

Title: Final Proposal to encode the Tai Yo Script
Author: Viet Khoi Nguyen
 Cong Danh Sam
 Frank van de Kasteelen

Date: 16th December 2022

1. Introduction

The *Tai Yo* language (ISO 639-3: **tyj**), usually mentioned as *Tay Dọ* or *Thái Yo* in Vietnamese, is one of the Tai languages spoken in some districts of two provinces Nghệ An and Thanh Hóa, Vietnam. There are also some Tai Yo varieties spoken in Laos by communities known as Tai Maen and Tai Pao. The *Tai Yo* language belongs to the Southwestern Tai group which includes Siamese (Central Thai), Lao, Tai Dam, Tai Don, and its geographically-related Tai Daeng and Tai Meuyay.

Native speakers of the Tai Yo language are called Tai Yo people, a subgroup of Thái ethnic group under Vietnam's classification. The Tai Yo people in fact have some other ethnonyms such as Tay Mương, Tay Hàng Tổng, and Tay Chiềng/Xiềng (only in Nghệ An). An estimated number of 292.000 native Tai Yo speakers (2019) in Nghệ An and Thanh Hóa is reported¹.

The script is natively called *Lai Tay* (ᨲᨶᨶᨶ ᨶᨶ), literally meaning ‘Tai script’. It is one of two scripts used for Tai Yo, the other natively being called *Lay Pao*, which means ‘Pao script’. For the sake of consistency with other Tai writing systems, such as Tai Tham, Tai Viet, Tai Le etc., the script in this proposal will be called *Tai Yo*. None of these scripts use their native names, which all mean ‘Tai script’. *Tai Yo* is also the name that has traditionally been used in western scholarship about the script. We do request that an annotation with the native name be made in the code charts.

The other script, for which a proposal is forthcoming, will similarly be called *Tai Pao*.

2. Usage

The *Tai Yo* script was traditionally used by some Tai Yo communities in Thanh Hóa province (Thường Xuân and Như Xuân districts, formerly belonging to an upland district called *Châu Thường Xuân*) and in Nghệ An province (Quế Phong, Quỳnh Châu, Quỳnh Hợp, Tân Kỳ districts, belonging to a former prefecture called *Phủ Quỳnh Châu*). Nowadays, in Nghệ An province, *Tai Yo* manuscripts are preserved in some personal collections and public places. Besides, teaching and learning activities of the *Tai Yo* script are proceeding with the support of the local government and teachers.

3. Writing system

The *Tai Yo* script remains relatively close to its neighboring scripts Thai, Lao and Tai Viet. There are three features, however, that structurally set it apart. The most obvious of these features is that

¹ Sầm Công Danh (2021: 30)

Tai Yo is written from top to bottom in columns running from right to left in a way similar to, and clearly influenced by, Chinese. In fact, many of the older manuscripts have their titles written in Chinese, and gloss uncommon words with the related Chinese character. The second differentiating feature is that in *Tai Yo*, the vowel is always written in phonetic order, following the consonant it modifies. There are five signs which do occur to the right of their associated consonant, but a vowel sign never precedes or circumvents its consonant as is common in other Tai scripts.

Thirdly, the *Tai Yo* script remains a true abugida, meaning that a bare consonant is pronounced as a syllable with the inherent vowel <o>, [ɔ], where Thai and Lao, for example, now explicitly mark this vowel and have thus developed more into alphabets.

Fourthly, the *Tai Yo* language is tonal, similar to other Tai languages. However, like other traditional scripts for Southwestern Tai languages, the traditional *Tai Yo* script does not use tone marks. Tone indicators are developed in modernized scripts of some Tai languages such as Lao, Black Tai, White Tai (sharing the Tai Viet Unicode block).

Especially the fact that vowels in general follow their associated consonant makes that the script is relatively straightforward to encode, requiring very little in terms of complex rendering. There are 29 consonants, of which only 8 occur in syllable final position. The script has 13 vowels, one of which is placed to the right of the consonant in an open syllable, but follows it in a closed syllable. There are 9 signs which represent a vowel-final consonant combination. Of those 4 are placed to the right of the consonant, one of which doubles as a medial glide [-w-]. Finally, there are 2 to 4 other signs, which represent iteration or whole words.

Throughout the proposal, whenever a single Tai Yo letter is shown embedded in Latin text, it is shown on its side in horizontal layout. Whenever characters are shown in a table, they are displayed in their vertical layout.

3.1. Consonants

Similar to other Tai languages in Vietnam, initial consonants of low class in *Tai Yo* are equivalent to voiceless consonants in Proto-Tai and those of high class are equivalent to voiced consonants in Proto-Tai. In the phonology of the modern *Tai Yo* language, most low-high consonant phonemic pairs are merged into one pronunciation attached with different tones.

3.1.1. Initial consonants

The following consonants are used for the onset:

Name	Glyph	IPA	Tai Dam equivalent	Siamese equivalent	Sukhothai equivalent
LETTER LOW KO	ᵿ	/k/	ᵿ	ก	ᵿ
LETTER HIGH KO	ᵿ		ᵿ	ค	ᵿ

LETTER LOW KHO	ꨀ	/k ^h /	ꨁ	ꨂ	ꨃ
LETTER HIGH KHO ²	ꨄ				
LETTER GO	ꨅ	/g/	ꨆ		
LETTER NGO	ꨇ	/ŋ/	ꨈ	ꨉ	ꨊ
LETTER CO	ꨋ	/c/	ꨌ	ꨍ	ꨎ
LETTER LOW XO	ꨏ	/s/	ꨐ	ꨑ	ꨒ
LETTER HIGH XO	ꨓ		ꨔ	ꨕ	ꨖ
LETTER LOW NYO	ꨗ	/ɲ/			
LETTER HIGH NYO	ꨙ		ꨚ	ꨛ	ꨜ
LETTER DO	ꨟ	/d/	ꨠ	ꨡ	ꨢ
LETTER LOW TO	ꨣ	/t/	ꨤ?	ꨥ/ꨦ?	ꨧ/ꨨ?
LETTER HIGH TO	ꨩ		ꨪ?	ꨫ/ꨬ?	ꨭ/ꨮ?
LETTER THO	ꨰ	/t ^h /	ꨱ	ꨲ	ꨳ
LETTER NO	ꨴ	/n/	ꨵ	ꨶ	꨷
LETTER BO	꨸	/b/	꨹	꨺	꨻
LETTER LOW PO	꨼	/p/	꨽	꨾	꨿
LETTER HIGH PO	꨿		ꯀ?	ꯁ/ꯂ?	ꯃ/ꯄ?
LETTER PHO	ꯆ	/p ^h /, /f ³ /	(ꯇ)	ꯈ	ꯉ
LETTER LOW FO	ꯊ	/f/	ꯋ	ꯌ	ꯍ
LETTER HIGH FO	ꯎ		ꯏ	ꯐ	ꯑ
LETTER MO	ꯓ	/m/	ꯔ	ꯕ	ꯖ
LETTER YO	ꯙ	/j/	ꯚ, ꯛ	ꯜ	ꯝ
LETTER LO	ꯟ	/l/	ꯠ	ꯡ	ꯢ

² Letter HIGH KHO is found only in modern teaching materials.

