

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: 2021-01-29

1. Introduction: The script proposed to be named ‘Tangsa’ is a script developed since 1990 by Mr Lakhum Mossang (died 11th July 2020), of Namphei Nong, Miao, Arunachal Pradesh. It was developed 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 (see Section 1.4 below for more detail about the Tangsa languages). It is an 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. Since the creation of the script in 1990, it has been promoted for many years by its creator, Mr Lakhum Mossang, in the early years mostly being taught in small face-to-face classes and in handwritten form. The creator produced handwritten documents

Tangsa is not a single language (see section 1.3 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 Muishvung¹. 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.

Throughout the 30 year history of the Tangsa script, the basic characters (vowels + tones and consonants, see below section 1.2) and the guiding principals of the script have remained stable. The relatively minor changes introduced in 2019 have been made to allow the script to cover all the possible consonants in all the Tangsa varieties and also to ensure that all the vowels are treated by the script in a consistent manner. The addition of the vowel Series TANGSA LETTER V, discussed briefly in Section 1.3 below.

Note that earlier versions of this proposal named the script as LAKHUM MOSSANG TANGSA, so that, for example the first letter would be named LAKHUM MOSSANG TANGSA LETTER OZ (U+16A70). This name was felt to be too long and was thus reduced. It should be noted (as discussed further below) that there are other writing systems in use by members of the Tangsa community, mostly Roman based orthographies. The script developed by the late Lakhum Mossang, however, is the only native script that has a long history of usage (since 1990) and the support of community organisations, as detailed below.

1.1 Brief history of the usage of the Tangsa script 1990-2020

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 be notated as /ə/) should replace the letter a. It would be realised as [muw²ʃəuŋ²] 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 its own name for each other group, so the Mossang are also called [mjo²xaŋ²] by the Cholim, [mjaŋ²sa²] by the Lauchäng, [muw²ʃa²] by the Shecya &c.

The script has been taught in a handwritten form by the late Mr. Lakhum Mossang over the past 30 years. His first draft of the inventory letters, containing 73 symbols, is presented in Figure 1 below. Lakhum Mossang taught the letters by numbers, and these numbers appear on Figure 1. He also produced a number of hand-written documents with traditional songs (Figure 2) and a traditional prayer (Figure 3), both as examples for his students and to show the community as a whole how a script like this could be used to preserve culture.

Only one photograph of his efforts in teaching the script has been found, and is included as Figure 4. This is dated to approximately 2005.

Lakhum Mossang also regularly attended festivals and public community events in which he would demonstrate his script with the aid of posters, such as can be seen in Figure 5, where he is standing with Stephen Morey, and two of his most committed students, Mr. Manpo Kelim and Mr. Jokkam Jokkah.

Since at least 2003, Lakhum Mossang has been discussing with community organisations and with the Government of Arunachal Pradesh about how the script can be promoted so as to give the Tangsa community more access to literacy and the protection of their culture. Figure 6 is a letter from 2005 relating to meetings about script and literacy developed, and Figure 7 is a letter from the Government of Arunachal Pradesh requesting Mr Lakhum Mossang's attendance at a meeting about script development.

The number of users of the script is growing rapidly, as at the beginning of 2021 and is estimated at at least 100. The users include school students in at least one Primary School (see Figure 19), community members attending a 15 day workshop being held in January-February 2021, members of the Tangsa Script Development Committee and in addition a small number of dedicated students who have learned to use the script fluently over many years, all of them members of the Muishvung (Mossang) community. Examples of the kind of work produced by these students are found in Figure 8 which includes stories and words lists written for the script classes and showing corrections, Figure 9 shows the use of the script in every day situations, daily sales records, receipts and marketing lists and Figure 10 includes traditional songs, and stories written by Mr. Manpo Kelim. Note that in Figure 10.3 we can see not only the notebook in which he has written a story in Tangsa script, but underneath it a published document on which he has written using the script. This is an indication of the extent to which some individuals are using the script.

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 Muishvung (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 in Figure 11 below in the 'Examples' section. One of the decisions of the Script Development committee was a slight alteration to the script, with addition of one more vowel character (the Tangsa Letter V series, and a couple of consonants to be used to write Tangsa varieties other than Muishvung, see further section 1.3 below). The revised list of characters, which is the list that forms the basis of this proposal, is given in Figure 12.

Throughout 2020, a primer to teach the script, at least for the Muishvung variety was under preparation (this preparation was somewhat delayed by the very hard lockdown in India in April-May 2020). Early draft versions of this primer are shown in both Figures 13 and 15, and the version presented to Mr Tapi Gao, Director of Elementary Education, Government of Arunachal Pradesh, November 2020 is given as Figure 17. (A photograph of the draft primer being presented is Figure 16).

The preparation of the primer was a request of the Arunachal Pradesh Directorate of Elementary Education (DEE). The script was submitted to the DEE in April 2020 and in August 2020, the members of the Tangsa Script Development Committee were asked by the Directorate of Elementary Education to produce these teaching materials, Note that the development of these teaching materials, which the DEE will now produce, is a pre-requisite for the introduction of the script into school education; as Wanglung Mossang wrote on 1st

December 2020, this is “not happening yet, but the book has been handed to the DEE, and when the Education department finishes printing the books, he (Director of the DEE) will give an official order and then only the schools will start teaching”

In addition, some online and digital materials have been produced in the script for teaching and for communication. In 2012, the late Lakhum Mossang commissioned and produced a series of CDs that would explain the script for persons with access to laptop computers. .mp4 versions of the first of these CDs (consisting of a short introduction and the main video of one hours duration) have been placed on a YouTube channel at <https://youtu.be/IJeLzLWGM3E> and <https://youtu.be/vHFGg5uJM5I>.

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).

In December 2020, the Tangsa Script Development Committee released a 2021 Calendar using the Tangsa script, reported on Facebook at: <https://www.facebook.com/felixaantony/posts/10214353599193688>.

1.2 The form and structure of the Tangsa script

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 / Muishvung 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) (See Figure 18 below for a sample of Rev Gam Win’s primer).

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 Juinz Luik Dap Hewa Naga*), gave them names in Muishvung 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 Muishvung	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 Muishvung word for ‘to go’ 𑜀𑜢𑜤𑜰 (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 Muishvung). 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, 𑜀𑜢𑜤𑜰 𑜀𑜢𑜤𑜰 𑜀𑜢𑜤𑜰 𑜀𑜢𑜤𑜰 (TANGSA LETTER UIUZ U+16A98 through to TANGSA LETTER UIUX, U+ 16A9B) which are actually diphthongs, the combination of two vowels together [u] 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 𑜀𑜢𑜤𑜰 (TANGSA LETTER NGA, U+16AA3), but in syllable final position it is written with 𑜀𑜢𑜤𑜰 (TANGSA LETTER FINAL NG, U+16A90). The latter symbol is listed with the vowels. Thus the word for ‘feel jealous’ is written as 𑜀𑜢𑜤𑜰𑜃𑜫𑜀𑜢𑜤𑜰 (TANGSA LETTER NGA, U+16AA3, TANGSA LETTER EX, U+16A7F, and TANGSA LETTER FINAL NG, U+16A90) [ŋen²]. 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 𑜀𑜢𑜤𑜰𑜃𑜫𑜀𑜢𑜤𑜰 [k^hun²nen²] (a song language word meaning ‘tendency of a person to desire others not to get a benefit’), where in the second syllable, 𑜀𑜢𑜤𑜰, TANGSA LETTER NA, U+16AAC is used in both initial and syllable final position.

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 Muishvung 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 Muishvung (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 Muishvung 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 Muishvung.

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 Muishvung (Mossang) as [ɔ] (U+16A8A, TANGSA LETTER AWQ) and for the letter [z] (U+16ABE, TANGSA LETTER ZA.) 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 11. The revised set of symbols promoted by that committee is included as Figure 12.

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 Muishvung would be treated. This is discussed in section 9 below:

1.3 The addition of the TANGSA LETTER V series

The most significant change to the script that was undertaken in 2019 revision was to add a single vowel, TANGSA LETTER V. However, as already discussed, since the vowels in the language each have four symbols, for each of the four tones, this necessitated the addition of four new symbols, TANGSA LETTER VZ (U+16A78), TANGSA LETTER VC (U+16A79), TANGSA LETTER VQ (U+16A7A) and TANGSA LETTER VX (U+16A7B).

These new vowels were created by taking the TANGSA LETTER A series (U+16A74 to 16A77) and adding a small ‘crossing dash’ to each of the letters in that series.

The reason for the creation of the new series of TANGSA LETTER V is that in the initial creation of the script, Lakhum Mossang did not make a clear distinction between the long [a] like sound and the shorter sound that is more similar to the schwa [ə]. Prior to the revision of the script, to write words with the short TANGSA LETTER V, it was necessary to add a TANGSA LETTER I to the end of the syllable. Thus the word for ‘rice’ /tsəm¹/ʳᵐᵐ was early written TANGSA LETTER TSA (16AB6). TANGSA LETTER AZ (16A74), TANGSA LETTER MA (16AAB) and TANGSA LETTER IZ (16A80). This could also be pronounced /tsa¹mi¹/ and was potentially ambiguous. The new way of writing this word is ʳᵐᵐ TANGSA LETTER TSA (16AB6). TANGSA LETTER VZ (16A78) and TANGSA LETTER MA (16AAB) and is not ambiguous.

