng	𖪤 s	𖪥 y	𖪦 w	𖪧 p
𖪨 ny	𖪩 ph	𖪪 b	𖪫 m	𖪬 n	𖪭 h	𖪮 l	𖪯 ht
𖪰 t	𖪱 d	𖪲 r	𖪳 nh	𖪴 sh	𖪵 j	𖪶 ts	𖪷 gh
𖪸 htt	𖪹 th	𖪺 x	𖪻 f	𖪼 dh	𖪽 ch	𖪾 z	

𖪦𖩵𖪮𖪦𖩵𖪮 (Numerals)

𖫁	𖫂	𖫃	𖫄	𖫅	𖫆	𖫇	𖫈	𖫉	𖫁𖫀
𖫁𖫁	𖫁𖫂	𖫁𖫃	𖫁𖫄	𖫁𖫅	𖫁𖫆	𖫁𖫇	𖫁𖫈	𖫁𖫉	𖫂𖫀
𖫂𖫁	𖫂𖫂	𖫂𖫃	𖫂𖫄	𖫂𖫅	𖫂𖫆	𖫂𖫇	𖫂𖫈	𖫂𖫉	𖫃𖫀
𖫃𖫁	𖫃𖫂	𖫃𖫃	𖫃𖫄	𖫃𖫅	𖫃𖫆	𖫃𖫇	𖫃𖫈	𖫃𖫉	𖫄𖫀
𖫄𖫁	𖫄𖫂	𖫄𖫃	𖫄𖫄	𖫄𖫅	𖫄𖫆	𖫄𖫇	𖫄𖫈	𖫄𖫉	𖫅𖫀
𖫅𖫁	𖫅𖫂	𖫅𖫃	𖫅𖫄	𖫅𖫅	𖫅𖫆	𖫅𖫇	𖫅𖫈	𖫅𖫉	𖫆𖫀
𖫆𖫁	𖫆𖫂	𖫆𖫃	𖫆𖫄	𖫆𖫅	𖫆𖫆	𖫆𖫇	𖫆𖫈	𖫆𖫉	𖫇𖫀
𖫇𖫁	𖫇𖫂	𖫇𖫃	𖫇𖫄	𖫇𖫅	𖫇𖫆	𖫇𖫇	𖫇𖫈	𖫇𖫉	𖫈𖫀
𖫈𖫁	𖫈𖫂	𖫈𖫃	𖫈𖫄	𖫈𖫅	𖫈𖫆	𖫈𖫇	𖫈𖫈	𖫈𖫉	𖫉𖫀
𖫉𖫁	𖫉𖫂	𖫉𖫃	𖫉𖫄	𖫉𖫅	𖫉𖫆	𖫉𖫇	𖫉𖫈	𖫉𖫉	𖫁𖫀𖫀

(Lesson 2)