³ Some dialects of Tai Yo language have no distinction between /p^h/ and /f/, thus in these dialects, this LOW PH is used as LOW F.

LETTER VO	ᨆ	/v/	ᨁ	ᨂ	ᨃ
LETTER LOW HO	ᨄ	/h/	ᨅ	ᨆ	ᨇ
LETTER HIGH HO	ᨈ		ᨉ	ᨊ	ᨋ
LETTER QO	ᨌ	/ʔ/	ᨍ	ᨎ	ᨏ
LETTER LOW KVO	ᨐ	/kw/	ᨑ	ᨒ	ᨓ
LETTER HIGH KVO	ᨔ		ᨕ	ᨖ	ᨗ

The following initial consonants are missing in old manuscripts but are mentioned in one modern teaching material (**Figure 13**), which lists them but gives no examples of their use:

Name	Glyph	IPA	Alternative
LETTER SO	ᨚ	/s/	ᨛ
LETTER RO	᨜	/z/	᨝
LETTER TRO	᨟	/t/	ᨠ/ᨡ

The phonemes these letter represent are not native in Tai Yo and found in loanwords from Vietnamese. However, phoneme /s/ is found in *Sàm* (a popular surname of Tai Yo people) and was frequently spelled with LETTER THO (ᨆ) in some old materials (**Figure 14**). /z/ and /t/ sometimes merge into /d/ and /t/ respectively by some Tai Yo speakers. These phonemes are expected to be alternatively represented by DO as RO and HIGH or LOW TO as TRO, but as far as we know, these phonemes are not found in known old manuscripts. Since their actual use is still to be proven, they are for now not included in this proposal, but space will be left in the table for addition at a later time.

The letter HIGH KHO is similarly a recent addition to the *Tai Yo* script. This letter is used for a phoneme native to Tai Yo, although it's quite rare. Since this glyph has no alternative, it has been included in this proposal. The fact that it's missing from the older manuscripts is mainly due to the fact that it's quite rare in and occurs in vocabulary not related to the mainly religious context of these manuscripts. Additionally, there are other rare phonemes native to Tai Yo, which are HIGH PHO, HIGH THO, HIGH CO, HIGH YO, and HIGH QO. Since they are missing in known manuscripts, they are not included in this proposal.

3.1.2. Final consonants

Codas are optional for most vowels and written at the end of the syllable. Only 8 following consonants, which are also in the above-mentioned initial consonant list, can be used for codas:

Name	Glyph	IPA
LETTER LOW NYO		/-j/
LETTER NGO		/-ŋ/
LETTER NO		/-n/
LETTER MO		/-m/

Name	Glyph	IPA
LETTER VO		/-w/
LETTER LOW KO		/-k/
LETTER DO		/-t/
LETTER BO		/-p/

3.2. Vowels

3.2.1. Vowels with Zero initial

Like other Tai scripts, there is no independent vowel glyph. Like ə in Siamese Thai, *Tai Yo* uses LETTER QO for the zero initial or glottal stop onset. While Siamese Thai uses ə for both the glottal stop onset and vowel /-ɔ:-/ in non-initial position, *Tai Yo* uses a different glyph for vowel /-ɔ:-/: LETTER OO (O). When vowel OO /-ɔ:-/ is followed by no coda, the whole syllable is written as a standalone initial consonant.

The vowel UE /-i(:)-/ has two different forms. It occurs to the right of the consonant in open syllables, but sometimes follows it in closed syllables. This is comparable behavior to Thai SARA A and MAI HAN-AKAT, which have the same distribution. It therefore makes sense, like the Thai examples, to encode them separately and to not depend on complex behavior to choose the required form.

<i>mue</i>	<i>chue</i>	<i>pue</i>	<i>fue</i>
‘hand’	‘character’	‘arrow’	‘firewood’

3.2.2. Vowel placement

The way vowels are written in *Tai Yo* is different from other Tai scripts. Vowels succeed (are below) the initial consonant.

3.2.3. List of vowels

๕ (like ๗ in Tai Dam, ๗ in Siamese and ๕ in Sukhothai) is used in place of the initial consonant.

Succeeding vowels	Glyph	IPA	Tai Dam equivalent	Siamese equivalent	Sukhothai equivalent
LETTER AUE	๕	/-ə(:)-/	๗	ไก	๕๗
LETTER AA	๕	/-a:-/	๗	กา	๕๗
LETTER I	๕	/-i(:)-/	๕	กี	๕๗
LETTER U	๕	/-u(:)-/	๗	กู	๕๗
LETTER IA	๕	/-iə-/	๕		
LETTER E	๕	/-e(:)-/	๗	เก	๕๗
LETTER AE	๕	/-ɛ(:)-/	๗	แก	๕๗
LETTER UEA	๕	/-iə-/			
LETTER UA	๕	/-uə-/	๗		๕๗๗
LETTER O	๕	/-o(:)-/	๕	โก	๕๗
LETTER OO	๕	/-ɔ:-/	๗	กอ	๕๗
LETTER UE	๕	/-i(:)-/	๕	กี	๕๗

3.2.4. Rime signs

Phonetically, vowel short A /a/ only occurs in closed syllables. Rimes with short A /a/ are written a single glyph. This is the only instance where vowel length is indicated in Tai Yo, all other vowel signs double as long and short vowels:

			
<i>chap</i> 'snake skin'	<i>hak</i> 'root'	<i>mat</i> 'birthmark'	<i>nam</i> 'thorn'
			
<i>chăp</i> 'perched'	<i>hăk</i> 'love'	<i>măt</i> 'vermin'	<i>năm</i> 'water'

When vowel /-o-/ proceeds coda M forming rime OM /-om/, a single rime sign, called SIGN OM (◌̣), is used as an alternative of LETTER O (◌̣) + LETTER MO (◌̣).

The following glyphs represent such rimes:

Succeeding rimes	Glyph	IPA	Tai Dam equivalent	Siamese equivalent	Sukhothai equivalent
LETTER AN		/-an/	นั	กัน	ฉนณ
LETTER AM		/-am/	ทั	กำ	ฉา
LETTER AK		/-ak/	ทั	กัก	ฉฉฉ
LETTER AT		/-at/	ทั ⁴	กัต	ฉตต
LETTER AP		/-ap/	ทั	กั	ฉปป

Rime marks	Glyph	IPA	Tai Dam equivalent	Siamese equivalent	Sukhothai equivalent
SIGN UE		/-i:/	ทั	กี	ฉฉ

⁴ นั in Tai Don.