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

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 Cham-pang). 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).

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 definitely not mutually intelligible.

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


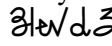
2. Structure: The characters are all written left to right. Many of the words in the language are monosyllabic, and these syllables take the form of:



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

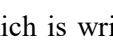
VOWEL + TONE

OPTIONAL SECOND VOWEL + TONE


OPTIONAL FINAL CONSONANT

For example, the word for ‘sky’, written as [rauŋ²] phonemically, is given as  (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 [rauŋ²] phonemically, is given as  (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  (TANGSA LETTER KA (U+16AA0), TANGSA LETTER AQ (U+16A76), or it can be written with an additional vowel symbol, as  (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 /ha¹/  (TANGSA LETTER HA (U+16AAd), TANGSA LETTER AZ (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) [o³] 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 U+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 (Muishvung) 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 = [ʏ]

TANGSA LETTER UIU = [ɯɯ]

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

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 Muishvung 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 Muishvung 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	ᨣ	TANGSA LETTER HTA	[t ^h]	
16AB3	ᨡ	TANGSA LETTER NHA	[ɲ]	dental nasal
16AB5	ᨢ	TANGSA LETTER CA	[tɕ]	written with <j> in Gam Win's system
16AB7	ᨣ	TANGSA LETTER GHA	[ɣ]	
16AB8	ᨤ	TANGSA LETTER HTTA	[tʰ]	aspirated voiceless dental stop
16AB9	ᨥ	TANGSA LETTER THA	[t]	unaspirated voiceless dental stop
16ABC	ᨦ	TANGSA LETTER DHA	[ð]	a voiced dental fricative
16ABD	ᨧ	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 Muishvung (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	ᨢᨣ	TANGSA LETTER FA
16ABC	ᨢᨦ	TANGSA LETTER DHA
16ABD	ᨢᨧ	TANGSA LETTER CHA
16ABE	ᨢᨣᨣ	TANGSA LETTER ZA

For example, the TANGSA LETTER FA is made up of the 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 Muishvung, 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 𑜄𑜰 TANGSA LETTER LONG UEX
 16A92 𑜄𑜱 TANGSA LETTER SHORT UEZ
 16A93 𑜄𑜲 TANGSA LETTER SHORT AWX

The two prosodic symbols, 16A91 and 16A93 are used as follows. A long falling tone is used in the phrase 𑜀𑜃𑜫 𑜁𑜨𑜃𑜫 [nɔː² kəra²], ‘**very very** far’ where the first word uses TANGSA LETTER NA, U+16AAC and TANGSA LETTER LONG UEX, U+16A91 rather than being 𑜀𑜃𑜫 [nɔ²], 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 𑜀𑜃𑜫 𑜁𑜨 [vrɔ² lɔ²], 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 𑜁𑜨 [lɔ²], an imperative particle, is written with TANGSA LETTER LA, U+16AAE, and TANGSA LETTER SHORT AWX, U+16A93.

In the Muishvung 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 𑜀𑜃𑜫 (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 (𑜄𑜱) (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	təɣui¹	𑜀𑜃𑜫𑜁𑜨𑜃𑜫
creator	shvkex	ʃəke²	𑜀𑜃𑜫𑜁𑜨
likewise	kvrulc mvrenc	kərul³ məren³	𑜀𑜃𑜫𑜁𑜨𑜃𑜫𑜁𑜨𑜃𑜫
others	wvghanc	βəɣan³	𑜀𑜃𑜫𑜁𑜨𑜃𑜫
lastly	vtvsuip	ətəsuiɸ	𑜀𑜃𑜫𑜁𑜨𑜃𑜫𑜁𑜨

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 Muishvung (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 𑜄𑜰𑜫 TANGSA LETTER MC
 16A9E 𑜄𑜱𑜫 TANGSA LETTER MQ
 16A9F 𑜄𑜲𑜫 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 [ts^h], [j] (where it is a distinct phoneme from [dʒ] / [dʒ̥] and possibly [tʃ]). 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 Muishvung (Mossang)

As mentioned earlier, the vowel symbols combine vowel and tone, but the tones of other Tangsa varieties are different from Muishvung (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 Muishvung (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 Muishvung 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 [tʃ] 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 different sub-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.

U+16A70

Tangsa Script

U+16ACF

	16A7	16A8	16A9	16AA	16AB	16AC
0	ᄀ	ᄁ	ᄂ	ᄃ	ᄄ	ᄅ
1	ᄆ	ᄇ	ᄈ	ᄉ	ᄊ	ᄋ
2	ᄌ	ᄍ	ᄎ	ᄏ	ᄐ	ᄑ
3	ᄒ	ᄓ	ᄔ	ᄕ	ᄖ	ᄗ
4	ᄙ	ᄚ	ᄛ	ᄜ	ᄝ	ᄞ
5	ᄟ	ᄠ	ᄡ	ᄢ	ᄣ	ᄤ
6	ᄥ	ᄦ	ᄧ	ᄨ	ᄩ	ᄪ
7	ᄫ	ᄬ	ᄭ	ᄮ	ᄯ	ᄰ
8	ᄱ	ᄲ	ᄳ	ᄴ	ᄵ	ᄶ
9	ᄷ	ᄸ	ᄹ	ᄺ	ᄻ	ᄼ
A	ᄽ	ᄾ	ᄿ	ᄿ	ᄿ	
B	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	
C	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	
D	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	
E	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	
F	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	

Vowels

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

16A9A	ɔ	TANGSA LETTER UIUQ
16A9B	ɔ	TANGSA LETTER UIUX
16A9C	ɔ	TANGSA LETTER MZ
16A9D	ɔ	TANGSA LETTER MC
16A9E	ɔ	TANGSA LETTER MQ
16A9F	ɔ	TANGSA LETTER MX

Consonants

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

16ABE ꨑ TANGSA LETTER ZA

Digits

16AC0 ① TANGSA DIGIT ZERO

16AC1 ୨ TANGSA DIGIT ONE

16AC2 ୩ TANGSA DIGIT TWO

16AC3 ୪ TANGSA DIGIT THREE

16AC4 ୫ TANGSA DIGIT FOUR

16AC5 ୬ TANGSA DIGIT FIVE

16AC6 ୭ TANGSA DIGIT SIX

16AC7 ୮ TANGSA DIGIT SEVEN

16AC8 ୯ TANGSA DIGIT EIGHT

16AC9 ୦ 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 documents in the script are those that have been produced in 2019-20, such as the document produced by the Script Committee in 2020, samples of which appear in Figure 11 and the draft primer in Figures 15 and 17)

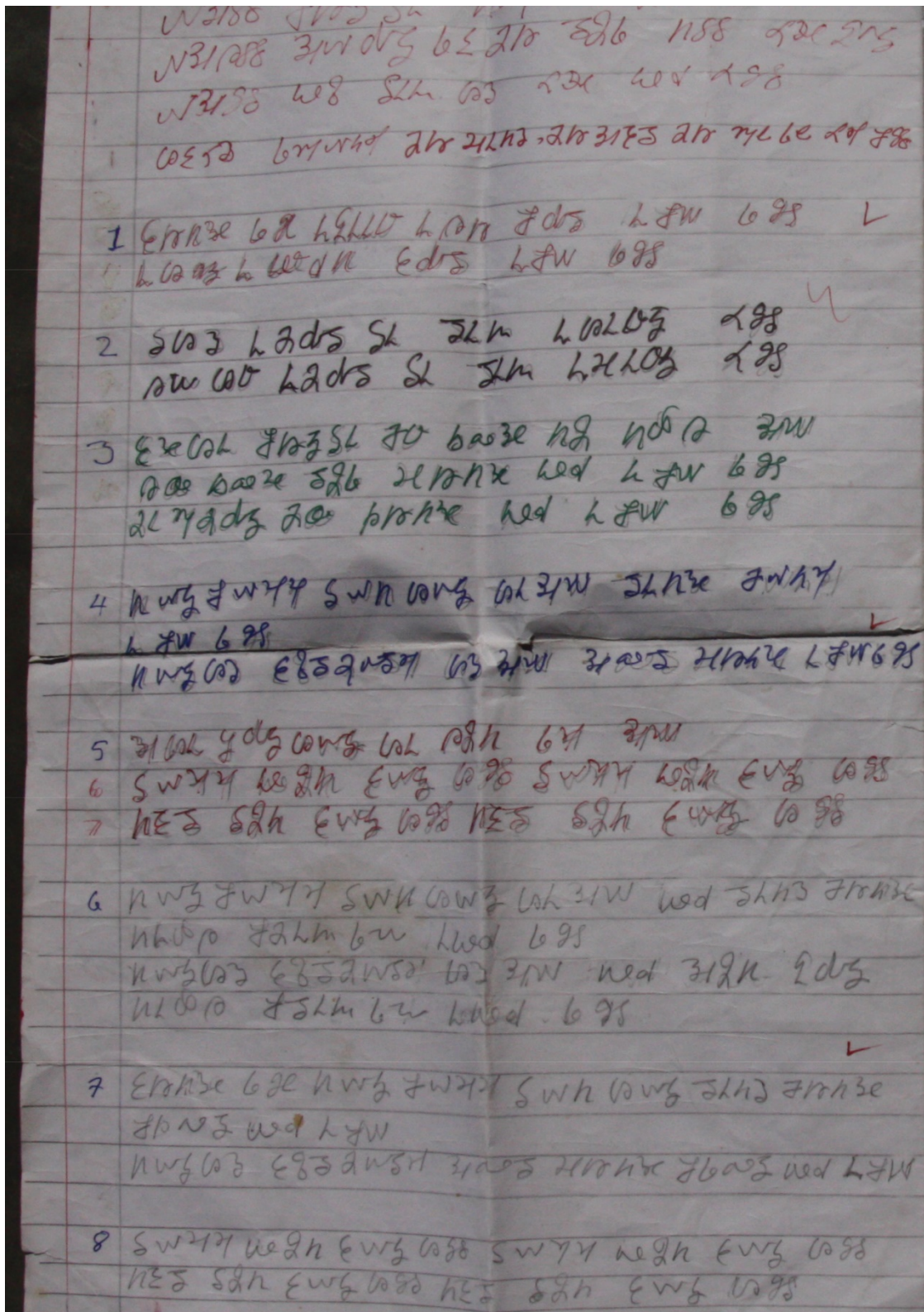
Figure 1: Full list of Lakhum Mossang's script (2003 version)