ಕ ಕಾ ಕಿ ಕೀ ಕು ಕೂ ಕೃ ಕೆ ಕೇ ಕೈ ಕೊ ಕೋ ಕೌ

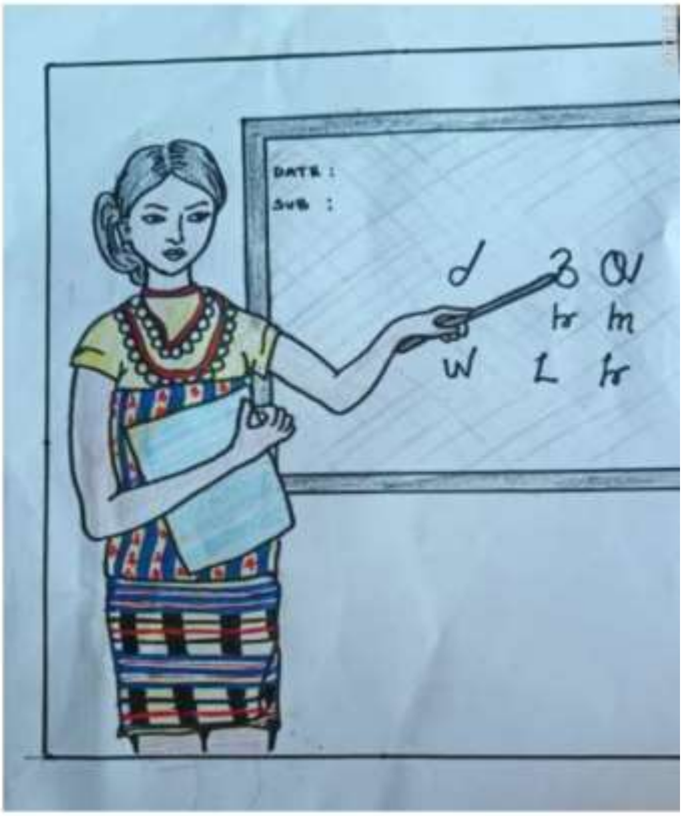
ಕ್ಕ ಕ್ಕಾ ಕ್ಕಿ ಕ್ಕೀ	ಕ್ಕು ಕ್ಕೂ ಕ್ಕೃ ಕ್ಕೆ	ಕ್ಕೇ ಕ್ಕೈ ಕ್ಕೊ ಕ್ಕೋ	ಕ್ಕೌ ಕ್ಕಂ ಕ್ಕಃ	ಕ್ಕ್ ಕ್ಕಾ ಕ್ಕಿ ಕ್ಕೀ
ಖ್ಖ ಖ್ಖಾ ಖ್ಖಿ ಖ್ಖೀ	ಖ್ಖು ಖ್ಖೂ ಖ್ಖೃ ಖ್ಖೆ	ಖ್ಖೇ ಖ್ಖೈ ಖ್ಖೊ ಖ್ಖೋ	ಖ್ಖೌ ಖ್ಖಂ ಖ್ಖಃ	ಖ್ಖ್ ಖ್ಖಾ ಖ್ಖಿ ಖ್ಖೀ
ಗ್ಗ ಗ್ಗಾ ಗ್ಗಿ ಗ್ಗೀ	ಗ್ಗು ಗ್ಗೂ ಗ್ಗೃ ಗ್ಗೆ	ಗ್ಗೇ ಗ್ಗೈ ಗ್ಗೊ ಗ್ಗೋ	ಗ್ಗೌ ಗ್ಗಂ ಗ್ಗಃ	ಗ್ಗ್ ಗ್ಗಾ ಗ್ಗಿ ಗ್ಗೀ
ಘ್ಘ ಘ್ಘಾ ಘ್ಘಿ ಘ್ಘೀ	ಘ್ಘು ಘ್ಘೂ ಘ್ಘೃ ಘ್ಘೆ	ಘ್ಘೇ ಘ್ಘೈ ಘ್ಘೊ ಘ್ಘೋ	ಘ್ಘೌ ಘ್ಘಂ ಘ್ಘಃ	ಘ್ಘ್ ಘ್ಘಾ ಘ್ಘಿ ಘ್ಘೀ
ಙ್ಙ ಙ್ಙಾ ಙ್ಙಿ ಙ್ಙೀ	ಙ್ಙು ಙ್ಙೂ ಙ್ಙೃ ಙ್ಙೆ	ಙ್ಙೇ ಙ್ಙೈ ಙ್ಙೊ ಙ್ಙೋ	ಙ್ಙೌ ಙ್ಙಂ ಙ್ಙಃ	ಙ್ಙ್ ಙ್ಙಾ ಙ್ಙಿ ಙ್ಙೀ
ಚ್ಚ ಚ್ಚಾ ಚ್ಚಿ ಚ್ಚೀ	ಚ್ಚು ಚ್ಚೂ ಚ್ಚೃ ಚ್ಚೆ	ಚ್ಚೇ ಚ್ಚೈ ಚ್ಚೊ ಚ್ಚೋ	ಚ್ಚೌ ಚ್ಚಂ ಚ್ಚಃ	ಚ್ಚ್ ಚ್ಚಾ ಚ್ಚಿ ಚ್ಚೀ
ಛ್ಛ ಛ್ಛಾ ಛ್ಛಿ ಛ್ಛೀ	ಛ್ಛು ಛ್ಛೂ ಛ್ಛೃ ಛ್ಛೆ	ಛ್ಛೇ ಛ್ಛೈ ಛ್ಛೊ ಛ್ಛೋ	ಛ್ಛೌ ಛ್ಛಂ ಛ್ಛಃ	ಛ್ಛ್ ಛ್ಛಾ ಛ್ಛಿ ಛ್ಛೀ
ಜ್ಜ ಜ್ಜಾ ಜ್ಜಿ ಜ್ಜೀ	ಜ್ಜು ಜ್ಜೂ ಜ್ಜೃ ಜ್ಜೆ	ಜ್ಜೇ ಜ್ಜೈ ಜ್ಜೊ ಜ್ಜೋ	ಜ್ಜೌ ಜ್ಜಂ ಜ್ಜಃ	ಜ್ಜ್ ಜ್ಜಾ ಜ್ಜಿ ಜ್ಜೀ