SIGN ANG	ᨧᩃ	/-aŋ/	ັ້ᨧ	ັ້ᨧ	ᨧᩃᩃ
SIGN AU	ᨧᩁ	/-aw/	ᨧᩁ	ᨧᩁ	ᨧᩁ
SIGN AY	ᨧᩁ+	/-aj/	ᨧᩁ	ᨧᩁ	ᨧᩁ
SIGN OM	ᨧᩁ)	/-om/	ᨧᩁᩃ	ᨧᩁᩃ	ᨧᩁᩃ

3.3. Medial W

When the phonemes LOW K and HIGH K are followed by medial /w/, LOW KVO (ᨧᩁᩃ, a ligature of LOW KO and LOW VO) (**Figure 9**) and HIGH KVO (ᨧᩁᩃ, a ligature of HIGH KO and LOW VO) (**Figure 10**) are used respectively.

When other initial consonants are followed by medial /w/, the glyph of SIGN AY is used (**Figure 3**). Thus, SIGN AY is interpreted as rime /-aj/ with no succeeding glyph and as medial /-w-/ with succeeding glyphs for the rime.

As /-aj/	As /-w-/
ᨧᩁ+	ᨧᩁ+ ᨧᩁ+ ᨧᩁ
/ŋaj/	/ŋwɔ̃n/

3.4. Special symbols

Name	Glyph	Tai Dam equivalent	Siamese equivalent	Notes
SYMBOL XAM LAI	ᨧᩁ	ᨧᩁ	ᨧᩁ	= Repetition mark.
SYMBOL MEUANG	ᨧᩁ			Used for syllable <i>muōng</i> /miəŋ ^{A2} / = village, country, region.

LETTER HIGH NYO (ᨧᩁ) is also used for syllable *nhi* /ni:^{B2}/ as in the proper name of *Taaw Nyi* [*Khun Cheuang Epic*] and a Sino-Vietnamese word *nhi* (二).

SYMBOL XAM LAI (ꨧ) has a stylistic variant: ꨧ. Use of this variant makes no difference in meaning and most authors consistently use either one or the other. One text is known in which the author switches variant halfway through the text on a new page.

DEFAULT SYMBOL XAM LAI	VARIANT SYMBOL XAM LAI
ꨧ	ꨧ

3.5. Punctuation

Most traditional Tai Yo manuscripts are written without punctuation marks. However, some aiding symbols are found in some manuscripts. All of these marks are the same as traditional Chinese punctuation marks and should be unified with them (**Figure 2**).

FULL STOP DOT	ENUMERATION COMMA
◦	、
Codepoint: U+3002	Codepoint: U+3001

Modern texts sometimes make use of western punctuation marks in a way similar to vertical Chinese and Japanese (**Figure 15**). For this, the Latin punctuation marks () , : < > [] are used, rotated as expected. In addition, full-width U+FF01 (!) and U+FF1F (?) are used upright. The centred forms are preferred.

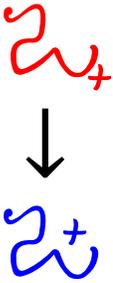
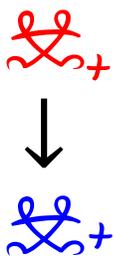
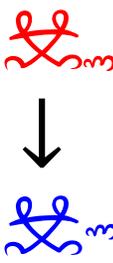
3.6. Numerals

Tai Yo does not use numerals, as numbers are generally spelled out. The use of numerals is predicted for modern texts, where it is expected that longer numbers will be written rotated in line with the text while individual numerals would be written upright. For the latter case full-width numerals (U+FF10 ... U+FF19) should be used.

4. Rendering

4.1. Mark positioning

Vowel marks should be rendered appropriately with the initial consonant.

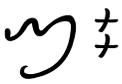
HIGH TO + SIGN AY	HIGH FO + SIGN AY	HIGH TO + SIGN ANG	HIGH FO + SIGN ANG
			
/taj/	/faj/	/taŋ/	/faŋ/

Since HIGH TO and HIGH FO are wider than other consonants, fonts without correct rendering function may have some positioning issues as in the red text: the vowel marks are too close to the tail of their preceding consonants. The blue ones are preferred and observed in manuscripts.

4.2. Medial W with vowel marks

When a word has SIGN AY as the medial /-w-/ and its vowel glyph is a right-side mark, the two marks are stacked vertically to avoid overlapping. Since one syllable has no more than one vowel mark, the right-side mark stack has no more than two layers. As far as we know, only /ŋwaj/ and /ʔwaj/ are found in Tai Yo vocabulary for such cases.

This behaviour is similar to Greek combining marks where, for example, U+0313 (◌̣) and U+0300 (◌̂) appear next to each other rather (◌̣̂) than stacked (*◌̣̂).

NGO + SIGN AY + SIGN AY	QO + SIGN AY + SIGN ANG
	
/ŋwaj/	/ʔwaj/

4.3. Tone marks

Although Tai Yo is a tonal language, the traditional *Tai Yo* script uses no tone indicators. Readers have to guess the tone and meaning of a word from context. Modernizing attempts consider adding

a tone mark system based on its sister scripts. Tai languages follow a common tone system, conventionally described as having 4 phonemic tones: A, B, C and D, where D is only for dead syllables (always with plosive codas). Tones in Siamese Thai have been indicated since Sukhothai era as in the Ram Khamhaeng stele, which dates from 1283-1290 AD. In the Sukhothai script and the modern Thai script, tone A and D have no tone mark while tone B and tone C are represented by a diacritic on top of the syllable. The modern Lao script developed the same tone indicating system. Similarly since 1970s, modern scripts of Black Tai and White Tai have also developed tone marks for tones B and C which are included in Tai Viet Unicode block.

Tone marks	Tone marks of Tai Yo in a modernizing attempt	Tai Dam equivalent	Siamese equivalent	Lao equivalent	Sukhothai equivalent
MARK 1 (TONE B)	ᵀ	ᵀ	ᵀ	ᵀ	ᵀ
MARK 2 (TONE C)	ᶇ	ᶇ	ᶇ	ᶇ	ᶇ

In one modernizing attempt (**Figure 16**), there are 2 tone marks for tones B and C. Both are written to the left of the onset of a syllable. We are not proposing these tone marks in this proposal but leaving them as a potential future additions.