1990 - 2010 - TANGSA (TANG-SHANG) NAGA EDUCATION DEVELOPMENT COMMON SCRIPT				PART - II				NUMBERS	
PART - I LINGUAL - 73									
d 1	2	3	4	47	48	49	50	1.	9
5	6	7	8	51	52	53	54	2	h
9	10	11	12	55	56	57	58	3	e
13	14	15	16	59	60	61	62	4	u
17	18	19	20	63	64	65	66	5	e
21	22	23		67	68	69	70	6	u
24	25	26	27	71	72	73		7	e
28								8	9
29	30	31	32					9	0
33	34	35							
36	37	38	39						
40	41	42	43						
44	45	46							

Script Founder
LAKHUM MOSSANG
P.O. Namphai - I

9402445754 @ 3863340254 ©

Figure 2: Wihu Song Manuscript written by Lakhum Mossang c. 2005



SL. 18, DATE: 28-12-2010
SL 48 12/1/2013

- [illegible]

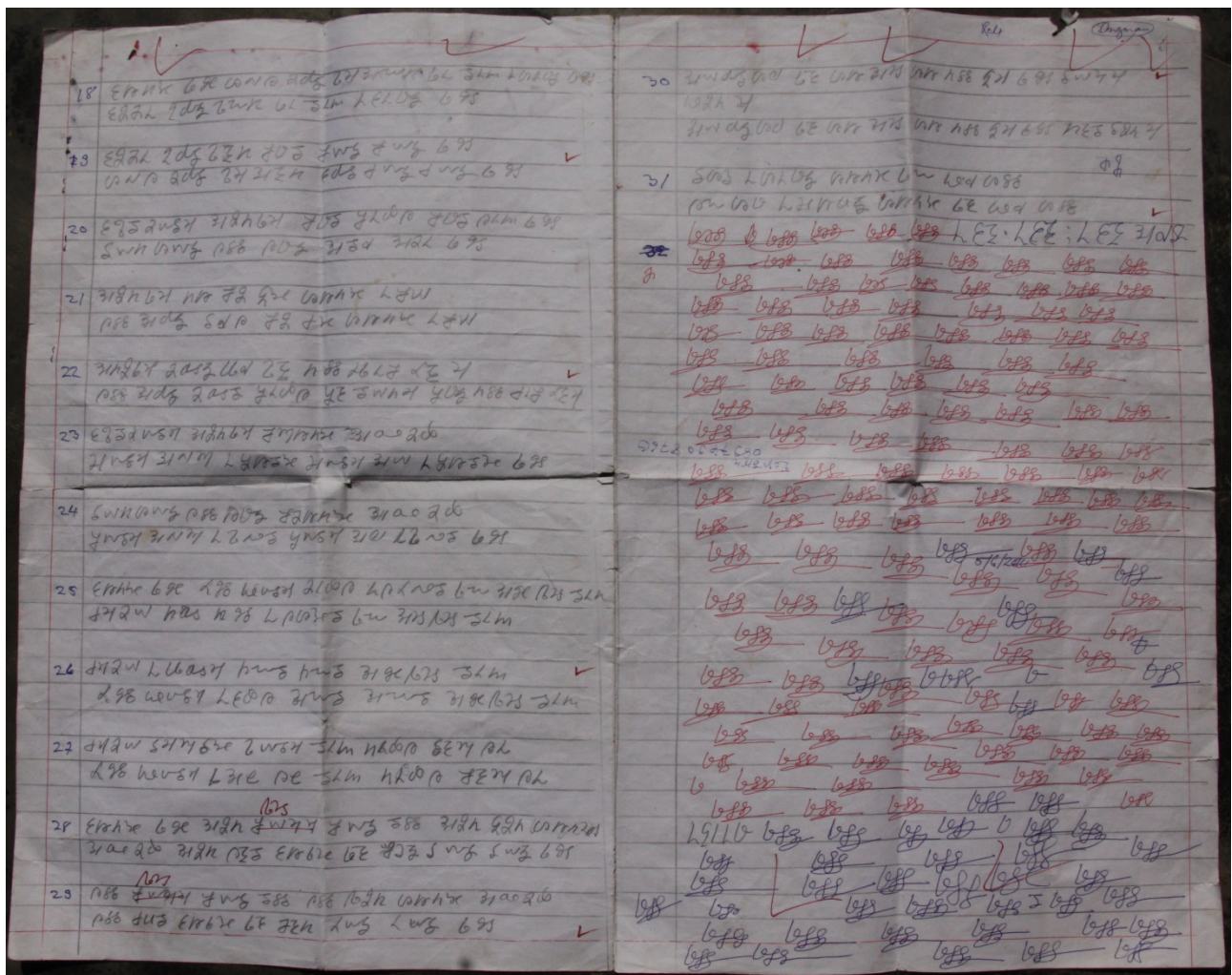


Figure 3: Traditional Prayer (Rimrim) written by Lakhum Mossang (before 2010)