ಝ್ಝ ಝ್ಝಾ ಝ್ಝಿ ಝ್ಝೀ	ಝ್ಝು ಝ್ಝೂ ಝ್ಝೃ ಝ್ಝೆ	ಝ್ಝೇ ಝ್ಝೈ ಝ್ಝೊ ಝ್ಝೋ	ಝ್ಝೌ ಝ್ಝಂ ಝ್ಝಃ	ಝ್ಝ್ ಝ್ಝಾ ಝ್ಝಿ ಝ್ಝೀ
ಞ್ಞ ಞ್ಞಾ ಞ್ಞಿ ಞ್ಞೀ	ಞ್ಞು ಞ್ಞೂ ಞ್ಞೃ ಞ್ಞೆ	ಞ್ಞೇ ಞ್ಞೈ ಞ್ಞೊ ಞ್ಞೋ	ಞ್ಞೌ ಞ್ಞಂ ಞ್ಞಃ	ಞ್ಞ್ ಞ್ಞಾ ಞ್ಞಿ ಞ್ಞೀ
ಟ್ಟ ಟ್ಟಾ ಟ್ಟಿ ಟ್ಟೀ	ಟ್ಟು ಟ್ಟೂ ಟ್ಟೃ ಟ್ಟೆ	ಟ್ಟೇ ಟ್ಟೈ ಟ್ಟೊ ಟ್ಟೋ	ಟ್ಟೌ ಟ್ಟಂ ಟ್ಟಃ	ಟ್ಟ್ ಟ್ಟಾ ಟ್ಟಿ ಟ್ಟೀ
ಠ್ಠ ಠ್ಠಾ ಠ್ಠಿ ಠ್ಠೀ	ಠ್ಠು ಠ್ಠೂ ಠ್ಠೃ ಠ್ಠೆ	ಠ್ಠೇ ಠ್ಠೈ ಠ್ಠೊ ಠ್ಠೋ	ಠ್ಠೌ ಠ್ಠಂ ಠ್ಠಃ	ಠ್ಠ್ ಠ್ಠಾ ಠ್ಠಿ ಠ್ಠೀ
ಡ್ಡ ಡ್ಡಾ ಡ್ಡಿ ಡ್ಡೀ	ಡ್ಡು ಡ್ಡೂ ಡ್ಡೃ ಡ್ಡೆ	ಡ್ಡೇ ಡ್ಡೈ ಡ್ಡೊ ಡ್ಡೋ	ಡ್ಡೌ ಡ್ಡಂ ಡ್ಡಃ	ಡ್ಡ್ ಡ್ಡಾ ಡ್ಡಿ ಡ್ಡೀ
ಢ್ಢ ಢ್ಢಾ ಢ್ಢಿ ಢ್ಢೀ	ಢ್ಢು ಢ್ಢೂ ಢ್ಢೃ ಢ್ಢೆ	ಢ್ಢೇ ಢ್ಢೈ ಢ್ಢೊ ಢ್ಢೋ	ಢ್ಢೌ ಢ್ಢಂ ಢ್ಢಃ	ಢ್ಢ್ ಢ್ಢಾ ಢ್ಢಿ ಢ್ಢೀ
ಣ್ಣ ಣ್ಣಾ ಣ್ಣಿ ಣ್ಣೀ	ಣ್ಣು ಣ್ಣೂ ಣ್ಣೃ ಣ್ಣೆ	ಣ್ಣೇ ಣ್ಣೈ ಣ್ಣೊ ಣ್ಣೋ	ಣ್ಣೌ ಣ್ಣಂ ಣ್ಣಃ	ಣ್ಣ್ ಣ್ಣಾ ಣ್ಣಿ ಣ್ಣೀ
ತ್ತ ತ್ತಾ ತ್ತಿ ತ್ತೀ	ತ್ತು ತ್ತೂ ತ್ತೃ ತ್ತೆ	ತ್ತೇ ತ್ತೈ ತ್ತೊ ತ್ತೋ	ತ್ತೌ ತ್ತಂ ತ್ತಃ	ತ್ತ್ ತ್ತಾ ತ್ತಿ ತ್ತೀ
ಥ್ಥ ಥ್ಥಾ ಥ್ಥಿ ಥ್ಥೀ	ಥ್ಥು ಥ್ಥೂ ಥ್ಥೃ ಥ್ಥೆ	ಥ್ಥೇ ಥ್ಥೈ ಥ್ಥೊ ಥ್ಥೋ	ಥ್ಥೌ ಥ್ಥಂ ಥ್ಥಃ	ಥ್ಥ್ ಥ್ಥಾ ಥ್ಥಿ ಥ್ಥೀ
