

Positioning of Tai Tham Vowels Below

Author: Richard Wordingham

Date: 29 April 2018

Table of Contents

Summary.....	1
Description.....	1
Evidence.....	4
Acknowledgements.....	9
References.....	10

Summary

The position of Tai Tham vowels below in orthographic syllables with two or more consonants is variable. While they are normally positioned below the Roman baseline, U+1A69 TAI THAM SIGN U and U+1A6A TAI THAM SIGN UU may be rendered as spacing glyphs to the right of the base character (as seen for example in Figure 10). Whether this happens is a stylistic decision.

Description

The Tai Tham script has four characters or near characters that may reasonably be considered as vowels below, shown on 𐞀 U+1A20 TAI THAM LETTER HIGH KA in the list below:

𐞁 U+1A69 TAI THAM VOWEL SIGN U

𐞂 U+1A6A TAI THAM VOWEL SIGN UU

𐞃 U+1A6C TAI THAM VOWEL SIGN OA BELOW

𐞄 <U+1A60 TAI THAM SIGN SAKOT, U+1A45 TAI THAM LETTER WA>

The first three are classified as single characters in Unicode, and accordingly have an Indic Positional Category; they have the non-trivial value of Bottom. The property does not extend to sequences. The first two occur in both Pali and the Tai languages employing the script and the latter two occur in the Tai languages but not in Pali. The sequence <SAKOT, WA> may also represent a medial consonant, and should always be treated as such for the placement of tone

marks. In this rôle it occurs in both Tai and Pali, in the latter most commonly in the absolutive suffix -tvā. (There is also a widespread view that SIGN OA BELOW is simply a subscript form of U+1A4B TAI THAM LETTER A.)

When a subscript consonant does not extend to the right, with a tail extending to the hanging baseline, the simplest treatment of subscript consonant and the vowel symbol below is for them to form a vertical stack. This is a common treatment in Northern Thai fonts and is also seen in hand-written text. For example, in Figure 3 there is ຂຸ in the middle of the second phrase (of three) on

line 1, Pali ຍາດຣີຣິຕິ on line 2 and ຈຸ in the middle of the final phrase on line 3. Similarly, there

is ທູ້ at the end of line 2 of the left hand column of Figure 6. The disadvantage of this approach is that the descending stack may clash with ascending marks above in the line below, especially in the Tai languages. The natural solution is to increase the line-spacing, but this wastes writing medium, running contrary to techniques such as condensing Pali words by stacking syllables, such as condensing *vedanā* as ດອຣ် rather than writing plain ດອຣ (which paradoxically can now merely have the effect of merely adding 2 code points to the backing store).

Now, especially in native Tai words, an intervening subscript may, possibly optionally, have a relatively small depth (NGA, BA, LOW YA, MEDIAL LA, LA) or may be able to enfold the vowel below. Many fonts have a small subscript WA to assist with this vertical condensation. This preserves the structure of a vertical stack, albeit at the possible cost of additional contextual forms or ligatures in a font, while not requiring any greater depth than most Pali subscripts. Examples of enfoldings include the word ທູ້ <HIGH HA, SAKOT, MA, SIGN UU, TONE-1> appearing as the second syllable of

the right hand column of the first line of Figure 7 and the word ທູ້ <HIGH HA, SAKOT, MA, OA BELOW, MAI KANG, TONE-2> in the middle of the right hand columns of both the second and third lines shown in Figure 7. Similar enfolding or condensation can be seen in the Tai Khün example in Figure 2, where the renderings of ທູ້ in the middle of the final phrase on line 1 and of ພູ້ at the end of the sixth line from the bottom respect the vertical boundaries of the orthographic syllables.