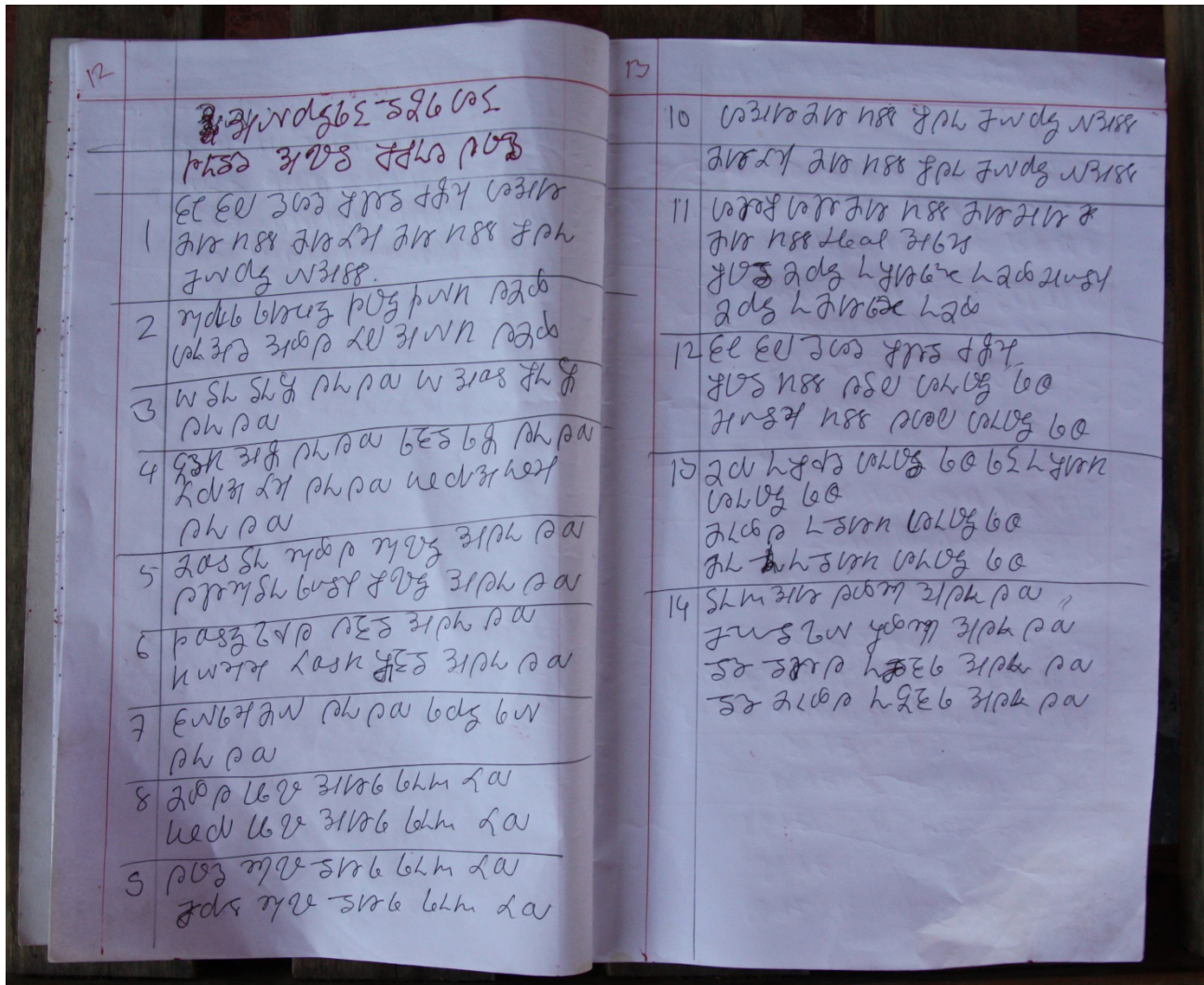
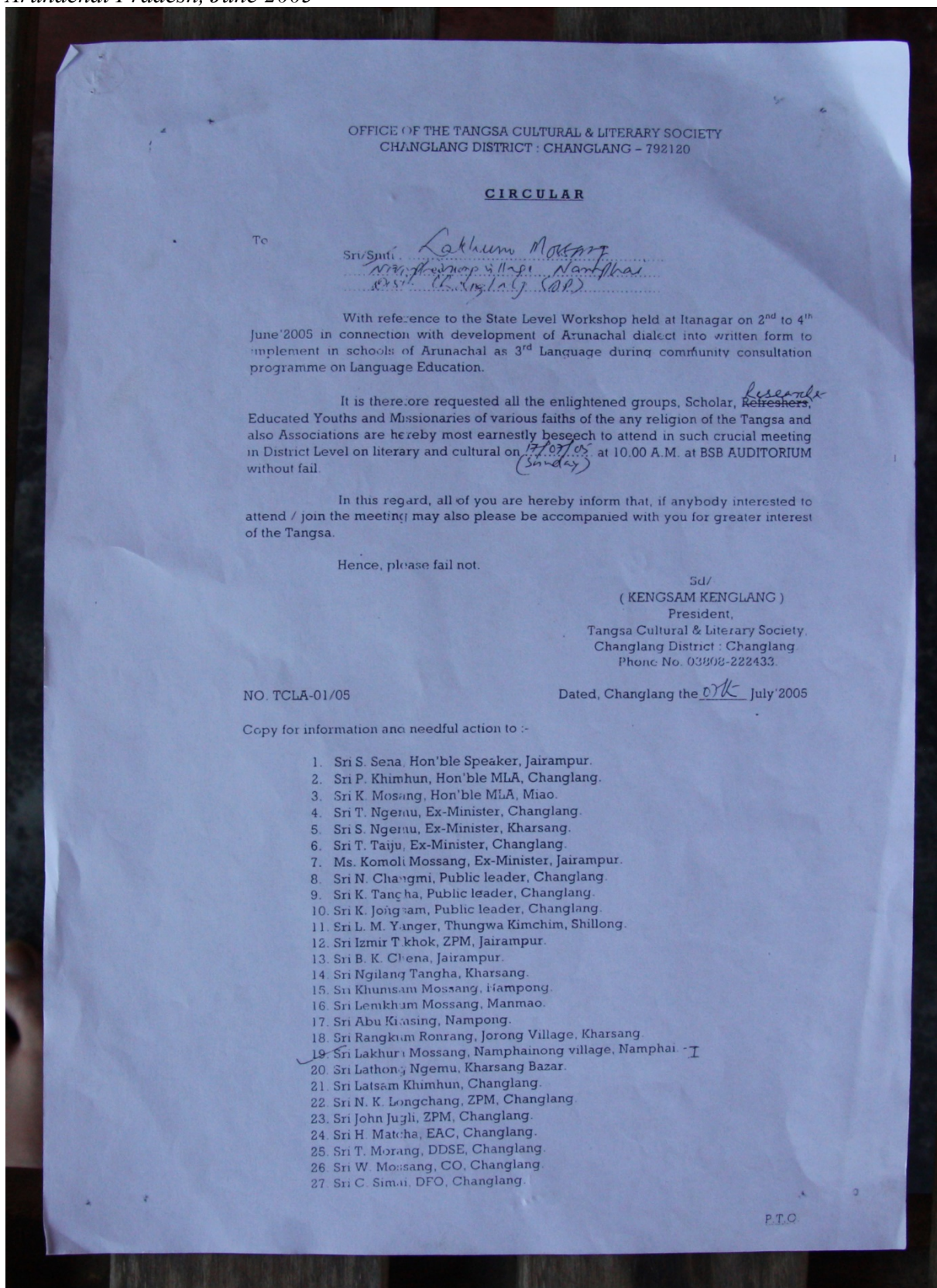


Figure 5: Lakhum Mossang standing with his students and Dr. Stephen Morey in front of the document showing his script, 2013



Figure 6: Invitation to Lakhum Mossang to attend a workshop on the development of literacy in Arunachal Pradesh, June 2005



29

28. Sri D. Chhangmai, I/c DDI, Changlang.
29. Sri L. Tongha, HDO, Changlang.
30. Sri T. Teng ha, OC SB Branch, Changlang.
31. Sri Hon kain Tungkhah, JT, Changlang.
32. Sri C. Tak ul, Changlang.
33. Sri K. D. Nyeinu, DC's Office, Changlang.
34. Sri Dhujiang Sainang, Changlang.
35. Sri Wangram Zongsam, New Changlang village.
36. Sri Maitu Morang, Nampong.
37. Sri S. Nge nu, Changlang.
38. Sri Samtai Zongsam, Yungchen village, Miao.
39. Sri Khoimt : Khonrang, Yungchen village, Miao.
40. Sri Chomtang Kimshing, Miao.
41. Sri Rangput Jugli, Kharsang.
42. Sri Khachan Jugli, Kharsang.
43. The Language Officer, Changlang District. Changlang for information and necessary action. He is requested to kindly arrange sitting arrangements at BSD auditorium without fail.

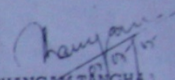

(THANGAM TANCHIA)
Secretary,
Tangsa Cultural & Literary Society,
Changlang District : Changlang
Phone No. 03808-222594.

Figure 7: Letter from the Government of Arunachal Pradesh to Mr Lakhum Mossang requestion his attendance at a meeting relating to script development, 16th May 2018



**GOVERNMENT OF ARUNACHAL PRADESH
STATE COUNCIL OF EDUCATIONAL RESEARCH AND TRAINING
GOHPUR TINALI, ITANAGAR - 791113**

NO.SCERT/GEN-1/

Dated, Itanagar the 16th May'2018

To,

The Director of Elementary Education
Govt. of Arunachal Pradesh
Itanagar.

Sub: Representation of Shri Lakhum Jogkah Mossang for development of Tangsa Language.

Sir,

I am forwarding herewith original letter received from Shri Lakhum Jogkah Mossang duly forwarded to SCERT by Hon'ble Minister (Education). In this regard, I would like to highlight that this is a matter of development of Tangsa Language which falls under third language for Elementary Education.

Therefore, necessary guidance in the form of financial help for development of Tangsa language for use in Tangsa inhabited area may be kindly looked into for third language.

This is for your kind information and necessary action please.

Yours faithfully

(Moto Nyori)
Jt. Director
SCERT, Itanagar

NO.SCERT/GEN-1/

Dated, Itanagar the 16th May'2018

Copy to:

1. The PS to Hon'ble Minister (Education), Govt. of Arunachal Pradesh, Itanagar for your kind information.
2. Shri Lakhum Jogkah Mossang, for information.
3. Office copy.