ದ್ದ ದ್ದಾ ದ್ದಿ ದ್ದೀ	ದ್ದು ದ್ದೂ ದ್ದೃ ದ್ದೆ	ದ್ದೇ ದ್ದೈ ದ್ದೊ ದ್ದೋ	ದ್ದೌ ದ್ದಂ ದ್ದಃ	ದ್ದ್ ದ್ದಾ ದ್ದಿ ದ್ದೀ
ಧ್ಧ ಧ್ಧಾ ಧ್ಧಿ ಧ್ಧೀ	ಧ್ಧು ಧ್ಧೂ ಧ್ಧೃ ಧ್ಧೆ	ಧ್ಧೇ ಧ್ಧೈ ಧ್ಧೊ ಧ್ಧೋ	ಧ್ಧೌ ಧ್ಧಂ ಧ್ಧಃ	ಧ್ಧ್ ಧ್ಧಾ ಧ್ಧಿ ಧ್ಧೀ
ನ್ನ ನ್ನಾ ನ್ನಿ ನ್ನೀ	ನ್ನು ನ್ನೂ ನ್ನೃ ನ್ನೆ	ನ್ನೇ ನ್ನೈ ನ್ನೊ ನ್ನೋ	ನ್ನೌ ನ್ನಂ ನ್ನಃ	ನ್ನ್ ನ್ನಾ ನ್ನಿ ನ್ನೀ
ಪ್ಪ ಪ್ಪಾ ಪ್ಪಿ ಪ್ಪೀ	ಪ್ಪು ಪ್ಪೂ ಪ್ಪೃ ಪ್ಪೆ	ಪ್ಪೇ ಪ್ಪೈ ಪ್ಪೊ ಪ್ಪೋ	ಪ್ಪೌ ಪ್ಪಂ ಪ್ಪಃ	ಪ್ಪ್ ಪ್ಪಾ ಪ್ಪಿ ಪ್ಪೀ
ಫ್ಫ ಫ್ಫಾ ಫ್ಫಿ ಫ್ಫೀ	ಫ್ಫು ಫ್ಫೂ ಫ್ಫೃ ಫ್ಫೆ	ಫ್ಫೇ ಫ್ಫೈ ಫ್ಫೊ ಫ್ಫೋ	ಫ್ಫೌ ಫ್ಫಂ ಫ್ಫಃ	ಫ್ಫ್ ಫ್ಫಾ ಫ್ಫಿ ಫ್ಫೀ
ಬ್ಬ ಬ್ಬಾ ಬ್ಬಿ ಬ್ಬೀ	ಬ್ಬು ಬ್ಬೂ ಬ್ಬೃ ಬ್ಬೆ	ಬ್ಬೇ ಬ್ಬೈ ಬ್ಬೊ ಬ್ಬೋ	ಬ್ಬೌ ಬ್ಬಂ ಬ್ಬಃ	ಬ್ಬ್ ಬ್ಬಾ ಬ್ಬಿ ಬ್ಬೀ
ಭ್ಭ ಭ್ಭಾ ಭ್ಭಿ ಭ್ಭೀ	ಭ್ಭು ಭ್ಭೂ ಭ್ಭೃ ಭ್ಭೆ	ಭ್ಭೇ ಭ್ಭೈ ಭ್ಭೊ ಭ್ಭೋ	ಭ್ಭೌ ಭ್ಭಂ ಭ್ಭಃ	ಭ್ಭ್ ಭ್ಭಾ ಭ್ಭಿ ಭ್ಭೀ
ಮ್ಮ ಮ್ಮಾ ಮ್ಮಿ ಮ್ಮೀ	ಮ್ಮು ಮ್ಮೂ ಮ್ಮೃ ಮ್ಮೆ	ಮ್ಮೇ ಮ್ಮೈ ಮ್ಮೊ ಮ್ಮೋ	ಮ್ಮೌ ಮ್ಮಂ ಮ್ಮಃ	ಮ್ಮ್ ಮ್ಮಾ ಮ್ಮಿ ಮ್ಮೀ
ಯ್ಯ ಯ್ಯಾ ಯ್ಯಿ ಯ್ಯೀ	ಯ್ಯು ಯ್ಯೂ ಯ್ಯೃ ಯ್ಯೆ	ಯ್ಯೇ ಯ್ಯೈ ಯ್ಯೊ ಯ್ಯೋ	ಯ್ಯೌ ಯ್ಯಂ ಯ್ಯಃ	ಯ್ಯ್ ಯ್ಯಾ ಯ್ಯಿ ಯ್ಯೀ