5. Unicode Character Data

5.1. Collation

ᵀ LOW KO < ᵀ HIGH KO < ᵀ LOW KHO < ᵀ HIGH KHO < ᵀ GO < ᵀ NGO < ᵀ CO < ᵀ LOW XO < ᵀ HIGH XO < ᵀ LOW NYO < ᵀ HIGH NYO < ᵀ DO < ᵀ LOW TO < ᵀ HIGH TO < ᵀ THO < ᵀ NO < ᵀ BO < ᵀ LOW PO < ᵀ HIGH PO < ᵀ PHO < ᵀ LOW FO < ᵀ HIGH FO < ᵀ MO < ᵀ YO < ᵀ LO < ᵀ VO < ᵀ LOW HO < ᵀ HIGH HO < ᵀ QO < ᵀ LOW KVO < ᵀ HIGH KVO < ᵀ AA < ᵀ I < ᵀ UE < ᵀ SIGN UE < ᵀ U < ᵀ AE < ᵀ SIGN AU < ᵀ O < ᵀ E < ᵀ IA < ᵀ UEA < ᵀ UA < ᵀ OO < ᵀ AUE < ᵀ SIGN AY < ᵀ SIGN ANG < ᵀ AN < ᵀ AM < ᵀ AK < ᵀ AT < ᵀ AP < ᵀ SIGN OM

5.2. Line breaking and space

Word spacing between words is optional in *Tai Yo* manuscripts. It appears that word spacing was more common practice among higher ranking writers as in **Figure 2** and **Figure 3**. In the revitalization effort, spacing between words is the default, although non-spaced texts can still be found.

In manuscripts without word spacing, line breaking could occur anywhere, even in the middle of a syllable between letters and only excludes line breaking before non-spacing marks.

5.3. Directionality

The *Tai Yo* script is generally written from top to bottom in columns running from right to left. In modern publications, if vertical text is not possible, the line is simply turned on its side and runs from left to right. The letters don't stay upright like in Japanese or Chinese, but are turned on their side, much like Mongolian.

According to UAX#50, the vertical property of *Tai Yo* is R. The characters are displayed sideways in the code charts as they would appear in a horizontal layout and should be rotated 90 degrees in a vertical layout.

5.4. Character properties

```
1E6C0;TAI YO LETTER LOW KO;Lo;0;L;;;;N;;;;;
1E6C1;TAI YO LETTER HIGH KO;Lo;0;L;;;;N;;;;;
1E6C2;TAI YO LETTER LOW KHO;Lo;0;L;;;;N;;;;;
1E6C3;TAI YO LETTER HIGH KHO;Lo;0;L;;;;N;;;;;
1E6C4;TAI YO LETTER GO;Lo;0;L;;;;N;;;;;
1E6C5;TAI YO LETTER NGO;Lo;0;L;;;;N;;;;;
1E6C6;TAI YO LETTER CO;Lo;0;L;;;;N;;;;;
1E6C7;TAI YO LETTER LOW XO;Lo;0;L;;;;N;;;;;
1E6C8;TAI YO LETTER HIGH XO;Lo;0;L;;;;N;;;;;
1E6C9;TAI YO LETTER LOW NYO;Lo;0;L;;;;N;;;;;
1E6CA;TAI YO LETTER HIGH NYO;Lo;0;L;;;;N;;;;;
1E6CB;TAI YO LETTER DO;Lo;0;L;;;;N;;;;;
1E6CC;TAI YO LETTER LOW TO;Lo;0;L;;;;N;;;;;
1E6CD;TAI YO LETTER HIGH TO;Lo;0;L;;;;N;;;;;
1E6CE;TAI YO LETTER THO;Lo;0;L;;;;N;;;;;
1E6CF;TAI YO LETTER NO;Lo;0;L;;;;N;;;;;
1E6D0;TAI YO LETTER BO;Lo;0;L;;;;N;;;;;
1E6D1;TAI YO LETTER LOW PO;Lo;0;L;;;;N;;;;;
1E6D2;TAI YO LETTER HIGH PO;Lo;0;L;;;;N;;;;;
1E6D3;TAI YO LETTER PHO;Lo;0;L;;;;N;;;;;
1E6D4;TAI YO LETTER LOW FO;Lo;0;L;;;;N;;;;;
1E6D5;TAI YO LETTER HIGH FO;Lo;0;L;;;;N;;;;;
1E6D6;TAI YO LETTER MO;Lo;0;L;;;;N;;;;;
1E6D7;TAI YO LETTER YO;Lo;0;L;;;;N;;;;;
1E6D8;TAI YO LETTER LO;Lo;0;L;;;;N;;;;;
1E6D9;TAI YO LETTER VO;Lo;0;L;;;;N;;;;;
1E6DA;TAI YO LETTER LOW HO;Lo;0;L;;;;N;;;;;
1E6DB;TAI YO LETTER HIGH HO;Lo;0;L;;;;N;;;;;
1E6DC;TAI YO LETTER QO;Lo;0;L;;;;N;;;;;
1E6DD;TAI YO LETTER LOW KVO;Lo;0;L;;;;N;;;;;
1E6DE;TAI YO LETTER HIGH KVO;Lo;0;L;;;;N;;;;;
1E6E0;TAI YO LETTER AA;Lo;0;L;;;;N;;;;;
1E6E1;TAI YO LETTER I;Lo;0;L;;;;N;;;;;
1E6E2;TAI YO LETTER UE;Lo;0;L;;;;N;;;;;
1E6E3;TAI YO SIGN UE;Mn;230;NSM;;;;N;;;;;
1E6E4;TAI YO LETTER U;Lo;0;L;;;;N;;;;;
1E6E5;TAI YO LETTER AE;Lo;0;L;;;;N;;;;;
1E6E6;TAI YO SIGN AU;Mn;230;NSM;;;;N;;;;;
1E6E7;TAI YO LETTER O;Lo;0;L;;;;N;;;;;
1E6E8;TAI YO LETTER E;Lo;0;L;;;;N;;;;;
1E6E9;TAI YO LETTER IA;Lo;0;L;;;;N;;;;;
```

```

1E6EA;TAI YO LETTER UEA;Lo;0;L;;;;N;;;
1E6EB;TAI YO LETTER UA;Lo;0;L;;;;N;;;
1E6EC;TAI YO LETTER OO;Lo;0;L;;;;N;;;
1E6ED;TAI YO LETTER AUE;Lo;0;L;;;;N;;;
1E6EE;TAI YO SIGN AY;Mn;230;NSM;;;;N;;;
1E6EF;TAI YO SIGN ANG;Mn;230;NSM;;;;N;;;
1E6F0;TAI YO LETTER AN;Lo;0;L;;;;N;;;
1E6F1;TAI YO LETTER AM;Lo;0;L;;;;N;;;
1E6F2;TAI YO LETTER AK;Lo;0;L;;;;N;;;
1E6F3;TAI YO LETTER AT;Lo;0;L;;;;N;;;
1E6F4;TAI YO LETTER AP;Lo;0;L;;;;N;;;
1E6F5;TAI YO SIGN OM;Mn;230;NSM;;;;N;;;
1E6FE;TAI YO SYMBOL MEUANG;Lo;0;L;;;;N;;;
1E6FF;TAI YO XAM LAI;Lm;0;L;;;;N;;;

```

5.5. Indic properties

Due to the linear nature of the script, no Indic character data is needed.