(Moto Nyori)
Jt. Director
SCERT, Itanagar

8.2 Document showing word list as written by one of the students and corrected by Lakhum Mossang (c. 2015)



Figure 9: Use of the Tangsa script for daily sale records, receipts, marketing lists &c. 2020

⑦ Thankai Maa

Total - Rs. 302/-

18/11/21 S32h31 - Rs. 20/-

20/11/21 S32h31 Rs. 20/-

12/12/21 Shh3e Rs. 5/-

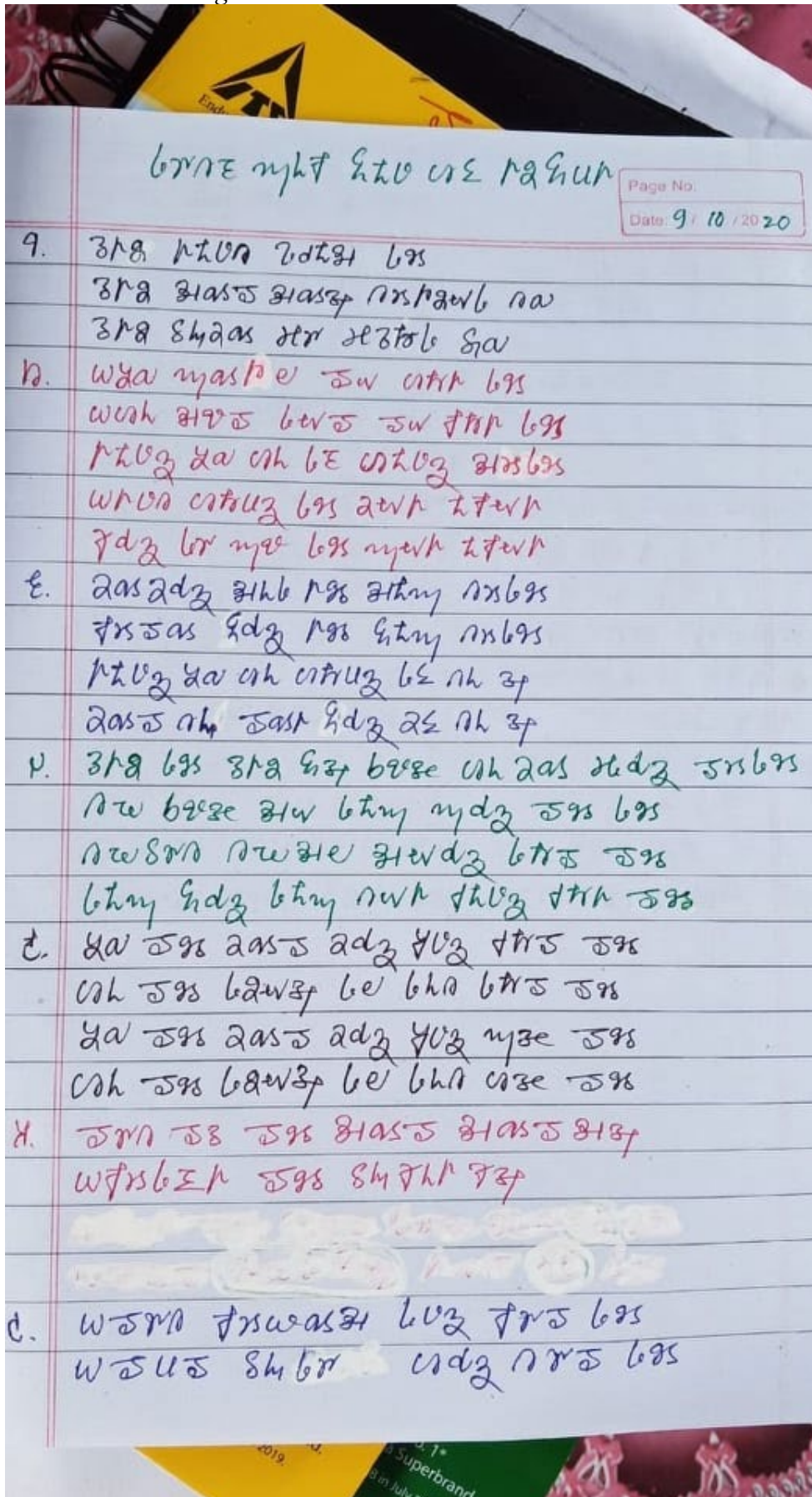
24/12/21 S32h31 Rs. 20/-

25/12/21 S32h31 Rs. 40/-

12/1/22 Shh3e Rs. 10/-

S32h31 Rs. 40/-

Figure 10: Documents produced in language by Mr. Manpo, 2020
10.1 Traditional song



[illegible]

Figure 11: Document produced by Tangsa Script Development Committee, January 2020
11.1 p1

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

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

Diverse way of expressing in Tangsa dialects

Mossang- How are you/ Are you fine?

ᐱᐸᐱᐸ ᐸᐸ

Moklum-

ᐱᐸᐱᐸ ᐸᐸ

Longchang-

ᐱᐸᐸ ᐸᐸ

Kimsing-

ᐱᐸᐸ ᐸᐸ

Longri-

ᐱᐸᐸ ᐸᐸ

Mungrey-

ᐱᐸᐸ ᐸᐸ

Cholim-

ᐱᐸᐸ ᐸᐸ

Figure 12: Revised list of Tangsa Script characters, January 2020

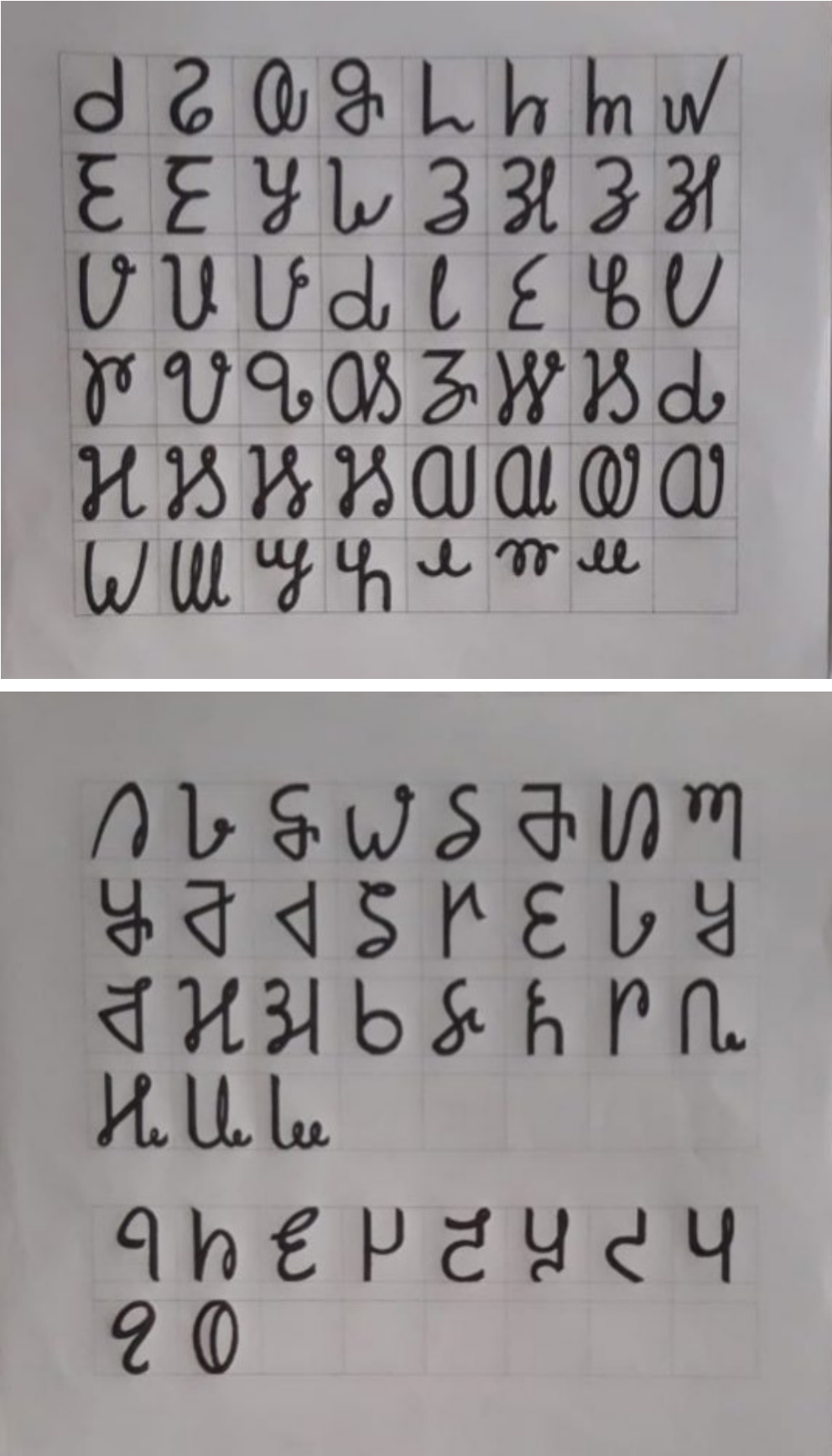


Figure 13: Samples of first two ‘lessons’ of the draft Primer, April 2020

ᐱᐱᐱᐱᐱ ᠑

᠑

(Lesson 1)

ᐱᐱᐱᐱᐱ (Vowels)