ರ್ರ ರ್ರಾ ರ್ರಿ ರ್ರೀ	ರ್ರು ರ್ರೂ ರ್ರೃ ರ್ರೆ	ರ್ರೇ ರ್ರೈ ರ್ರೊ ರ್ರೋ	ರ್ರೌ ರ್ರಂ ರ್ರಃ	ರ್ರ್ ರ್ರಾ ರ್ರಿ ರ್ರೀ
ಲ್ಲ ಲ್ಲಾ ಲ್ಲಿ ಲ್ಲೀ	ಲ್ಲು ಲ್ಲೂ ಲ್ಲೃ ಲ್ಲೆ	ಲ್ಲೇ ಲ್ಲೈ ಲ್ಲೊ ಲ್ಲೋ	ಲ್ಲೌ ಲ್ಲಂ ಲ್ಲಃ	ಲ್ಲ್ ಲ್ಲಾ ಲ್ಲಿ ಲ್ಲೀ
ವ್ವ ವ್ವಾ ವ್ವಿ ವ್ವೀ	ವ್ವು ವ್ವೂ ವ್ವೃ ವ್ವೆ	ವ್ವೇ ವ್ವೈ ವ್ವೊ ವ್ವೋ	ವ್ವೌ ವ್ವಂ ವ್ವಃ	ವ್ವ್ ವ್ವಾ ವ್ವಿ ವ್ವೀ
ಶ್ಶ ಶ್ಶಾ ಶ್ಶಿ ಶ್ಶೀ	ಶ್ಶು ಶ್ಶೂ ಶ್ಶೃ ಶ್ಶೆ	ಶ್ಶೇ ಶ್ಶೈ ಶ್ಶೊ ಶ್ಶೋ	ಶ್ಶೌ ಶ್ಶಂ ಶ್ಶಃ	ಶ್ಶ್ ಶ್ಶಾ ಶ್ಶಿ ಶ್ಶೀ
ಷ್ಷ ಷ್ಷಾ ಷ್ಷಿ ಷ್ಷೀ	ಷ್ಷು ಷ್ಷೂ ಷ್ಷೃ ಷ್ಷೆ	ಷ್ಷೇ ಷ್ಷೈ ಷ್ಷೊ ಷ್ಷೋ	ಷ್ಷೌ ಷ್ಷಂ ಷ್ಷಃ	ಷ್ಷ್ ಷ್ಷಾ ಷ್ಷಿ ಷ್ಷೀ
ಸ್ಸ ಸ್ಸಾ ಸ್ಸಿ ಸ್ಸೀ	ಸ್ಸು ಸ್ಸೂ ಸ್ಸೃ ಸ್ಸೆ	ಸ್ಸೇ ಸ್ಸೈ ಸ್ಸೊ ಸ್ಸೋ	ಸ್ಸೌ ಸ್ಸಂ ಸ್ಸಃ	ಸ್ಸ್ ಸ್ಸಾ ಸ್ಸಿ ಸ್ಸೀ
ಹ್ಹ ಹ್ಹಾ ಹ್ಹಿ ಹ್ಹೀ	ಹ್ಹು ಹ್ಹೂ ಹ್ಹೃ ಹ್ಹೆ	ಹ್ಹೇ ಹ್ಹೈ ಹ್ಹೊ ಹ್ಹೋ	ಹ್ಹೌ ಹ್ಹಂ ಹ್ಹಃ	ಹ್ಹ್ ಹ್ಹಾ ಹ್ಹಿ ಹ್ಹೀ
ಳ್ಳ ಳ್ಳಾ ಳ್ಳಿ ಳ್ಳೀ	ಳ್ಳು ಳ್ಳೂ ಳ್ಳೃ ಳ್ಳೆ	ಳ್ಳೇ ಳ್ಳೈ ಳ್ಳೊ ಳ್ಳೋ	ಳ್ಳೌ ಳ್ಳಂ ಳ್ಳಃ	ಳ್ಳ್ ಳ್ಳಾ ಳ್ಳಿ ಳ್ಳೀ
ೞ್ೞ ೞ್ೞಾ ೞ್ೞಿ ೞ್ೞೀ	ೞ್ೞು ೞ್ೞೂ ೞ್ೞೃ ೞ್ೞೆ	ೞ್ೞೇ ೞ್ೞೈ ೞ್ೞೊ ೞ್ೞೋ	ೞ್ೞೌ ೞ್ೞಂ ೞ್ೞಃ	ೞ್ೞ್ ೞ್ೞಾ ೞ್ೞಿ ೞ್ೞೀ