5.6. Line breaking properties

```

1E6C0.. 1E6E2;AL
1E6E3;CM
1E6E4;AL
1E6E5;AL
1E6E6;CM
1E6E7.. 1E6ED;AL
1E6EE;CM
1E6EF;CM
1E6F0.. 1E6F4;AL
1E6F5;CM
1E6FE;AL
1E6FF;AL

```

5.7. Script metadata

Name	Tai Yo
Size	55
Sample	ꨀ
Sample_Code	1E6D5
Origin_Country	Vietnam
Density	1
Likely language	Tai Yo
Lang_Code	tyj
ID Usage	Limited_Use
RTL?	No
LB letters?	Yes
Shaping req?	Yes
IME?	No
Case?	No

6.

	1E6C	1E6D	1E6E	1E6F
0	ㄱ	ㄴ	ㄷ	ㄹ
1	ㅁ	ㅂ	ㅅ	ㅇ
2	ㄷ	ㄹ	ㅁ	ㅂ
3	ㄷ	ㅁ	ㅂ	ㅅ
4	ㅁ	ㅂ	ㅅ	ㅇ
5	ㄱ	ㄴ	ㄷ	ㄹ
6	ㄷ	ㄹ	ㅁ	
7	ㅂ	ㅅ	ㅇ	
8	ㄴ	ㅇ	ㄷ	
9	ㄹ	ㄷ	+	
A	ㅅ	ㅇ	ㄱ	
B	ㅇ	ㄷ	ㅂ	
C	ㄷ	ㅁ	ㅇ	
D	ㄷ	ㅂ	ㅅ	
E	ㄱ	ㅁ	ㅂ ⁺	ㄹ
F	ㄹ		ㅂ ^ㅁ	ㄷ

Letters

1E6C0	𑜀	TAI YO LETTER LOW KO
1E6C1	𑜁	TAI YO LETTER HIGH KO
1E6C2	𑜂	TAI YO LETTER LOW KHO
1E6C3	𑜃	TAI YO LETTER HIGH KHO
1E6C4	𑜄	TAI YO LETTER GO
1E6C5	𑜅	TAI YO LETTER NGO
1E6C6	𑜆	TAI YO LETTER CO
1E6C7	𑜇	TAI YO LETTER LOW XO
1E6C8	𑜈	TAI YO LETTER HIGH XO
1E6C9	𑜉	TAI YO LETTER LOW NYO
1E6CA	𑜊	TAI YO LETTER HIGH NYO
1E6CB	𑜋	TAI YO LETTER DO
1E6CC	𑜌	TAI YO LETTER LOW TO
1E6CD	𑜍	TAI YO LETTER HIGH TO
1E6CE	𑜎	TAI YO LETTER THO
1E6CF	𑜏	TAI YO LETTER NO
1E6D0	𑜐	TAI YO LETTER BO
1E6D1	𑜑	TAI YO LETTER LOW PO
1E6D2	𑜒	TAI YO LETTER HIGH PO
1E6D3	𑜓	TAI YO LETTER PHO
1E6D4	𑜔	TAI YO LETTER LOW FO
1E6D5	𑜕	TAI YO LETTER HIGH FO
1E6D6	𑜖	TAI YO LETTER MO
1E6D7	𑜗	TAI YO LETTER YO
1E6D8	𑜘	TAI YO LETTER LO
1E6D9	𑜙	TAI YO LETTER VO
1E6DA	𑜚	TAI YO LETTER LOW HO
1E6DB	𑜛	TAI YO LETTER HIGH HO
1E6DC	𑜜	TAI YO LETTER QO
1E6DD	𑜝	TAI YO LETTER LOW KVO
1E6DE	𑜞	TAI YO LETTER HIGH KVO
1E6DF		<reserved>

Vowels and Finals

1E6E0	𑜀	TAI YO LETTER AA
1E6E1	𑜁	TAI YO LETTER I
1E6E2	𑜂	TAI YO LETTER UE
1E6E3	𑜃	TAI YO SIGN UE
1E6E4	𑜄	TAI YO LETTER U
1E6E5	𑜅	TAI YO LETTER AE
1E6E6	𑜆	TAI YO SIGN AU
1E6E7	𑜇	TAI YO LETTER O
1E6E8	𑜈	TAI YO LETTER E
1E6E9	𑜉	TAI YO LETTER IA
1E6EA	𑜊	TAI YO LETTER UEA
1E6EB	𑜋	TAI YO LETTER UA
1E6EC	𑜌	TAI YO LETTER OO
1E6ED	𑜍	TAI YO LETTER AUE
1E6EE	𑜎	TAI YO SIGN AY
1E6EF	𑜏	TAI YO SIGN ANG
1E6F0	𑜐	TAI YO LETTER AN
1E6F1	𑜑	TAI YO LETTER AM
1E6F2	𑜒	TAI YO LETTER AK
1E6F3	𑜓	TAI YO LETTER AT
1E6F4	𑜔	TAI YO LETTER AP
1E6F5	𑜕	TAI YO SIGN OM
1E6F6		<reserved>
1E6F7		<reserved>
1E6F8		<reserved>
1E6F9		<reserved>
1E6FA		<reserved>
1E6FB		<reserved>
1E6FC		<reserved>
1E6FD		<reserved>

Symbols

1E6FE	𑜀	TAI YO SYMBOL MEUANG
1E6FF	𑜁	TAI YO XAM LAI

7. Materials



Figure 1: Two pages from a typical traditional *Tai Yo* manuscript. The text is written without space and punctuation.

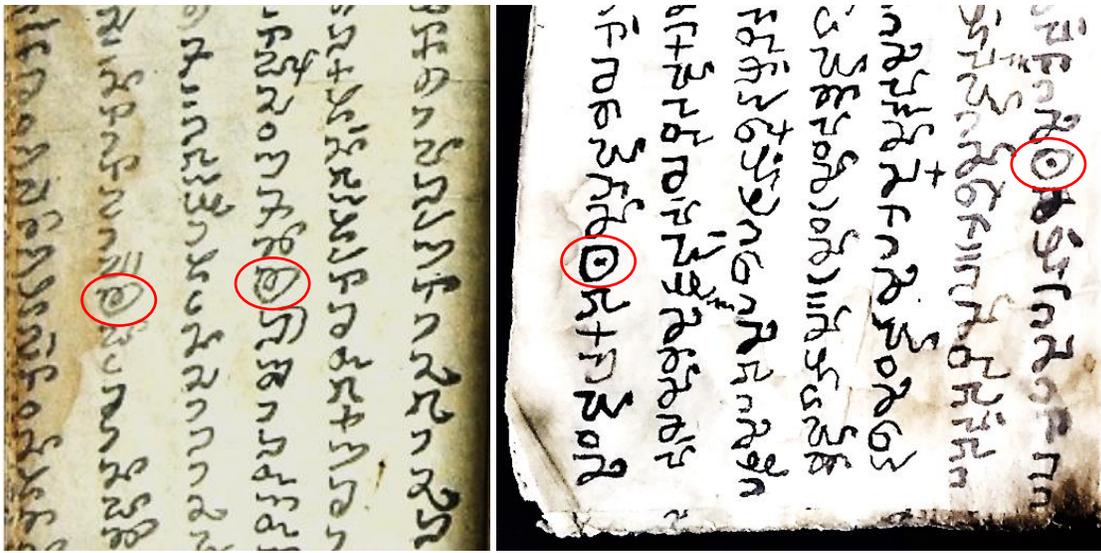


Figure 4: Examples of the symbol of *Muōng* (☉ - TAI YO SYMBOL MEUANG).

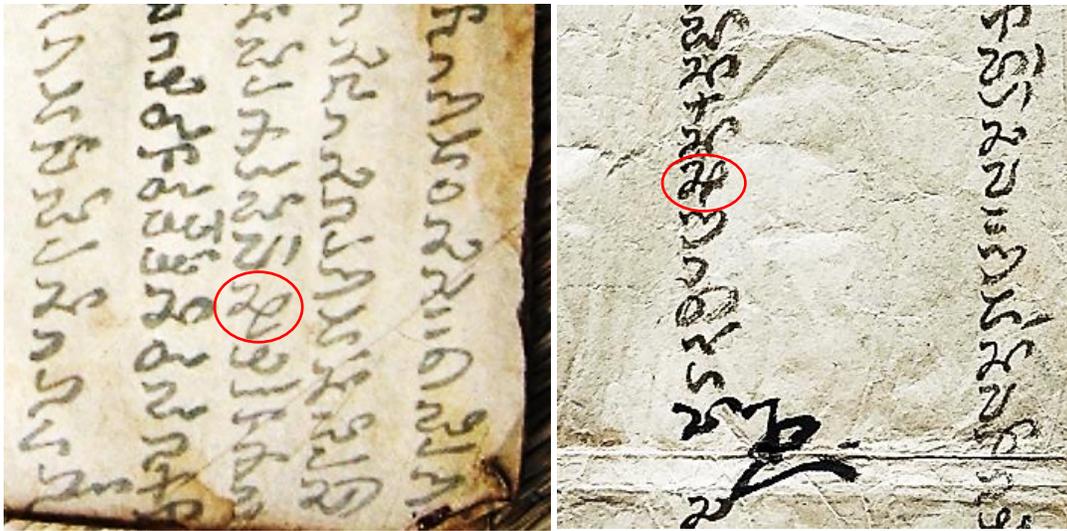


Figure 5: Examples of the symbol of *Nhi* (☼ - TAI YO SYMBOL NYI).

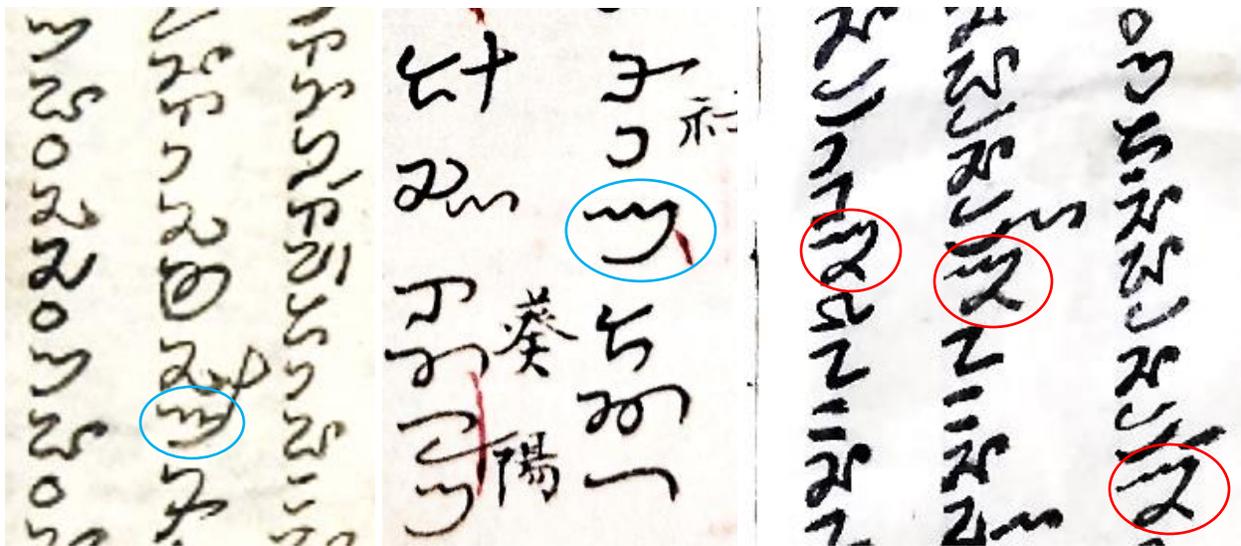


Figure 6: Examples of the symbol of Xăm lāi: TAI YO SYMBOL REPETITION (ꨀ) - blue circle) and TAI YO SYMBOL REPETITION ALTERNATIVE (ꨁ - red circle).

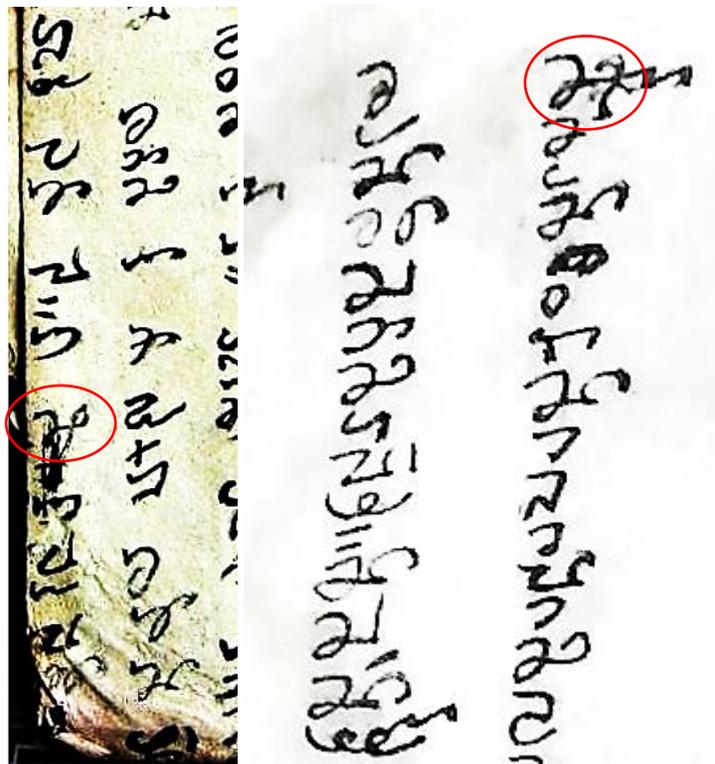


Figure 7: Examples of TAI YO LETTER HIGH NYO (ꨀ).

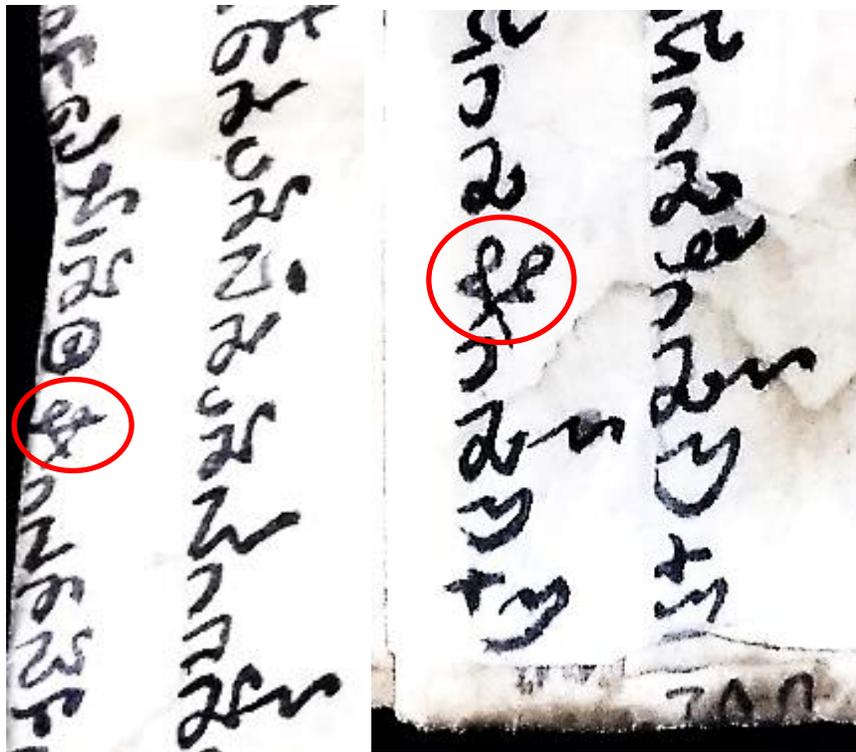


Figure 8: Examples of TAI YO LETTER HIGH FO (ꨲ).

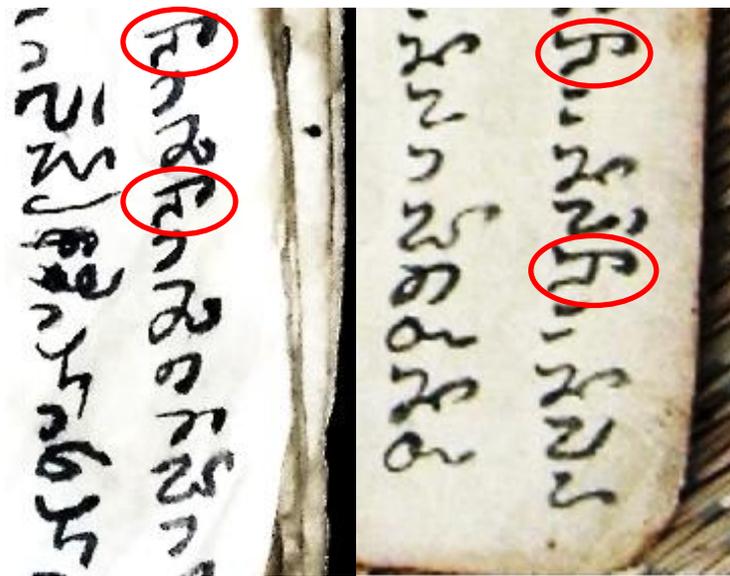


Figure 9: Examples of TAI YO LETTER LOW KVO (ꨳ).

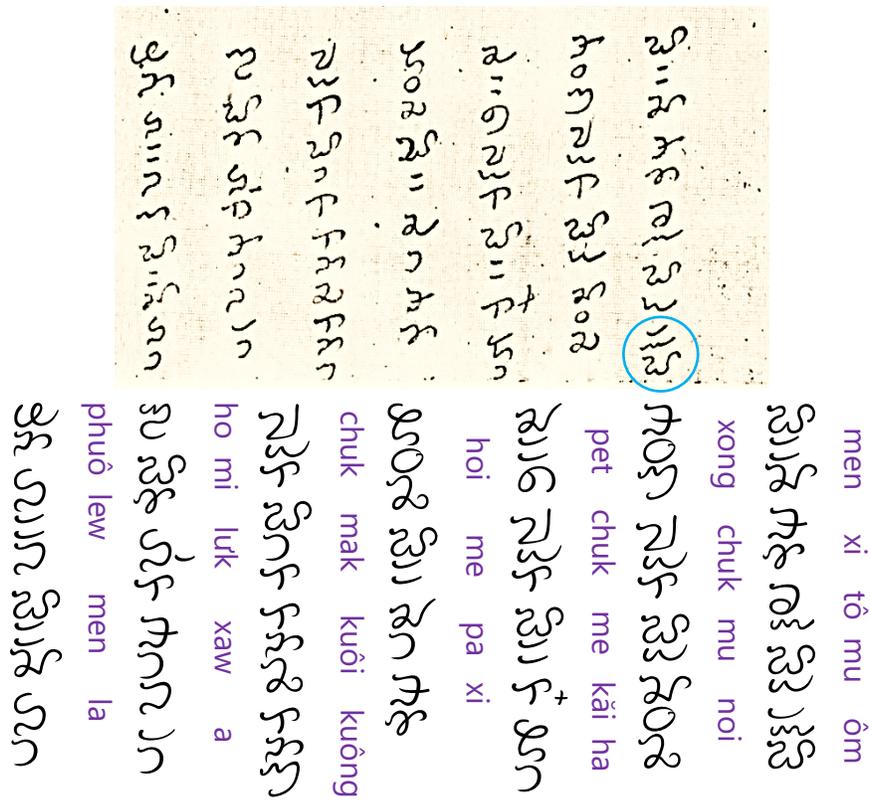


Figure 11: Examples of TAI YO SIGN OM (◌̃ - red circle) and TAI YO LETTER O + TAI YO LETTER HIGH MO (◌̃ - blue circle) as rime OM /-om/.

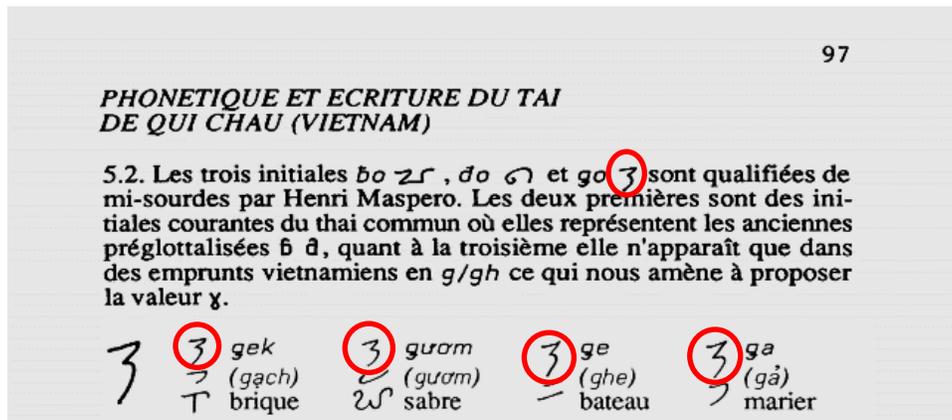


Figure 12: TAI YO LETTER HIGH GO (◌̃) in Ferlus, 1993.

N.	Chữ cái	Tên chữ	Chữ-Việt	N.	Chữ cái	Tên chữ	Chữ-Việt
33	o	cọc	o	45	๒	tró	tr
34	๓	cố	ô	46	๓	cú	u
35	๔	cớ	ơ	47	๔	cúa	ua
36	๕	pó	p	48	๕	cứ	ư
37	๖	po	p	49	๖	cúra	ura
38	๗	phó	ph	50	๗	vó	v
39	๘	'phó	'ph	51	๘	xó	x
40	๙	ró	r	52	๙	xo	x
41	๑๐	só	s	53	๑๐	xăm mo	-
42	๑๑	tó	t	54	๑๑	ó	-
43	๑๒	to	t	55	๑๒	p'ha nha	p'h
44	๑๓	thó	th				

Figure 13: TAI YO LETTER HIGH SO (๙๐, yellow), TAI YO LETTER LOW RO (๑๐, green), and TAI YO LETTER LOW TRO (๒๐, blue) in modern materials – a Lai Tay teaching textbook (Sầm Văn Bình, 2014).

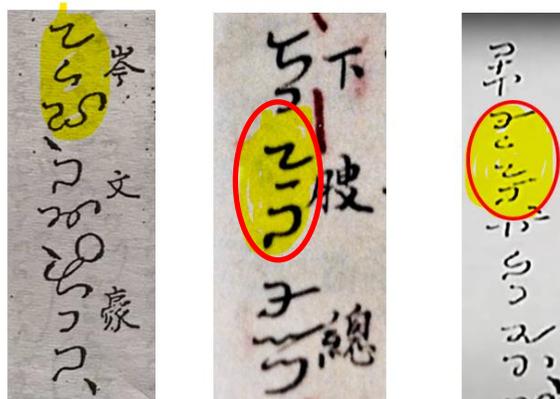


Figure 14: TAI YO LETTER THO (๑๒) for phoneme HIGH SO in old materials for SÀM (๑๒๑) and SUU (๑๒) - red circle).

BIBLIOGRAPHY

1. Sầm, Văn Bình (2014), *Tài liệu dạy học chữ Thái (Hệ Lai Tay) – Sách giáo khoa* [A Teaching document of Learning Tai script (Lai Tay) – Textbook], Vinh University Press.
2. Sầm, Công Danh (2021). *Tones of Tai Yo language via experimental parameter*. Graduation Thesis (BA). Faculty of Linguistics, University of Social Sciences and Humanity – Hanoi VNU (In Vietnamese; a summary in English: <http://dx.doi.org/10.13140/RG.2.2.18266.88007>)
3. Ferlus, Michel (1993). Phonétique et écriture du Tai de Qui Châu (Vietnam). In: *Cahiers de linguistique – Asie Orientale*. Vol 22 (1). <https://doi.org/10.3406/clao.1993.1432>
4. Ferlus, Michel (2008). Les dialectes Tai du Nghệ An, Vietnam (Tay Daeng, Tay Yo, Tay Muong). In: Diller, Anthony V.N.; Edmondson, Jerold A.; Luo, Yongxian (ed.). *The Tai-Kadai languages*, Routledge, pp. 298-315, <halshs-01617791>
5. Sầm, Công Danh (2021). *Bộ chữ Lai Tay và chữ Tay Dọ - Một số vấn đề và giải pháp*. Memorandum of National Scientific Conference named: ‘Summarizing the research results and submitting the proposals to Government and related state agencies for emergency problems in preserving and developing the role, identity of the language and writing of ethnic minorities’
6. *TAI YO - Language and Relations - Beyond and Inside*. Website: <https://www.facebook.com/Taiyocore>

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646⁵.**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://std.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://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest *Roadmaps*.

A. Administrative

1. Title:	Proposal to encode Tai Yo		
2. Requester's name:	<i>F. van de Kasteelen, Sam Cong Danh, Nguyen Viet Khoi</i>		
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual contribution</i>		
4. Submission date:	<i>13-07-2022</i>		
5. Requester's reference (if applicable):			
6. Choose one of the following:			
This is a complete proposal:	<input checked="" type="checkbox"/>		
(or) More information will be provided later:	<input type="checkbox"/>		

B. Technical – General

1. Choose one of the following:			
a. This proposal is for a new script (set of characters):	<input checked="" type="checkbox"/>	<i>Yes</i>	
Proposed name of script:	<i>Tai Yo</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>55</i>		
3. Proposed category (select one from below - see section 2.2 of P&P document):			
A-Contemporary <input checked="" type="checkbox"/> B.1-Specialized (small collection) <input type="checkbox"/> B.2-Specialized (large collection) <input type="checkbox"/>			
C-Major extinct <input type="checkbox"/> D-Attested extinct <input type="checkbox"/> E-Minor extinct <input type="checkbox"/>			
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/> G-Obscure or questionable usage symbols <input type="checkbox"/>			
4. Is a repertoire including character names provided?	<input checked="" type="checkbox"/>		
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<input checked="" type="checkbox"/>		
b. Are the character shapes attached in a legible form suitable for review?	<input checked="" type="checkbox"/>		
5. Fonts related:			
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Viet Khoi Nguyen</i>		
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):			
6. References:			
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<input checked="" type="checkbox"/>		
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<input checked="" type="checkbox"/>		
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)?	<input checked="" type="checkbox"/>		
8. Additional Information:			
<p>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 Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.</p>			

⁵ Form number: N4502-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, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain	No
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.)? If YES, with whom? If YES, available relevant documents:	Yes Sam Cong Danh
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	Yes 1. / 2.
4. The context of use for the proposed characters (type of use; common or rare) Reference:	Rare 2.
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	Yes Vietnam, 1. / 2.
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:	No
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? If YES, is a rationale for its inclusion provided? If YES, reference:	No
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	No
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	No
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:	Yes Yes 3.2.4
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	No
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:	No