᠔	᠔	᠒	᠔	᠕	᠎	᠎	᠕
oz	oc	oq	ox	az	ac	aq	ax
᠕	᠎	᠎	᠕	᠑	᠑	᠕	᠕
vz	vc	vq	vx	ez	ec	eq	ex
᠔	᠔	᠔	᠔	᠕	᠕	᠕	᠕
iz	ic	iq	ix	uz	uc	uq	ux
᠕	᠕	᠕	᠕	᠕	᠕	᠕	᠕
uiz	uic	uiq	uix	awz	awc	awq	awx
᠔	᠕	᠕	᠕	᠕	᠕	᠕	᠕
eng	uex	uezz	awx	uec	uec	ueq	uex
᠒	᠒	᠒	᠒	᠕	᠕	᠕	᠕
uiuz	uiuc	uiuq	uiux	mz	mc	mq	mx

ᐱᐱᐱᐱᐱ (Consonants)

᠒	᠕	᠕	᠕	᠕	᠕	᠕	᠕
k	kh	g	ng	s	y	w	p
᠕	᠕	᠕	᠕	᠕	᠕	᠕	᠕
ny	ph	b	m	n	h	l	ht
᠕	᠕	᠕	᠕	᠕	᠕	᠕	᠕
t	d	r	nh	sh	j	ts	gh
᠕	᠕	᠕	᠕	᠕	᠕	᠕	
htt	th	x	f	dh	ch	z	

᠕᠕᠕᠕᠕᠕᠕ (Numerals)

᠑	᠎	᠑	᠕	᠑	᠕	᠑	᠕	᠑	᠑᠒
᠑᠑	᠑᠎	᠑᠑	᠑᠕	᠑᠑	᠑᠕	᠑᠑	᠑᠕	᠑᠑	᠎᠒
᠎᠑	᠎᠎	᠎᠑	᠎᠕	᠎᠑	᠎᠕	᠎᠑	᠎᠕	᠎᠑	᠑᠒
᠑᠑	᠑᠎	᠑᠑	᠑᠕	᠑᠑	᠑᠕	᠑᠑	᠑᠕	᠑᠑	᠕᠒
᠕᠑	᠕᠎	᠕᠑	᠕᠕	᠕᠑	᠕᠕	᠕᠑	᠕᠕	᠕᠑	᠑᠒
᠑᠑	᠑᠎	᠑᠑	᠑᠕	᠑᠑	᠑᠕	᠑᠑	᠑᠕	᠑᠑	᠕᠒
᠕᠑	᠕᠎	᠕᠑	᠕᠕	᠕᠑	᠕᠕	᠕᠑	᠕᠕	᠕᠑	᠑᠒
᠑᠑	᠑᠎	᠑᠑	᠑᠕	᠑᠑	᠑᠕	᠑᠑	᠑᠕	᠑᠑	᠕᠒
᠕᠑	᠕᠎	᠕᠑	᠕᠕	᠕᠑	᠕᠕	᠕᠑	᠕᠕	᠕᠑	᠑᠒
᠑᠑	᠑᠎	᠑᠑	᠑᠕	᠑᠑	᠑᠕	᠑᠑	᠑᠕	᠑᠑	᠕᠒

Figure 14: Coronavirus leaflet in Muishvung Tangsa using the Lakhum script, April 2020

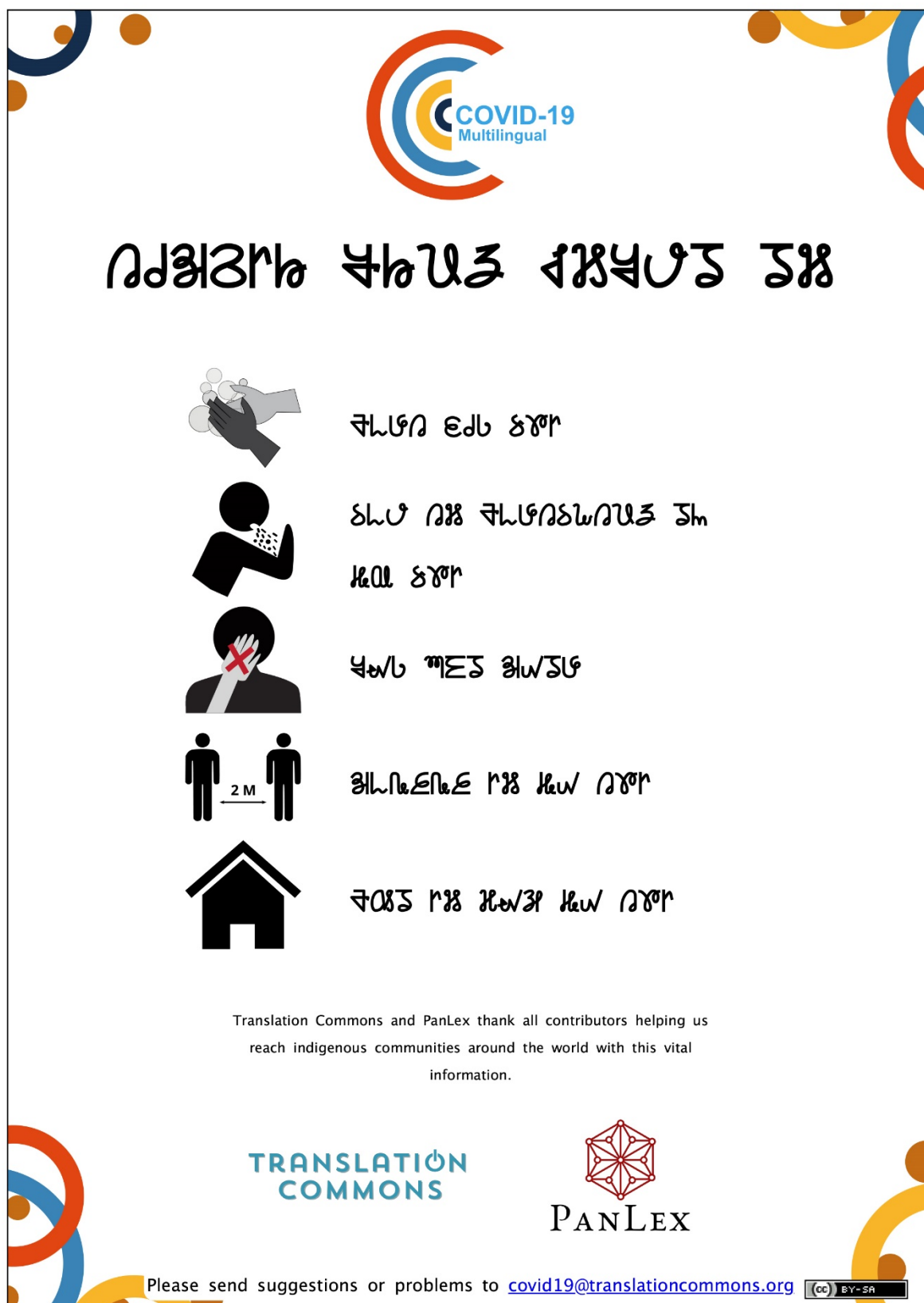


Figure 15: Draft primer being prepared in September 2020 (front cover, draft pages 1, 2 and 9_15.1 cover

ԼՎԻՈՒՅՑ ԵՃՈՁԷԼՈՊ



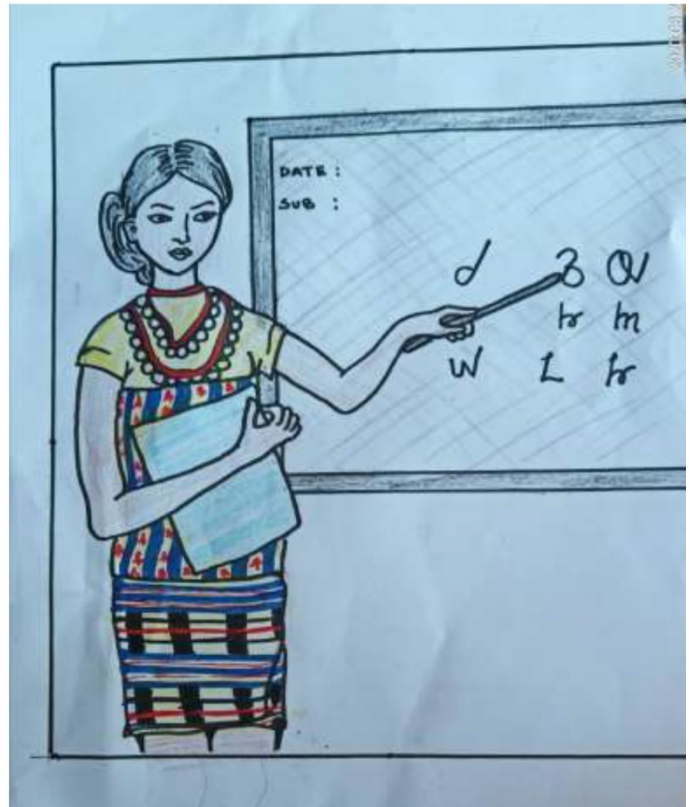
ବିଷୟ ୯ / Lesson 1

ବିଷୟ / Vowels

ଡ ଡ ଡ ଡ ଡ ଡ ଡ
 ଡ ଡ ଡ ଡ ଡ ଡ ଡ
 ଡ ଡ ଡ ଡ ଡ ଡ ଡ
 ଡ ଡ ଡ ଡ ଡ ଡ ଡ
 ଡ ଡ ଡ ଡ ଡ ଡ ଡ
 ଡ ଡ ଡ ଡ ଡ ଡ ଡ