Figure 9: Coronavirus leaflet in Muishvng Tangsa using the Lakhum script, April 2020

Figure 10: Draft primer being prepared in September 2020 (front cover, draft pages 1, 2 and 9

/ Lesson 1

/ Vowels



/ Consonants

[illegible] / Lesson 2

[illegible]/ Word forming (Unitelligible)

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

ᐋᐱᑲᐯᐢ ᕽ / Lesson 4

ᐅᑲᓇ ᐢᓂ ᐋᐱᑲᐯᐢ / Learning through Pictures

ᑎ

ᑎᔦᕽ, ᑎᐢᕽ

Kaan, kin

Mountain

ᐯ

ᐯᓇᑲ

Khel

Goat

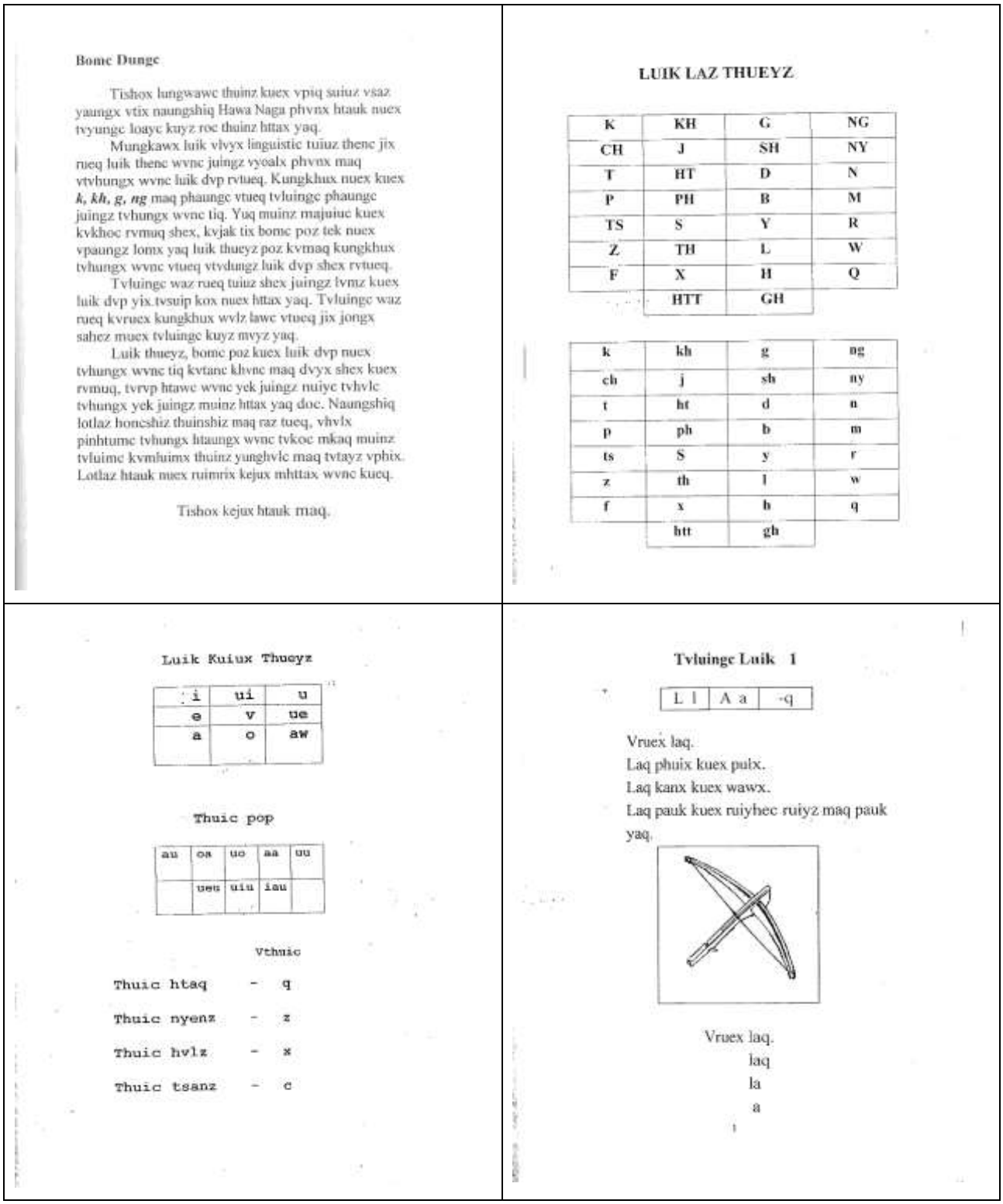
ᖻ

ᖻᕮᕮᕮ

Gehay

Dog

Figure 11: Sample from Rev. Gam Win's Primer for the Muishvung (Mossang) language.



Bomc Dungc

Tishox lungwawc thuinz kuex vpiq suiuz vsaz yaungx vtix naungshiq Hawa Naga phvnx htauk nuex tvyungc loayc kuyz roc thuinz httax yaq.

Mungkawx luik vlvyx linguistic tuiuz thenc jix rueq luik thenc wvnc juingz vyoalx phvnx maq vtvhungx wvnc luik dvp rvtueq. Kungkhux nuex kuex *k, kh, g, ng* maq phaungc vtueq tvluingc phaungc juingz tvhungx wvnc tiq. Yuq muinz majuiuc kuex kvkhoc rvmuq shex, kvjak tix bomc poz tek nuex vpaungz lomx yaq luik thueyz poz kvmaq kungkhux tvhungx wvnc vtueq vtvdungz luik dvp shex rvtueq.