An alternative technique is to write the subscript consonants and vowels below side-by-side. A problem of this technique, more so for fonts than for human writers, is that the space below the next base character may already be occupied. This problem does not apply to modern Tai Khün SIGN OA BELOW, for it now only occurs in closed syllables. Accordingly, the coda consonant, which is written as a base consonant, will reserve space for the sideways-displaced vowel. (It is just conceivable that foreign loan words might defeat the system.) Examples given here are:

Word	Figure	Location	Language
သမ္မုဇ္ဈ	1	Line 1, 3 rd phrase	Pali

Word	Figure	Location	Language
ငွေ	4	Line 2, right column, first word	Northern Thai
ညွှတ်	4	Line 5, right column, third word	Northern Thai
ညွှ	5	Line 6, middle of right column	Northern Thai
ညွှ	6	Line 2, penultimate syllable of left column	Northern Thai
ညွှ	2	Line 12, second syllable	Tai Khün
ညွှ	10	End of line 2	Tai Khün

This method may overlap with the enfolding technique, as in ညွှ at the start of the second phrase on line 1 in Figure 3.

The final solution in Northern Thai is to transfer a subscript consonant to the baseline. Thus instead of writing fully assimilated /hu:p/ as ညွှ, in Northern Thai one may write it as ညွှပ, which is completely ambiguous with the stem form of its source, Pali *rūpa*. This entails a change of coding, for the difference between subscript and non-subscript usually indicates a difference in pronunciation, and therefore the choice cannot be left to the rendering system.

Another solution, widely adopted in Tai Khün, is to move the vowel to the right. For vowels, there is no evidence that the rendering choice for a given style depends on any more than the sequence of current Unicode characters. An example from the Tai Tham script's encoding proposal is provided by the word ငွေ (ငွေ in a Tai Khün font) in the middle of the third phrase of line 7 of Figure 2. An

example in a Pali stack is provided by the word ဘိက္ခု 'bhikkhu' (Tai Khün rendering ဘိက္ခု) in the second phrase on line 1 of Figure 10. In combination with a tone mark, another example can be seen in the third line from the end in the same figure in the form of the word ဟုံဟုံ 'village'. The same font in the same book is also used to write ဟုံဟုံ with a non-spacing SIGN UU, as can be seen in the last third of the middle line of Figure 9. (The published text predates the encoding of the script.) For SIGN U and SIGN UU, the vowel in a closed syllable can be written on the right, though in current practice this seems to be restricted to before final, subscript BA and before final, subscript LOW YA. Thus, in Figure 8, we have ညွှ /hu:p/ (Tai Khün rendering ညွှ) on lines 1 and 3,

but လူက /lu:k/ at the end of line 2. For a renderer, the choice of form in a closed syllable is constrained by the presence or absence of U+1A60 TAI THAM SAKOT, which determines whether the final consonant is written as a subscript or as a base form. By contrast, the choice of a non-spacing or a spacing form of the vowel in an open syllable is up to the renderer, though I would suggest that ZWJ and ZWNJ could be used to influence the decision when either might be present.

Evidence

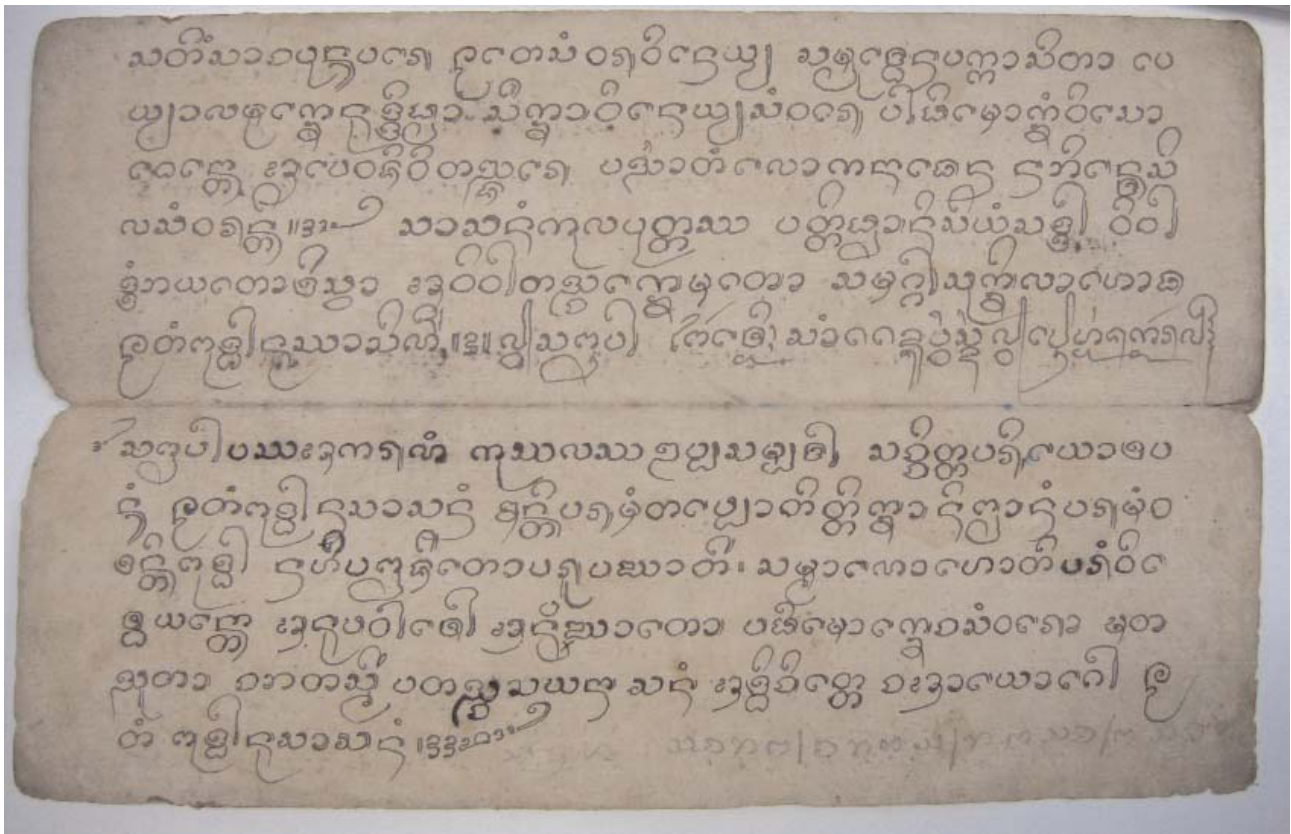


Figure 1 Pali manuscript from northern Thailand

The figure above is Figure 9a in Reference 1.

Shows side-by-side subscript consonant and vowel on line 1.

ခံခဲ့ရကုန်ဖို့၊ ဂိဇ္ဈသျှိသျှိန

ရေခဲသေတ္တာ (ပိုက်လင်း)

(လေခေတ္တ)

[illegible]

ကမ္ဘာ့ကုမ္ပဏီ ဂရင်းဂရိပ်ကသာသေ့

ရက်စွဲ ၂၁.၁၀.၂၀၀၂

Figure 2 Sample from a reader in Tai Khün

The figure above is Figure 11 in Reference 1.

Shows: Subscript consonant enfolding vowel on line 1.

Spacing SIGN U on line 7

Subscript consonant and SIGN OA BELOW side-by-side on line 12

Condensed subscript and consonant on sixth line from bottom

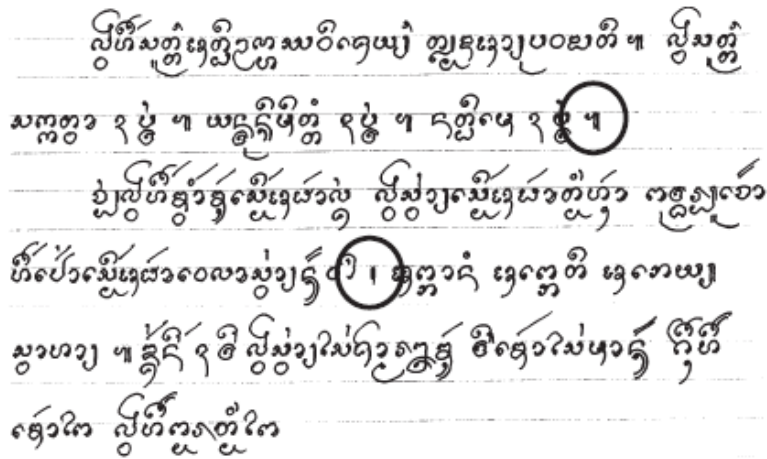


Figure 3 Sample from a reader in Northern Thai

The figure above is Figure 4 in Reference 1.

Shows 3-element vertical stacks on lines 1, 2 and 3.

ascending subscript consonant partially enfolding vowel and displacing it to beneath next base consonant on line 1

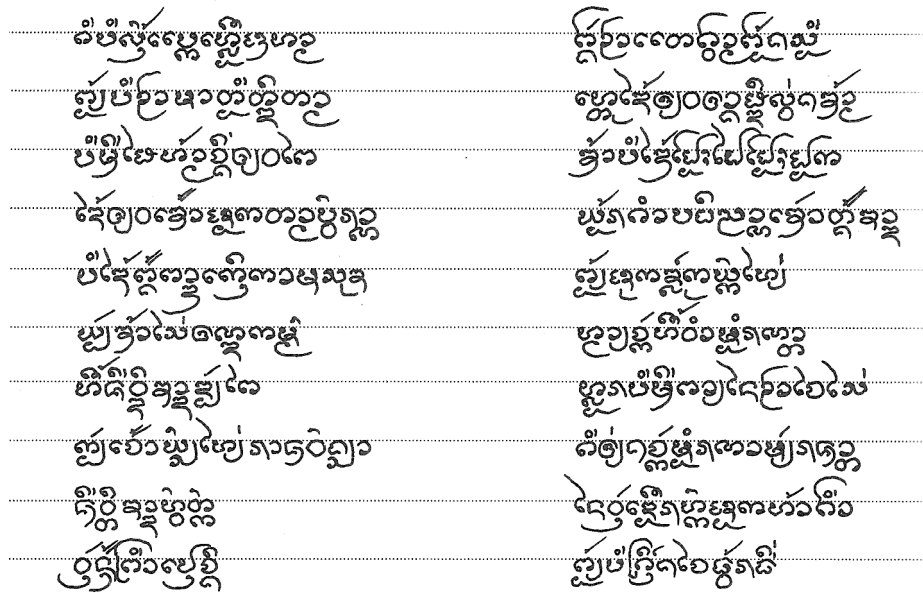


Figure 4 Northern Thai reading exercise: first 10 lines of p. 280 of Reference 3

Shows 4-level stack (1 above, 2 below) on line 1

subscript consonant and vowel side-by-side on lines 2 and 5.

3-level stack (2 below) on line 7

ຂົວຜາຂຽວໂງດວຣັດ	ສັດຂຽວກະເກ່ຍສັດຂຽວ
ປູນຊີ່ກັວນຸ່ງທູດເວ	ທູດກີ່ວເວຍກະເກ່ຍກະເກ່ຍ
ລຸກກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ນຸ່ງກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ

Figure 5 Northern Thai reading exercise: first 10 lines of p. 281 of Reference 3

Shows subscript consonant enfolding vowel on lines 1 and 2

subscript consonant and vowel side-by-side on line 6

ຂົວຜາຂຽວໂງດວຣັດ	ສັດຂຽວກະເກ່ຍສັດຂຽວ
ປູນຊີ່ກັວນຸ່ງທູດເວ	ທູດກີ່ວເວຍກະເກ່ຍກະເກ່ຍ
ລຸກກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ນຸ່ງກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ
ກະເກ່ຍກະເກ່ຍກະເກ່ຍ	ກະເກ່ຍກະເກ່ຍກະເກ່ຍ

Figure 6 Northern Thai reading exercise: first 10 lines of p. 282 of Reference 3

Shows Subscript consonant and vowel side-by-side on line 2

2 3-level stacks (2 levels below) on line 2 (2 examples)

သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ	သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ	သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ	သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ	သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ	သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ	သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ	သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ	သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ	သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ	သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ

Figure 7 Northern Thai reading exercise: last 10 lines of p. 310 of Reference 3

Shows subscript consonant enfolding vowel on lines 1, 2 and 3.

သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ

Figure 8 Tai Khün text: paragraph from p. 250 of Reference 2

Shows spacing SIGN UU with subscript coda consonant on lines 1 and 3, but non-spacing SIGN UU with base consonant serving as coda consonant on line 2.

သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ
သံဃိဗုဒ္ဓဗျာဓိဗုဒ္ဓဗျာဓိ

Figure 9 Tai Khün text: bottom of p. 257 of Reference 2

Shows subscript consonant enfolding SIGN U on line 3.

လွဲဖိဇွဲကမာသျှ လါကတိကျသံသယသေတုကယွဲဝဲရဟံဆာ
ဟါပာသဉ္စပေဖေဝဲဒဲကန္တဗေသဇဝဇ္ဇိမာဝ်သျှကျဉ္စ

ကန္တထံကကျသံသယသေတုကယွဲဝဲရဟံဆာ

ကန္တရုကဟျဉ္စသွမဝဲသံသယသေတုကယွဲဝဲရဟံဆာ
ဣန္ဒြေဇ္ဇိဝဲကန္တ န္တရုသိယုကတံဃာကွဲ သွမဝဲသံသယသေတု
ယဉ္စတံသျှပေသျှသျှဝဲ န္တဟွဲသျှပျာသျှသံသယသေတု
ရုမာသိသွမဝဲသံသယသေတု ဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှ
လါဣန္ဒြေဇ္ဇိဝဲသေတုကန္တဟွဲဝဲကွဲသျှသျှသံသယသေတု
ပဝဲဝဲဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှသျှသံသယသေတု
ဧဝဲဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှသျှသံသယသေတု
ကန္တရုကဟျဉ္စသွမဝဲသံသယသေတုကယွဲဝဲရဟံဆာ
မု နသျှပဝဲကတံဃာကွဲသျှ ဧဝဲဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှ
ပဝဲဝဲဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှသျှသံသယသေတု

ဝဲဝဲဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှသျှသံသယသေတု
ဝဲဝဲဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှသျှသံသယသေတု
ဧဝဲဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှသျှသံသယသေတု
တုကဟွဲဝဲကွဲသျှသျှသံသယသေတု
ဧဝဲဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှသျှသံသယသေတု
ဧဝဲဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှသျှသံသယသေတု
ဧဝဲဣန္ဒြေဇ္ဇိဝဲကန္တဟွဲဝဲကွဲသျှသျှသံသယသေတု

Figure 10 Tai Khün text: p. 245 of Reference 2

Shows: Spacing SIGN U on line 1.

SIGN OA BELOW at bottom right on line 2

Spacing SIGN UU on 3rd line from bottom.

Acknowledgements

Figures 1, 3 and 2 are taken from Reference 1.

The glyphs for most of the Tai Tham in the text are taken from the *Hariphunchai* font by Ed Trager, a font targeted at Northern Thai. The Tai Khün font used is *A Tai Tham KH* by Sangdang Kengtung. Both fonts are available under the SIL Open Font License.

References

1. Michael Everson, Martin Hosken, and Peter Constable, *Revised proposal for encoding the Lanna script in the BMP of the UCS*, ISO/IEC JTC1/SC2/WG2, L2/07-007R, <http://www.unicode.org/L2/L2007/07007r-n3207r-lanna.pdf> (March 2007)
2. *Chiang Tung: Its Way of Life* ຄຊຸຊຍຸດວຸຈຍຸດວຸດ [Khemarattha Nakon Cheng Tung] เขมรัตน์นครดียงตุง [Khemarat Nakhon Chiang Tung] (in Thai, Tai Khün, French and English) (Chiang Mai: Wat Tha Kradas วัดท่ากระดาส) (1998)
3. Watcharasat, Bunkhit (2003). *แบบเรียนชุดพื้นฐาน เน้นกระบวนการเรียนรู้ ภาษาเมืองลัณนา, Basic Course for Learning the Lanna Language* (Revision 1). (Chiang Mai: Thara Thong Printing ธาราทองการพิมพ์). ISBN 974-85472-0-5