ବିଷୟ / Consonants

ଡ ଡ ଡ ଡ ଡ ଡ ଡ
 ଡ ଡ ଡ ଡ ଡ ଡ ଡ
 ଡ ଡ ଡ ଡ ଡ ଡ ଡ
 ଡ ଡ ଡ ଡ ଡ ଡ ଡ



ᑭᓂᓂᓂᓂ ᑭ / Lesson 2

ᑭᓂᓂᓂᓂᓂ ᑭᓂᓂᓂᓂᓂ/ Word forming (Unitelligible)

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ

ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ ᓂᓂ



ᑭᑭᑭᑭᑭ ᑭ / Lesson 4

ᑭᑭᑭ ᑭᑭᑭᑭᑭᑭ / Learning through Pictures

ᑭ

ᑭᑭᑭ, ᑭᑭᑭ

Kaan, kin

Mountain



ᑭ

ᑭᑭᑭ

Khel

Goat



ᑭ

ᑭᑭᑭᑭ

Gehay

Dog



Figure 16 Submission of draft Primer (PHVUNGCWVNC meaning 'Starting') to Mr Tapi Gao, Director of Elementary Education, Government of Arunachal Pradesh, November 2020. The primer was presented by two members of the Tangsa Script Development committee, Mr Sengkhum Muixshvungx and, standing at the back, Mr Nanman Jugli.



Figure 17: Selection of pages of the Primer as presented to Mr Tapi Gao, Director of Elementary Education, Government of Arunachal Pradesh, November 2020

འཕྲིན་ལྷན་ རྒྱུ་འཕྲིན་ལྷན་

(Text Book)

In Tangsa



ᐱᓄᓂᓂᓂ ᠑ - ᑭᓄᓂᓂᓂᓂᓂ (Lesson 1 – ALPHABETS)

With English alphabets to mark tones and sounds.

ᑭᓄᓂᓂᓂ (VOWELS)

Z – low, c – high, Q – glottal stop, x – middle

ᐅ Oz	ᐅ oc	ᐅ Oq	ᐅ Ox	ᐅ Az	ᐅ ac	ᐅ aq	ᐅ ax
ᐅ vz	ᐅ vc	ᐅ Vq	ᐅ Vx	ᐅ Ez	ᐅ ec	ᐅ eq	ᐅ ex
ᐅ iz	ᐅ ic	ᐅ Iq	ᐅ Ix	ᐅ Uz	ᐅ uc	ᐅ uq	ᐅ ux
ᐅ awz	ᐅ awc	ᐅ Awq	ᐅ Awx	ᐅ Uiz	ᐅ uic	ᐅ uiq	ᐅ uix
ᐅ eng	ᐅ uexx	ᐅ Uezz	ᐅ Awxx	ᐅ Uec	ᐅ uez	ᐅ ueq	ᐅ uex
ᐅ uiuz	ᐅ uiuc	ᐅ Uiuq	ᐅ Uiux	ᐅ Mmz	ᐅ mmc	ᐅ mmq	ᐅ mmx

ᑭᓄᓂᓂᓂ (CONSONANTS)

ᐅ Ko	ᐅ kho	ᐅ Go	ᐅ Ngo	ᐅ So	ᐅ yo	ᐅ wo	ᐅ Po
ᐅ nyo	ᐅ pho	ᐅ Bo	ᐅ Mo	ᐅ No	ᐅ ho	ᐅ lo	ᐅ hto
ᐅ To	ᐅ do	ᐅ Ro	ᐅ Nho	ᐅ Sho	ᐅ jo	ᐅ tso	ᐅ gho
ᐅ htto	ᐅ tho	ᐅ Qho	ᐅ Fa	ᐅ Dha	ᐅ cha	ᐅ yha	

ବିଷୟ ୩ - ବର୍ଷର ୬ ମାସ
(Lesson 3- Months of a year)

ဒါနာအား ခံယူပါ။

ଅଧ୍ୟାୟ ୪ ଓ ୫ ର ୧୫ ଓ ୧୬ ଶ୍ରମିକଙ୍କୁ ଅନ୍ୟତମ ଶ୍ରମିକଙ୍କ ସହ ସମାନ କରିବା ପାଇଁ.

(There are twelve months in a year)

ਚੰਡੀਗੜ੍ਹ(Yacremz).....January

ꠘꠞꠘꠞꠘ (Yacngaz).....February

ᐱᐱᐱᐱᐱᐱ (Rvwud).....March

୧୫୫୩୫ (Wvrac).....April

ચેહઠી(Yacshix).....May

ਚੰਨੌੜ(Yackuiz).....June

ਚੇਕੀਐਕਸ(Yacjuimx).....July

ચિહ્નક્રમ(Yacshulx).....August

ચેહઠીય(Yacshax).....September

ᠠᠨᠠᠵᠢᠨᠠᠵᠢᠨ(Wudwuz).....October

ਚੰਡੀਚੌਰੀ(Yacdac).....November

ਚੰਡੀਗੜ੍ਹ (Yacdvnz)December

[illegible]

(The months of a year is divided into 4 parts/sections-)

9. ବିକଳତା - ବିକଳିତ ବ୍ୟକ୍ତିର ଅବସ୍ଥା

h. ବିଦ୍ୟାଳୟ - ଗଣିତର ଚିହ୍ନ ଚିହ୍ନିତ

୧.୬୧୪୩୩- ଚାନ୍ଦିଓଡ଼ ଚାନ୍ଦିଓଡ଼ ଚାନ୍ଦିଓଡ଼

Ի.ԵԶԵՈ- ԴՐՉԴՐՉ ՉԻՉԻ ՉԻՉԻ

ԳԼԵՍԻՆ Կ - ԿԵՊԵԼՈՒՄ
(Lesson 4 - Days of a Week)

ՃՅՈՒՇԱՆ(shovncshalx).....Monday

ՄԻՈՒՇԱՆ(mancshalx).....Tuesday

ՓԻՈՒՇԱՆ(phancshalx).....Wednesday

ԿԻՈՒՇԱՆ(htancshalx).....Thursday

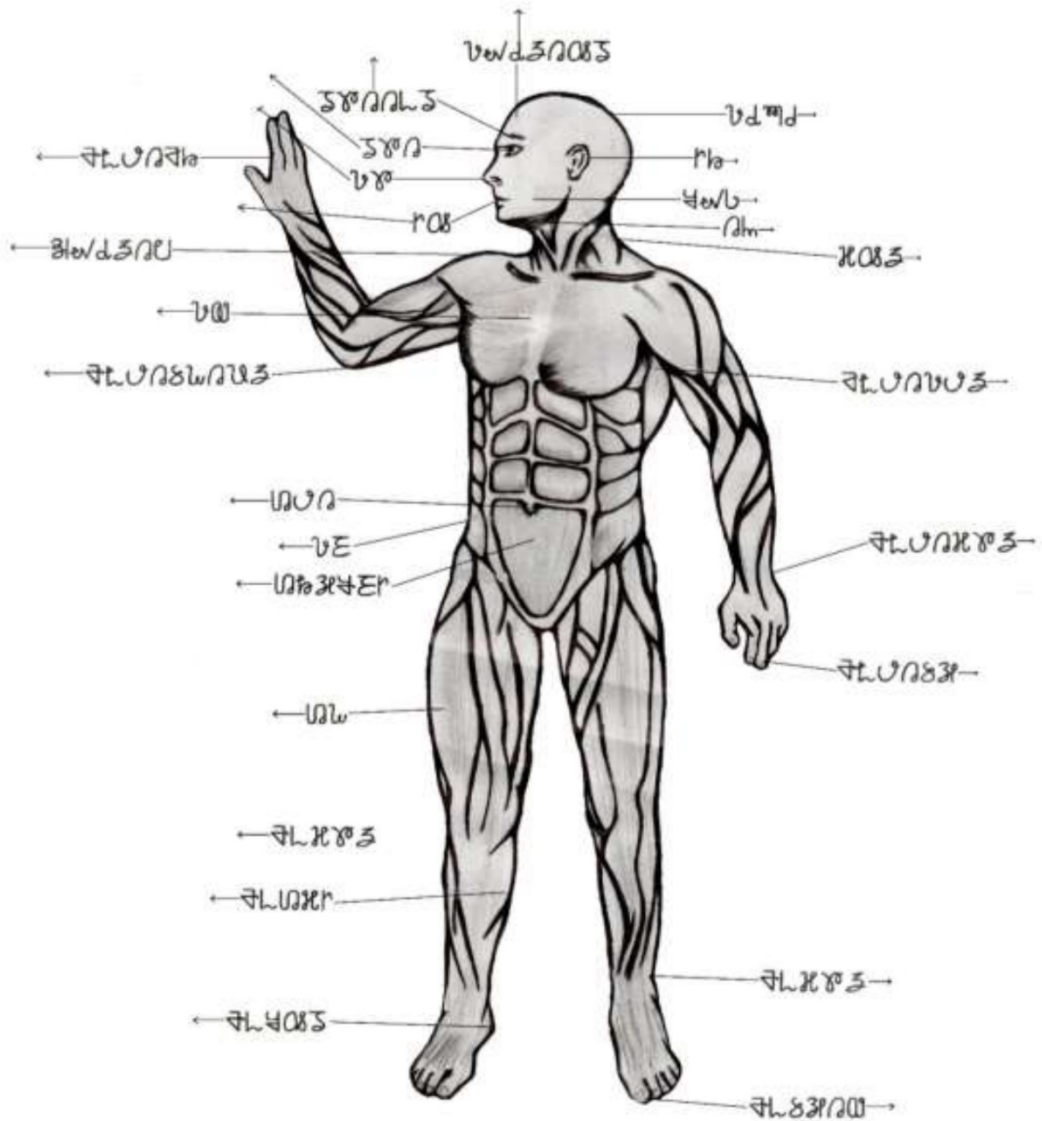
ՍԻՈՒՇԱՆ(wancshalx).....Friday

ՆԻՈՒՇԱՆ(ngancshalx).....Saturday

ՊԵՅՅՈՒՇԱՆ(peocshalx).....Sunday



48603 2 - 53643 503 48606
(Lesson 5 - Naming the Human Body parts)



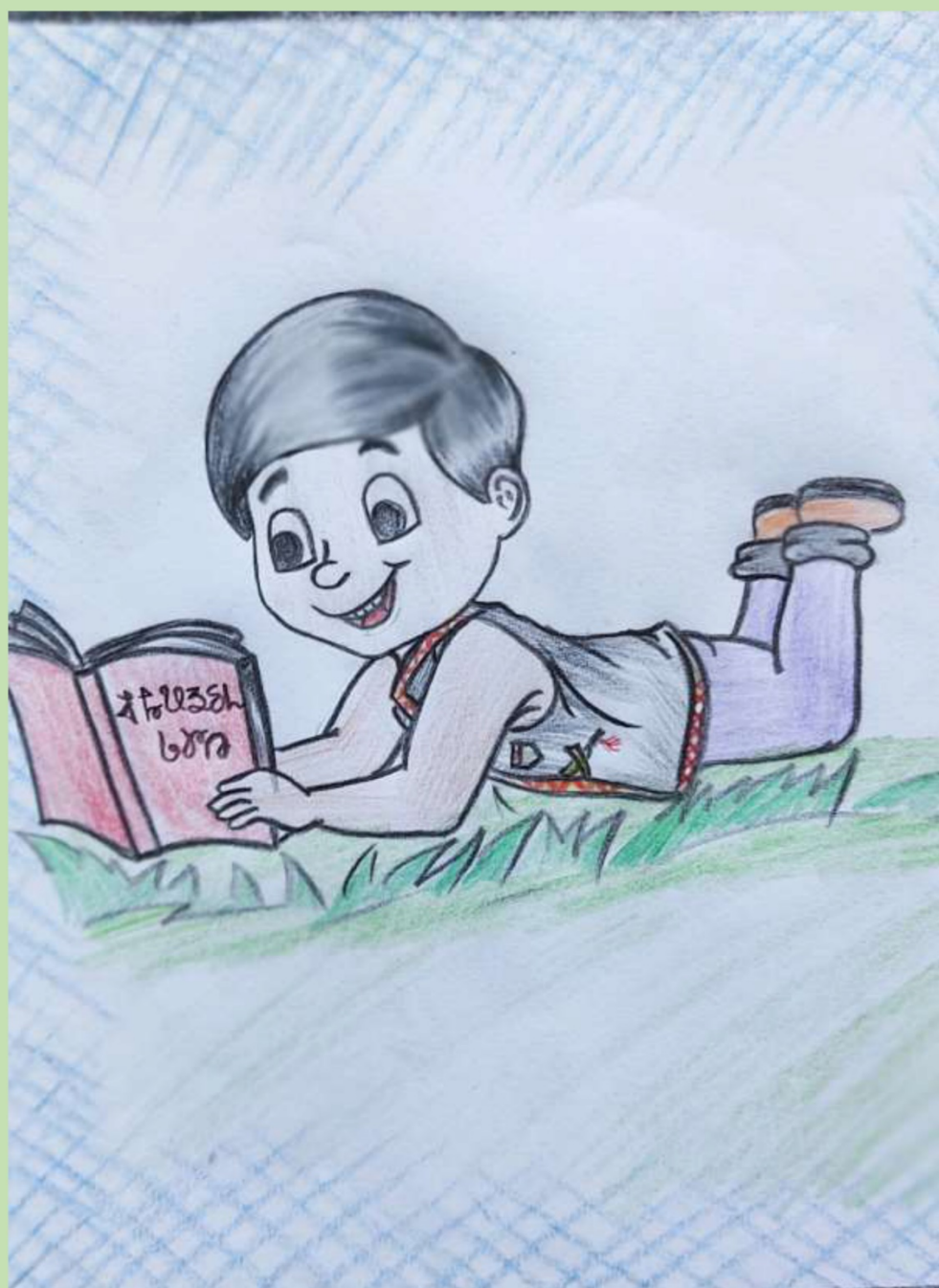


Figure 18: Sample from Rev. Gam Win's Primer for the Muishvung (Mossang) language.

Bomc Dunge

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

Mungkawx luik vlvix linguistic tuiuz thenc jix
rueq luik thenc wvnc juingz vyoalx phvnx maq
vtvungx 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 vtdungz 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 kvanc khvnc maq dvix shex kuex
rvmuq, tvrvp htawc wvnc yek juingz nuiyc tvhvlc
tvhungx yek juingz muinz httax yaq doc. Naungshiq
lotlaz honcshiz thuinshiz maq raz tueq, vhlx
pinhtumc tvhungx htaungx wvnc tvkoc mkaq muinz
tvluime kvmluimx 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	

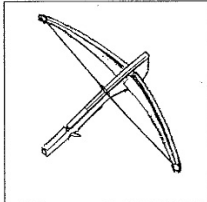
Vthuic

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

Figure 19. Photo of Primary School students learning Tangsa Script, January 2021



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, who 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. Mr Wanglung Mossang has reviewed multiple versions of this application and provided considerable additional information on request throughout the process of developing this proposal.

The initial font was developed by Karen Parker. Subsequent improvements to the font were made by Kellen Parker van Dam, who has also contributed to the application in discussions with the Tangsa Script Development Committee members.

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 10646²

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:	<i>Tangsa</i>	
2. Requester's name:	<i>Stephen Morey</i>	
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual contribution</i>	
4. Submission date:	<i>2020</i>	
5. Requester's reference (if applicable):	<i>?</i>	
6. Choose one of the following:		
This is a complete proposal:		<i>X</i>
(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):		<i>X</i>
Proposed name of script:	<i>Tangsa</i>	
b. The proposal is for addition of character(s) to an existing block:		
Name of the existing block:		
2. Number of characters in proposal:		<i>89</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary <i>?</i>	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?		<i>yes</i>
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?		<i>yes</i>
b. Are the character shapes attached in a legible form suitable for review?		<i>yes</i>
5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?	<i>Stephen Morey and Kellen Parker van Dam</i>	
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?		<i>yes</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>yes</i>	
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)?		<i>no</i>

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?	<input checked="" type="checkbox"/> yes
If YES explain	<i>Introductory proposal by Anshuman Pandey L2/13-231.</i>
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.)?	<input checked="" type="checkbox"/> yes
If YES, with whom?	<i>Tangsa Script Development Committee in India</i>
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?	<input checked="" type="checkbox"/> yes
Reference:	<i>this document</i>
4. The context of use for the proposed characters (type of use; common or rare)	<i>rare, but developing</i>
Reference:	
5. Are the proposed characters in current use by the user community?	<input checked="" type="checkbox"/> 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?	<input checked="" type="checkbox"/> 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)?	<input checked="" type="checkbox"/> yes
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	<input checked="" type="checkbox"/> 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?	<input checked="" type="checkbox"/> 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?	<input checked="" type="checkbox"/> 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?	<input checked="" type="checkbox"/> yes
If YES, is a rationale for such use provided?	<input checked="" type="checkbox"/> yes
If YES, reference:	<i>this document</i>
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	<input checked="" type="checkbox"/> no
If YES, reference:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	<input checked="" type="checkbox"/> no
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility character(s)?	<input checked="" type="checkbox"/> no
If YES, is the equivalent corresponding unified ideographic character(s) identified?	
If YES, reference:	