Tvluingc waz rueq tuiuz shex juingz lvmz kuex luik dvp yix.tvsuip kox nuex httax yaq. Tvluingc waz rueq kvruex kungkhux wvlz lawc vtueq jix jongx sahez muex tvluingc kuyz mvyz yaq.

Luik thueyz, bomc poz kuex luik dvp nuex tvhungx wvnc tiq kvtanc khvnc maq dvyx shex kuex rvmuq, tvrvp htawc wvnc yek juingz nuiyc tvhvlc tvhungx yek juingz muinz httax yaq doc. Naungshiq lotlaz honcshiz thuinshiz maq raz tueq, vhvlx pinhtumc tvhungx htaungx wvnc tvkoc mkaq muinz tvluimc kvmhuimx thuinz yunghvlc maq tvtayz vphix. Lotlaz htauk nuex ruimrix kejux mhttax wvnc kueq.

Tishox kejux htauk maq.

LUIK LAZ THUEYZ

K	KH	G	NG
CH	J	SH	NY
T	HT	D	N
P	PH	B	M
TS	S	Y	R
Z	TH	L	W
F	X	H	Q
	HTT	GH	

k	kh	g	ng
ch	j	sh	ny
t	ht	d	n
p	ph	b	m
ts	S	y	r
z	th	l	w
f	x	h	q
	htt	gh	

Luik Kuiux Thueyz

i	ui	u
e	v	ue
a	o	aw

Thuic pop

au	oa	uo	aa	uu
	ueu	uiu	iau	

Vthuic

Thuic htaq - q

Thuic nyenz - z

Thuic hvlz - x

Thuic tsanz - c

Tvluingc Luik 1

L l	A a	-q

Vruex laq.
Laq phuix kuex pulx.
Laq kanx kuex wawx.
Laq pauk kuex ruiyhec ruiyz maq pauk yaq.

Vruex laq.
laq
la
a

Bibliography

Morey, Stephen. (2017). ‘Tangsa’. In Graham Thurgood and Randy LaPolla (eds). *The Sino-Tibetan Languages*. London and New York: Routledge (2nd Edition). 350-368. ISBN 9781138783324

Morey, Stephen. (2015). ‘The internal diversity of Tangsa: vocabulary and morphosyntax.’ In Mark W. Post, Stephen Morey and Scott DeLancey (eds). *Language and culture in Northeast India and beyond: in honour of Robbins Burling.* Canberra, Asia-Pacific Linguistics, Language and Peoples of the Eastern Himalaya Region (series). 23-40. ISBN: 9781922185266, eISBN: 9781922185259.

Khämlan Binkhäm and W. Owen (2018). *Orthography and text engineering statement for the Tangshang Naga Unified Orthography*. Published by the authors.

van Dam, Kellen Parker. (2018). *The Tone System of Tangsa-Nocte and Related Northern Naga Varieties*. PhD Dissertation, Department of Languages and Linguistics, La Trobe University.

Acknowledgements

The first and major acknowledgement should be to Shri Lakhum Mossang, the creator of this script, wo has laboured for thirty years to bring this script into more widespread use. We also acknowledge the members of the Tangsa Script Development Committee, particularly the secretary, Mr. Sengkhum Mossang and Mr. Wanglung Mossang, who has spent many hours discussing this script with Stephen Morey and Kellen Parker van Dam and producing materials using the script.

The initial font was developed by Karen Parker. Subsequent improvements to the font were made by Kellen Parker van Dam, who has also contr

Deborah Anderson has been very encouraging and helpful with comments and suggestions throughout the preparation of this document. Her strong support for the work on smaller scripts throughout the world is much valued and appreciated. Deborah’s assistance has been made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project PR-253360-17 (part of the Script Encoding Initiative at UC Berkeley). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities. We also acknowledge Anshuman Pandey who produced an early draft some years ago.

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[2]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html
See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** ***Tangsa***
2. Requester's name: *Stephen Morey*
3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*
4. Submission date: *2020*
5. Requester's reference (if applicable): *?*
6. Choose one of the following:
   This is a complete proposal: *X*
   (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
   a. This proposal is for a new script (set of characters): *X*
      Proposed name of script: *Tangsa*
   b. The proposal is for addition of character(s) to an existing block:
      Name of the existing block:
2. Number of characters in proposal: *89*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary *?* B.1-Specialized (small collection) B.2-Specialized (large collection)
   C-Major extinct D-Attested extinct E-Minor extinct
   F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usage symbols
4. Is a repertoire including character names provided? *yes*
   a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? *yes*
   b. Are the character shapes attached in a legible form suitable for review? *yes*
5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? *Stephen Morey and Kellen Parker van Dam*
   If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
6. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *yes*
   b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *yes*
7. Special encoding issues:
   Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? *no*

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see http://www.unicode.org/Public/UNIDATA/UCD.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

[2] Form number: N3102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? *yes*
   If YES explain *Introductory proposal by Anshuman Pandey L2/13-231.*
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? *yes*
   If YES, with whom? *Tangsa Script Development Committee in India*
   If YES, available relevant documents:
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? *yes*
   Reference: *this document*
4. The context of use for the proposed characters (type of use; common or rare) *rare, but developing*
   Reference:
5. Are the proposed characters in current use by the user community? yes
   If YES, where? Reference:
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? *no*
   If YES, is a rationale provided?
   If YES, reference:
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? *yes*
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? *no*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? *no*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? *no*
    If YES, is a rationale for its inclusion provided?
    If YES, reference:
11. Does the proposal include use of combining characters and/or use of composite sequences? *yes*
    If YES, is a rationale for such use provided? *yes*
    If YES, reference: *this document*
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? *no*
    If YES, reference:
12. Does the proposal contain characters with any special properties such as control function or similar semantics? *no*
    If YES, describe in detail (include attachment if necessary)
13. Does the proposal contain any Ideographic compatibility character(s)? *no*
    If YES, is the equivalent corresponding unified ideographic character(s) identified?
    If YES, reference: