Writing Tai Don
Additional characters needed for the Tai Viet script

Jim Brase,
SIL Non-Roman Script Initiative (NRSI)

This paper explores the differences between the writing of the Tai Dam, Tai Don, and Jinping Dai languages,[1] or dialects. It makes the assumption that the Tai Don and Jinping scripts will be unified with the Tai Viet script, and seeks to make a determination of what Tai Don or Jinping characters need to be added to the Tai Viet character repertoire in order to write those languages.

Note that the current Tai Viet character repertoire is adequate for writing Tai Don if one adopts the Tai Dam orthographic conventions of Son La Province. This is what the Tai community of Vietnam is trying to standardize on. However, if one wishes to write with the classical Tai Don conventions of Lai Chau province, or for recording classical Tai Don documents, additional characters are needed.

1 Language and script background

The Tai Don language (also known as White Tai) is spoken in the provinces of Lai Chau and Dien Bien in northwestern Vietnam, and in the surrounding areas of Vietnam, Lao, and China. Jinping Dai is spoken in the Jinping district of southern Yunnan Province, China, which borders on Lai Chau Province, Vietnam. Tai Dam (Black Tai) is centered in Son La, the next province to the south of Lai Chau. Immigrant populations speaking the Tai Dam language, and to a lesser extent Tai Don, can be found in several other countries around the world.

Gedney (1989) and Fippinger (1970) provide in-depth phonological comparisons of Tai Dam and Tai Don. For the purpose of this paper, we will briefly look at those features which impact the writing.

1.1 Initial consonants

The Tai Don language has four initial aspirated consonants which do not occur in Tai Dam, as shown by Table 1.1. If the writing is an accurate guide, Jinping Dai has the same consonant inventory as Tai Don.

---

[1] Western linguists generally classify Tai Dam and Tai Don as two separate languages, but the Vietnamese government considers them to be two dialects of the same language. The linguistic classification for Jinping Dai is uncertain. The author had once undestood it to be a dialect of Tai Don, on the basis of the two having identical phoneme inventories, but he has since learned that there are significant differences in the tone systems.

Tai Don & Jinping Dai	p, t, tɕ, k, kʷ, ʔ, b, d, pʰ, tʰ, tɕʰ, kʰ, kʰʷ, f, s, x, xʷ, h, v, l, j, m, n, ɲ, ŋ, ŋʷ
Tai Dam[2]	p, t, tɕ, k, kʷ, ʔ, b, d, tʰ, f, s, x, xʷ, h, v, l, j, m, n, ɲ, ŋ, ŋʷ

Table 1.1: Initial consonant inventory

This difference is already reflected in the Tai Viet character repertoire, in the form three pairs of consonants that were added specifically for Tai Don:

TAI VIET LETTER LOW KHO   TAI VIET LETTER HIGH KHO
TAI VIET LETTER LOW CHO   TAI VIET LETTER HIGH CHO
TAI VIET LETTER LOW PHO   TAI VIET LETTER HIGH PHO

Tai Don can be written with the resulting character set, but only if one uses these six characters in conjunction with the orthographic conventions of Son La. If one wishes to write Tai Don using Lai Chau orthographic conventions, some additional characters will probably be needed.

1.2   Vowels

Gedney found 13 vowels in Tai Dam, 10 in Tai Don, as shown in Table 1.2. Jinping Dai has the same vowel inventory as Tai Don.

Tai Dam Vowels			Tai Don & Jinping Dai Vowels		
i	ɨ	u	i	ɨ	u
iᵊ	ɨᵊ	uᵊ			
e	ə	o	e	ə	o
ɛ	a	ɔ	ɛ	a	ɔ
	aa			aa	

Table 1.2: Vowels[3]

The Tai Dam vowel glides, /iᵊ/, /ɨᵊ/, and /uᵊ/ consistently change to /e/, /ə/, and /o/ in Tai Don. This is reflected in the writing, where the same symbols used for Tai Dam /iᵊ/ and /uᵊ/ are also used for Tai Don /e/ and /o/.

1.3   Tones

Differences in the tone systems of the three languages are not reflected in the writing. All three have six tones on unchecked syllables. On checked syllables (syllables which end with /p/, /t/, /k/, or /ʔ/) Tai Dam and Tai Don have only two tones, but Jinping Dai has three. However, until the latter half of the 20th century, none of these writing systems had any tone marks. In recent years, various communities have sought to correct the deficiency by borrowing from other writing systems. The differences in the

---

[2] For Tai Dam, Gedney lists the aspirated velar stops, kʰ and kʰʷ, but Fippinger analyzes them as the fricatives /x/ and /xʷ/. This may reflect a variation in dialects.

[3] The vowel shown here as /ɨ/ high-central unrounded, has sometimes been analyzed as /ɯ/ high-back unrounded.

tone marks, then, is the result of what sources were chosen for the borrowings rather than any differences in the languages.

1.3.1 Tonal notation

The conflicting notation used by various authors to identify the tones can be confusing. I will make a brief comparison of the notational systems here, in an effort to reduce the confusion. For our purposes, it is sufficient to look at only the tones on unchecked syllables.

Gedney and Fippinger numbered the tones of both Tai Dam and Tai Don from 1 through 6. Although the phonetic realization of the tones was different between the two languages, tones with the same number were cognates. Luo Mei Zhen (1992) also numbered the tones from 1 through 6, but the tone numbers were assigned in a different sequence. Table 1.3 compares the tone numbers used by Gedney with those used by Luo. I will follow the tone numbering used by Gedney.

There are also two different ways of identifying the tonal series of the consonants. In linguist circles, the distinction is made based on the historical derivation of the tones. Tones 1-3, which derived from the voiceless consonants of Proto-Tai, are designated as the "high series". Tones 4-6, which derived from the voiced consonants of Proto-Tai, are designated the "low series". But in the popular usage of the Tai people of Vietnam, the designations are based on the perceived pitch level of the tones. This results in the labels being reversed. Tones 1-3 are called "low" by the people, and tones 4-6 are called "high".

The proposed Unicode names for the Tai Viet characters are based on the popular usage.

To avoid confusion in this paper, I have avoided using the terms "high" and "low", except when referring to a specific Unicode character name. Otherwise I use phonemic notation for the consonants along with tone numbers. E.g. /k4/ for /k/ tones 4-6.

Gedney, Fippinger tone numbers	Luo tone numbers	Tone series, linguistic	Tone series, popular usage; (used in Unicode names)
1	1	high	low
2	5	high	low
3	3	high	low
4	2	low	high
5	6	low	high
6	4	low	high

Table 1.3: Comparison of tonal notation from various sources.

1.4   Script dialects and orthographic conventions

Not only is there a difference between the orthographic conventions of the Tai Don and the Tai Dam, there is considerable variation within the Tai Don community itself. I have been able to identify three different regional traditions in Vietnam, which I have classified as shown in Table 1.4. I have also wanted to compare the Jinping Dai to the Tai Don traditions, so I have treated it as a fourth tradition.

Tradition	Approximate Region	Description	Sources
Central	Lai Chau	central Lai Chau province	Điêu and Donaldson 1970 Minot 1940
Western	Mường Cha	district of northern Dien Bien province, formerly part of Lai Chau province and named Mường Lay	Các mẫu tự Thái… Lo 1966
Northern	Phong Thổ	district in northeastern Lai Chau province; the Tai name is Mường So	Các mẫu tự Thái… Lo 1966
Jinping	Jinping	Jinping district of Yunnan province, China; across the border from Lai Chau province	character charts and sample text in Chinese documents

Table 1.4: Regions and traditions of Tai Don writing

Note that my best data comes from the Central dialect. It is the only one for which I have a moderate amount of text, in addition to several code tables. I have four code tables and one short text from Jinping, along with the help of indirect contact with a Jinping expert. But I have only two code tables and no text for the Western and Northern dialects.

2 Character variations and contrasts

The four Tai Don traditions listed in section 1.4 use numerous glyph variations. I will attempt to be exhaustive in my examination of these glyphs, considering any glyph which may possible represent a new character, and trying to determine which ones truly represent unique characters and which ones are only stylistic variations.

In the following tables, the second and third columns show the glyphs from the TayViet font, which has been used for the representative glyphs in Unicode, and SIL's Tai Heritage font, in order to give a flavor of the stylistic variations one might expect in these characters. The glyphs shown in the other four columns are designed in the style of the Tai Heritage font, partly because that font was available to me for editing, and partly because my Jinping contacts were not familiar with the TayViet style.

The fist time that a new glyph appears in one of the tables, it is given a number. This is for reference purposes, and the reference number may be referred to throughout the paper.

After examining a glyph and its uses, I attempt to place each glyph into one of three categories:

1. Distinct form with contrasting usage—These glyphs need to be disunified.
2. Distinct form with common usage—This is the most common category. I am often uncertain whether these glyphs should be unified or disunified. Frequently the difference in form is so pronounced as to suggest they should be disunified, but the lack of contrast in usage undermines that position.
3. Similar form with common usage—The contrast between these glyphs and current characters are not as strong as in category 2. There is no contrast in usage with current characters. These glyphs should clearly be unified.

2.1 /k1/ and /t4/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)		Central (Lai Chau)	Northern (Phong Thổ)	Jinping	
/k/ tone 1-3	ꨆ LOW KO	ꨆ LOW KO	ꨆ		ꨆ	ꨆ	ꨆ	ꨆ Glyph 1
/t/ tone 4-6	ꨈ HIGH TO	ꨈ HIGH TO	ꨈ	ꨆ	ꨈ	ꨈ Glyph 2	ꨈ	

Table 2.1: /k1/ and /t4/

The tone 1-3 /k/ is fairly consistent in form, but shows some variation in how tightly the second stroke binds to the left stem (the left stem being the first stroke). This is illustrated by the difference between the LOW KO characters in the TayViet and Tai Heritage font. A similar variation exists with the tone 4-6 /t/. The result is that the visual contrast between the two characters can be lost. This is analogous to the lost of contrast between the LATIN LETTER CAPITAL I and LATIN LETTER SMALL L in some styles.

Glyph 1is sometimes an alternative glyph for tone 1-3 /k/, more often for tone 4-6 /t/, but it is not a new character.

In Tai Don and Jinping, Glyph 2 is an alternative glyph for TAI VIET LETTER HIGH TO (tone 4-6 /t/). It never contrasts with it, and is not a new character.

I classify these glyphs as follows:

- Glyph 1: Similar form with common usage—unify
- Glyph 2: Similar form with common usage—unify

2.2   /k1/, /k4/, and /x4/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)	Central (Lai Chau)	Northern (Phong Thổ)	Jinping	
/k/ tone 1-3	ഗ LOW KO	℩ LOW KO	℩	℩	℩	℩	
/k/ tone 4-6	တ HIGH KO	α HIGH KO	α	ℵ Glyph 3	ℵ	ℵ	η Glyph 4
/x/ tone 4-6	3 HIGH KHHO	3 HIGH KHHO	η	α	3	η Glyph 5	

Table 2.2: /k1/, /k4/, and /x4/

In Table 2.2, the Central dialect (Lai Chau) shows a three-way contrast between Glyph 3 (used for /k4/), TAI VIET LETTER LOW KO (used for /k1/), and TAI VIET LETTER HIGH KO (used for /x4/). At the same time, the northern dialect (Phong Thổ) provides contrast between Glyph 3 and TAI VIET LETTER HIGH KHHO. The conclusion is that Glyph 3 is a unique character and must be encoded.

In the Western dialect (Mường Cha), Glyph 4 contrasts with TAI VIET LETTER LOW KO and TAI VIET LETTER HIGH KO. But it does not contrast with TAI VIET LETTER HIGH KHHO. So we can conclude that in the Western dialect, Glyph 4 is a stylistic variation of TAI VIET LETTER HIGH KHHO.

One Jinping expert said that Glyph 4 is wrong—that the two legs of the character should be the same length. We can conclude that it either does not exist in the Jinping dialect, or it is a stylistic variation of Glyph 3. If it is a stylistic variation of Glyph 3, it represents a loss of visual contrast between Jinping /k4/ and Western /x4/.

In the Jinping dialect, Glyph 5 contrasts with TAI VIET LETTER LOW KO and Glyph 3. However, it has not been possible to establish a contrast with TAI VIET LETTER HIGH KHHO. Visually, the two are very different, and it is tempting to regard Glyph 5 as a unique character. But it would be possible to unify it with the HIGH KHHO.

I classify these glyphs as follows:

- Glyph 3: Distinct form with contrasting usage—disunify
- Glyph 4: Similar form with common usage—unify with TAI VIET LETTER HIGH KHHO
- Glyph 5: Distinct form with common usage—unify with TAI VIET LETTER HIGH KHHO

2.3 /x1/ and /kʰ1/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)	Central (Lai Chau)	Northern (Phong Thổ)		Jinping	
/x/ tone 1-3	ꪁ LOW KHHO	ꪁ LOW KHHO	ꪁ	ꪁ	Glyph 6	Glyph 7	Glyph 8	Glyph 9
/kʰ/ tone 1-3	ꪂ LOW KHO	ꪂ LOW KHO	ꪂ	ꪂ	Glyph 10		Glyph 11	Glyph 12

Table 2.3: /x1/ and /kʰ1/

The large number of forms used for the tone 1-3 /x/ and /kʰ/ in the Northern and Jinping dialects make it difficult to analyze how these sounds are written. The glyphs used for tone 1-3 /kʰ/ are clearly modification of those for tone 1-3 /x/, formed by adding a hook in the lower right corner. This device is used consistently across all the dialects. Therefore, we wills start our analysis with tone 1-3 /x/.

Our Jinping contact says that the use of Glyph 8 is rare. It contrasts with the TAI VIET LETTER HIGH PO. It is derived from the latter character by adding the head and horizontal tail at the top of the right-hand stem. Note that in the Tai Dam dialects of the script, the HIGH PO is sometimes written with the sharp angular style of the Tai Heritage font, and sometime with the rounded bowls of the TayViet font. In the Tai Don and Jinping dialects, the same stylistic variation exists, except that the character represents the tone 4-6 /m/ instead of the tone 4-6 /p/.

On the other hand, I have found no evidence that Glyph 8 ever contrasts with TAI VIET LETTER LOW KHHO. While the two forms are quite dissimilar visually, they always have the same orthographic value.

Glyph 6 is similar to Glyph 8. Like the latter form, it is derived from the TAI VIET LETTER HIGH PO, but the curved stroke that is added to the top of the glyph to make the derivation is an unusual device, which I have not seen used elsewhere. In no case does this glyph contrast with Glyph 8, and it can be unified with it.

Glyph 7 and Glyph 9 are a curious pair. Glyph 7 does not look anything like TAI VIET LETTER LOW KHO, until one compares it to Glyph 9. The latter is clearly a derivation of TAI VIET LETTER LOW KHO, formed by reshaping the tail and the beginning loop in the body of the character until the two cross one another. When that transformation is taken to its extreme limit, one ends up with Glyph 7. So Glyph 7 and Glyph 9 can be unified with each other.

Compared to TAI VIET LETTER LOW KHO, Glyph 7 and Glyph 9 are identical to the former orthographically. But visually Glyph 7 at least is very different. Is this difference enough to warrant disunifying Glyph 7 and Glyph 9 from the current character?

The resolution of Glyph 10-Glyph 12 should parallel that of Glyph 7-Glyph 9

I classify these glyphs as follows:

- Glyph 6 and Glyph 8: Unify with each other; with respect to TAI VIET LETTER LOW KHHO, they are a distinct form with common usage—uncertain whether to unify.

- Glyph 7 and Glyph 9: Unify with each other; with respect to TAI VIET LETTER LOW KHHO, they are a distinct form with common usage—uncertain whether to unify.
- Glyph 10: Distinct form with common usage—uncertain whether to unify with TAI VIET LETTER LOW KHO.
- Glyph 11 and Glyph 12: Unify with each other; with respect to TAI VIET LETTER LOW KHO, they are a distinct form with common usage—uncertain whether to unify.

2.4 /tɕ4/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)	Central (Lai Chau)	Northern (Phong Thổ)	Jinping
/tɕ/ tone 4-6	HIGH CO	HIGH CO	Glyph 13	Glyph 14		

Table 2.4: /tɕ4/

Glyph 13 is similar to TAI VIET LETTER HIGH CO except that the head on the right side of the latter character evolves into a complete extra loop in the new glyph. Glyph 14 is similar to Glyph 13 except that the bowls are inverted. This may seem to be a major difference, except that in hand-writing it is often difficult to tell which way the bowls are pointing anyway. To illustrate this, consider the following hand-written samples:

Lò Văn Mười (Mường Cha—western)	Daiyu Fangyan Yanjiu (Jinping)	Điêu & Donaldson (Lai Chau—central)	Minot (Lai Chau—central)

Table 2.4a: hand-written examples of /tɕ4/

The sample from Lò clearly has the bowls open at the bottom, and the one from Daiyu is open at the top. The Điêu and Donaldson sample is not quite as clear as Daiyu's, but can still be identified as being open at the top. But which style does the Minot sample belong to? It seems to be more of the angular style used in the Tai Heritage font for characters such as TAI VIET LETTER HIGH PO.

This confusion between angular styles and styles with bowls occurs with other letters, and is fairly common. There are two possible explanations. Either the forms with the bowls opening upward and those with the bowls opening downward started as contrasting forms, and merged into a single angular form, loosing contrast in the process, or the angular form was the initial form, and evolved into the two forms with the bowls. The author favors the latter explanation, because 1) the angular style of characters such as the TAI VIET LETTER HIGH PO seem to be the older styles. Those with bowls seem to have evolved quite late, and are still sometimes rejected by native writers as too modern. And 2) there is no evidence of contrast in usage between those forms with the bowls opening upward and those opening downward.

While the historical derivation of the character may help to inform our choice, it is the current form and usage on which we must base our decision to unify or disunify these glyphs. Although the forms are

quite distinct, in the absence of any evidence of contrast in usage, the author cannot recommend that they be disunified from TAI VIET LETTER HIGH CO.

I classify these glyphs as follows:

- Glyph 13: Distinct form with common usage—uncertain whether to unify with TAI VIET LETTER HIGH CO.
- Glyph 14: Distinct form with common usage—uncertain whether to unify with TAI VIET LETTER HIGH CO.

2.5   /tɕʰ4/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)	Central (Lai Chau)	Northern (Phong Thổ)	Jinping
/tɕʰ/ tone 1-3	ꪊ LOW CO	ꪊ LOW CO	ꪊ	ꪊ	ꪊ	ꪊ
/tɕʰ/ tone 4-6	ꪋ HIGH CO	ꪋ HIGH CO	ꪋ	ꪋ	ꪋ Glyph 15	ꪋ

Table 2.5: /tɕʰ4/

TAI VIET LETTER HIGH CO is distinguished from TAI VIET LETTER LOW CO solely by the hook which replaces the head in the lower right corner of the former. This device is used to distinguish these two characters in most dialects of the script. But in the Northern dialect, two diagonal strokes are added to the left side of the glyph of the tone 4-6 character, as shown by Glyph 15. Presumably, this is to provide redundancy in the contrast between the two characters, thereby making text more readable. However, there is no contrast in usage between Glyph 15 and TAI VIET LETTER HIGH CO.

I classify this glyphs as follows:

- Glyph 15: Similar form with common usage—unify with TAI VIET LETTER HIGH CO

2.6 /s1/ and /s4/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)	Central (Lai Chau)	Northern (Phong Thổ)	Jinping
/s/ tone 1-3	LOW SO	LOW SO	Glyph 16			
/s/ tone 4-6	HIGH SO	LOW SO	Glyph 17			

Table 2.6: /s1/ and /s4/

- Glyph 16: Distinct form with common usage—uncertain whether to unify with TAI VIET LETTER LOW SO.
- Glyph 17: Distinct form with common usage—uncertain whether to unify with TAI VIET LETTER HIGH SO.

2.7 /kʰ4/, /ɲ4/ and /ɲ1/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)	Central (Lai Chau)	Northern (Phong Thổ)	Jinping
/kʰ/ tone 4-6	HIGH KHO	HIGH KHO				Glyph 18
/ɲ/ tone 4-6	HIGH NYO	HIGH NYO	*Glyph 19	Glyph 20	Glyph 21	
/ɲ/ tone 1-3	LOW NYO	LOW NYO	Glyph 22	Glyph 23	Glyph 24	

Table 2.7: /kʰ4/, /ɲ4/ and /ɲ1/

The script is badly fragmented for these three sounds, creating considerable confusion.

The representative glyph that was chosen for TAI VIET LETTER HIGH KHO came from the Northern dialect. This was necessary because the Central and Western dialects represent the tone 4-6 /kʰ/ with the same symbol that Tai Viet uses for TAI VIET LETTER HIGH NYO.

Looking first at Glyph 18, we see that in Jinping it is used at /kʰ4/, but in the Western dialect is used as /ɲ4/. In the latter case, it contrasts in form and usage from TAI VIET LETTER HIGH NYO.

Glyph 19 is starred, because the data from which it is taken is of poor quality. The form of Glyph 19 is not clear, and taking only its visual appearance into account, it could be a stylistic variation of either the TAI VIET LETTER HIGH NYO or Glyph 18. However, from the orthographic values of the characters,

Glyph 19 must contrast with TAI VIET LETTER HIGH NYO (unless this dialect uses one letter for two sounds). But we can and will unify Glyph 19 with Glyph 18.

Glyph 20 contrasts with TAI VIET LETTER HIGH NYO in the Central dialect, and with Glyph 18 in Jinping.

The Northern dialect has contrast between TAI VIET LETTER HIGH KHO and Glyph 21.

Glyph 22, Glyph 23, and Glyph 24 are used by the Western, Central and Jinping, and Northern dialects, respectively for /ɲ1/. Each is formed by adding the TAI VIET LETTER LOW HO to the front a dialects' corresponding symbol for /ɲ4/, except for Glyph 21, which is formed from LOW HO plus HIGH KHO. If Glyph 18, Glyph 20, and Glyph 21 are all disunified, then it will be necessary to disunify Glyph 22-Glyph 24 as well. This would result in adding six new characters to the character repertoire.

However, we have not yet shown that Glyph 18 contrasts with TAI VIET LETTER HIGH KHO, or that Glyph 20 contrasts with Glyph 21. Let's assume for a moment that each of these pairs is unified. Unifying Glyph 20 and Glyph 21 would presumably result in the unification of Glyph 23 and Glyph 24 as well. But it would also appear that the unification of the TAI VIET LETTER HIGH KHO and Glyph 18 would result in the unification of Glyph 22 and Glyph 24, since the latter is a modification of the HIGH KHO. The result would be that Glyph 22 would be unified with Glyph 23. It would seem that this would be undesirable, since Glyph 22 and Glyph 23 are derived from glyphs that contrast with each other. However, if Glyph 18 is disunified from TAI VIET LETTER HIGH KHO, Glyph 20 can be unified with Glyph 21, and Glyph 23 can be unified with Glyph 24 without any problems.

- Glyph 18: Distinct form with contrasting usage—disunify.
- Glyph 19: Probably a similar form with common usage—unify with Glyph 18.
- Glyph 20: Distinct form with contrasting usage—disunify.
- Glyph 21: Distinct form with contrasting usage—unify with Glyph 20.
- Glyph 22: Distinct form with contrasting usage—disunify.
- Glyph 23: Distinct form with contrasting usage—disunify.
- Glyph 24: Distinct form with contrasting usage—unify with Glyph 23.

2.8 /d1/ and final /-at/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)	Central (Lai Chau)	Northern (Phong Thổ)	Jinping
/d/ tone 1-3	Ω LOW DO	∩ LOW DO	Ʋ6 Glyph 25	Ʋ6	Ʋ6	Ʋ6
/-at/	˘◌Ω	�'◌∩	◌Ʋ6 Glyph 26	◌Ʋ6	◌Ʋ6	◌Ʋ6

Table 2.8: /d1/ and final /-at/

Glyph 26 must be encoded as a distinct character, because there is no equivalent character currently in the Tai Viet character repertoire.

The contrast between Glyph 25 and Glyph 26, found in all of the Tai Don dialects, parallels the contrast between TAI VIET LETTER HIGH NO (�) and TAI VIET VOWEL AN (ᴜ́). If considered in isolation, it would be possible to classify Glyph 25 as a variant glyph for TAI VIET LETTER LOW DO. But when one takes into account Glyph 26, one sees that Glyph 25 has characteristics which the LOW DO does not possess, namely the ability to form the final /-at/ by adding a crossbar to the right-hand stem.

I classify these glyphs as follows:

- Glyph 25: Distinct form with contrasting usage—disunify
- Glyph 26: Distinct form with contrasting usage—disunify

2.9 /t^{h}4/ and /l4/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)	Central (Lai Chau)	Northern (Phong Thổ)	Jinping
/t^h/ tone 1-3	૧ HIGH THO	૧ HIGH THO	૧	૧	૧	૫ Glyph 27
/t^h/ tone 4-6	૫ HIGH THO	૫ HIGH THO	૫ Glyph 28	૫	૫	૪ Glyph 29
/l/ tone 4-6	૬ HIGH LO	૪ HIGH LO	ŋ Glyph 30	ŋ	ŋ	ŋ

Table 2.9: /t^{h}4/ and /l4/

Glyph 28 is derived from TAI VIET LETTER HIGH TO. Glyph 29 is derived from Glyph 2, ɲ, the Jinping equivalent to TAI VIET LETTER HIGH TO. Since Glyph 28 and Glyph 29 never occur in contrast, since the derivation for both glyphs uses the same device, and since TAI VIET LETTER HIGH TO has already been unified with Glyph 2, Glyph 28 and Glyph 29 must be unified with each other.

Taken in isolation, it would be relatively easy to make a case for unifying Glyph 28 and Glyph 29 with TAI VIET LETTER HIGH THO as well, since they never occur with contrasting usage, and since the HIGH THO is also a derivation of the HIGH TO. (One might even argue that the hook on Glyph 28 is a "disconnected" form of the tail that is on the HIGH THO.)

The difficulty occurs with Glyph 27, which is virtually identical to the design of the HIGH THO in the TayViet font. But if Glyph 27 is unified with the HIGH THO, and Glyph 28 and Glyph 29 are also unified with the HIGH THO, then Glyph 27 becomes unified with Glyph 29. But that cannot be done. The choice is to either unify Glyph 27 with the LOW THO, or disunify Glyph 28 and Glyph 29. My preference is to unify Glyph 27 with the LOW THO, and live with the loss of visual contrast between the resulting glyph for the LOW THO in a Jinping regional font versus the HIGH THO in a Tai Dam regional font.

Glyph 30 is visually quite different from the TAI VIET LETTER HIGH LO, but never occurs with contrasting usage. Thus, the two could be unified. The difficulty that this may create arises from the

visual similarity of Glyph 30 to the HIGH THO. If someone has only a standard Tai Viet font, with none of the glyph variations, will they use the TAI VIET LETTER HIGH THO in place of Glyph 30? While at this point I recommend that Glyph 30 and the HIGH LO be unified, I am open to any arguments to the contrary.

- Glyph 27: Distinct form with common usage—unify with TAI VIET LETTER LOW THO.
- Glyph 28: Similar form with common usage—unify with TAI VIET LETTER HIGH THO.
- Glyph 29: Similar form with common usage—unify with TAI VIET LETTER HIGH THO.
- Glyph 30: Distinct form with common usage—unify with TAI VIET LETTER HIGH LO.

2.10 /p1/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)	Central (Lai Chau)	Northern (Phong Thổ)	Jinping
/p/ tone 1-6	ꪜ LOW PO	ꪜ LOW PO	ꪜ Glyph 31	ꪜ Glyph 32	ꪜ	ꪜ

Table 2.10: /p1/

The tail of Glyph 32 has more curvature than the style used in the Tai Heritage font for TAI VIET LETTER LOW PO. In Glyph 31, the tail actually has a sharp joint in it. But neither of these ever occurs in contrast to TAI VIET LETTER LOW PO, and in neither case is the stylistic variation greater than that shown between the Tai Heritage and TayViet fonts.

- Glyph 31: Similar form with common usage—unify with TAI VIET CHARACTER LOW PO.
- Glyph 32: Similar form with common usage—unify with TAI VIET CHARACTER LOW PO.

2.11 /p4/, /f4/, /m4/, and /m1/

Phonemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)	Central (Lai Chau)		Northern (Phong Thổ)	Jinping	
/p/ tone 4-6	ꪝ HIGH PO	ꪝ HIGH PO	ꪝ Glyph 33	*ꪝ	*ꪝ Glyph 34	ꪝ	*ꪝ	*ꪝ
/f/ tone 4-6	ꪠ HIGH FO	ꪠ HIGH FO	ꪠ	ꪠ Glyph 35		ꪠ	ꪠ	
/m/ tone 4-6	ꪣ HIGH MO	ꪣ HIGH MO	ꪢ Glyph 36	ꪢ		ꪢ	ꪢ	
/m/ tone 1-3	ꪢꪣ LOW MO	ꪢꪣ LOW MO	ꪣꪢ Glyph 37	ꪣꪢ		ꪣꪢ	ꪣꪢ	

Table 2.11: / p4/, /f4/, /m4/, and /m1/

None of the Tai Don dialects use the TAI VIET LETTER HIIGH MO. Instead, they use the TAI VIET LETTER HIGH PO as their symbol for the tone 4-6 /m/.

Glyph 36 should be unified with TAI VIET LETTER HIGH PO. The rounding of the bowls at the bottom of Glyph 36 is a common stylistic variation in the Tai Viet script.

Glyph 37 should be disunified from TAI VIET LETTER LOW MO. It is derived from the HIGH PO, whereas the LOW MO is derived from the HIGH PO.

Glyph 33 should not be unified with the TAI VIET LETTER HIGH PO. It has contrasting usage from Glyph 36 (and hence with the HIGH PO) in the Western and Northern dialects. Note that the head in the lower right corner of character is a significant feature, even though in some dialects it is reduced from a loop to only a point.

In the Central and Jinping dialects, I have had considerable difficulty distinguishing between the tone 4-6 /p/ and the tone 4-6 /f/. My original data for the Central dialect showed the /p4/ with a long tail, and it appeared that the contrasting feature of these two characters was the shape of the tail. With this interpretation, the TAI VIET LETTER HIGH FO would be used to write /p4/, while Glyph 35 would be disunified and used to write /f4/.

The Jinping data seemed to be ambiguous on the length of the tail for /p4/, and I assumed it to be the same of as the Central Tai Don dialect.

However, recent discussions with my Jinping contacts have caused me to examine more closely some later data on the Central dialect, and it clearly shows the /p4/ being written with a short tail. Under that interpretation, the contrast between the /p4/ and /f4/ would be the length of the tail—not its shape. Glyph 33 and Glyph 34 would be unified with each other, and Glyph 35 would be unified with the TAI VIET LETTER HIGH FO. This is the interpretation that I am currently leaning towards, but I need confirmation from Jinping and/or Tai Don experts to be sure.

- Glyph 33: Similar form with common usage—unify with Glyph 34.
- Glyph 34: Distinct form with contrasting usage—disunify.
- Glyph 35: Similar form with common usage—unify with TAI VIET LETTER HIGH FO.
- Glyph 36: Similar form with common usage—unify with TAI VIET LETTER HIGH PO.
- Glyph 37: Distinct form with common usage, but a different origin—disunify.

2.12 /j1/

Pho-nemic value	*TAI VIET LETTER… (TayViet font)	*TAI VIET LETTER… (Tai Heritage)	Western (M. Cha)		Central (Lai Chau)		Northern (Phong Thổ)		Jinping
/j/ tone 1-3	ꪐ LOW YO	✓ LOW YO	ꪐ Glyph 38	ꪐ Glyph 39	ꪐ	ꪐ Glyph 40	*ꪐ Glyph 41	*ꪐ Glyph 42	ꪐ Glyph 43
/j/ tone 4-6	ꪑ HIGH YO	✓ HIGH YO	✓		✓		✓		✓

Table 2.12: /j1/

None of new glyphs that are introduced in Table 2.12 ever occur with a contrasting usage to TAI VIET LETTER LOW YO. Thus, they could all potentially be unified as alternate glyph forms of the LOW YO. However, there are some features of these glyphs that argue against unification, especially for Glyph 43.

The TAI VIET LETTER LOW YO is derived from the TAI VIET LETTER HIGH YO by lengthening the tail of the character. Glyph 43, on the other hand, is derived from the TAI VIET LETTER HIGH YO by appending the TAI VIET LETTER LOW HO to the left side of the character. Both of these devices are in common use in the Tai Viet script to distinguish between pairs of otherwise similar characters, but they are two distinct devices that generate contrasting characters, not glyph alternates of the same character. In other words, the tail of the TAI VIET LETTER HIGH NO, 𝓊, may be lengthened

to form the TAI VIET LETTER LOW CO, 𝓊. Or, the LOW HO may be appended to the left side of the

HIGH NO to form the TAI VIET LETTER LOW NO, 𝓋𝓊. The result is two glyphs that represent two distinct characters, the LOW CO and the LOW NO. It would be inconsistent, then, to use the same devices to generate the LOW YO and Glyph 43, and then say that the LOW YO and Glyph 43 are alternate glyphs of the same character. Thus, at a minimum, I would disunify Glyph 43.

Glyph 41 and Glyph 42 are puzzling to me. First, I suspect that one is an error, and that they should both be written the same way. Second, they are derived by adding the LOW HO to the left side of the TAI VIET LETTER LOW CHO and TAI VIET LETTER LOW PHO, respectively. But I do not understand why the LOW CHO or the LOW PHO would be used as a base for forming a tone 1-3 /j/. Given the very limited amount of data from the Northern Tai Don dialect, I would favor unifying these two glyphs with Glyph 43.

Glyph 40 differs from the TAI VIET LETTER LOW YO in two respects. I am not happy with my design of this glyph, but I think it does adequately show these two differences. First, it uses a different style of the TAI VIET LETTER LOW PO for its base, and second, it uses a slanted cross bar rather than a hat. Both differences are superficial stylistic variations, and Glyph 40 should be unified with TAI VIET LETTER LOW YO.

Glyph 38 and Glyph 39 are clearly stylistic variations of each other. They are similar to Glyph 40 in that they also use a cross bar rather than the hat of the LOW YO. However, their bases are taken from the TAI VIET LETTER LOW CO rather than from the LOW PO. At this point, there should be no problem unifying these glyphs with the TAI VIET LETTER LOW YO.

- Glyph 38: Distinct form with common usage—unify with TAI VIET LETTER LOW YO.
- Glyph 39: Distinct form with common usage—unify with TAI VIET LETTER LOW YO.
- Glyph 40: Similar form with common usage—unify with TAI VIET LETTER LOW YO.
- Glyph 41: Uncertain data, similar to Glyph 43—unify with Glyph 43.
- Glyph 42: Uncertain data, similar to Glyph 43—unify with Glyph 43.
- Glyph 43: Distinct form with common usage, but with distance character features from TAI VIET LETTER LOW YO—disunify.

2.13 Comparison of Tai Dam and Tai Don vowels

2.13.1 Vowels common to all dialects
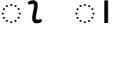
Phonemic value	*TAI VIET… (TayViet font)	*TAI VIET… (Tai Heritage)	Other common stylistic variations (found in all dialects)
/i/	VOWEL I	VOWEL I	
/ɛ/	VOWEL E	VOWEL E	
/aː/	VOWEL AA	VOWEL AA	
/aj/	VOWEL AY	VOWEL AY	
/ap/			
/ak/ and /aʔ/			
/an/	VOWEL AN	VOWEL AN	
/aŋ/			
/əw/	VOWEL AUE	VOWEL AUE	

Table 2.13 Vowels common to all dialects

Table 2.13 lists the vowels that are written the same way in all dialects. They are listed here in order to provide the reader with a complete view of the vowel system. No potentially new characters or changes in usage are represented in this section.

Note that the form used for writing short /a/ is frequently conditioned by the final consonant. For additional short /a/-final consonant sequences, see Section 2.13.4.

2.13.2 Tai Dam vowel glides and their cognates

Phonemic value	*TAI VIET… (TayViet font)	*TAI VIET… (Tai Heritage)	Tai Don (Vietnam)	Jinping		
/iᵊ/	VOWEL IA	VOWEL IA	/iᵊ/ becomes /e/ in Tai Don & Jinping Dai	/iᵊ/ becomes /e/ in Tai Don & Jinping Dai		
/e/						
/ɨᵊ/	VOWEL UEA	VOWEL UEA	/ɨᵊ/ becomes /ə/ in Tai Don & Jinping Dai	/ɨᵊ/ becomes /ə/ in Tai Don & Jinping Dai		
/ə/						
/uᵊ/	VOWEL UA	VOWEL UA	/uᵊ/ becomes /o/ in Tai Don & Jinping Dai	/uᵊ/ becomes /o/ in Tai Don & Jinping Dai		
/o/	VOWEL O	VOWEL O		Glyph 44	Glyph 45	Glyph 46

Table 2.14: Tai Dam vowel glides

As noted in section 1.2, Tai Dam has three high-to-mid vowel glides that are consistently pronounce as mid level vowels in Tai Don. Table 2.14 examines how these sounds are written. Note that the Tai Don symbol for /e/ is the same as the Tai Dam symbol for /iᵊ/. Then compare Glyph 44 with TAI VIET VOWEL UA. The resemblance between the two is evident, but Glyph 44 is slanted, or rotated, sharply to the right.

It is my opinion that Glyph 45 and Glyph 46 are also a variations of TAI VIET VOWEL UA. But in this case, instead of the character being rotated about 45 degrees, as with Glyph 44, it has been rotated a full 90 degrees.

Glyph 46 demonstrates that the rotation of the Jinping /o/, along with the other stylistic variations, can lead to a loss of visual contrast between it and other characters. Consider hand-written examples as shown by Table 2.14a.

Luo Mei Zhen *Orthographies of China's Minorities*	Minzu Wenzi Jinping Dai Character Chart	Daiyu Fangyan Yanjiu	Dao and He *Catchers on the Border…*

Table 2.14a: Hand-written Jinping symbols for /o/

The examples from Luo and Minzu Wenzi can easily by confused with the TAI VIET LETTER HIGH NO. The example from Dao and He is very similar in appearance to the TAI VIET LETTER HIGH PO.

One Jinping expert with whom I am in indirect contact considers the Jinping /o/ and the Jinping tone 4 /m/ (TAI VIET LETTER HIGH PO) to be the same character, distinguished only by position in the syllable. Thus, the word "ꪫꪫꪫ" is pronounced /mom/.

This poses a potential problem. If a font designer uses Glyph 46 as a stylistic variation of TAI VIET VOWEL UA, how does a user, looking only at the shape of the characters, tell it apart from TAI VIET LETTER HIGH PO? How does he know which one to keyboard. While a smart input method may be able to reduce the risk, it seems like the danger of data entry errors is high. On the other hand, if the TAI VIET VOWEL UA does not use Glyph 46, what is to prevent a user from using TAI VIET LETTER HIGH PO for /o/. Either way, the errors will show up when the user changes fonts, and characters that he thinks are tone 4 /m/s change to Glyph 45, or characters that he thinks are /o/s continue to look like /m/s. Thus, it may be best to unify this glyph with TAI VIET LETTER HIGH PO.

At this point, I classify these glyphs as follows:

- Glyph 44: Similar form with common usage—unify with TAI VIET VOWEL UA.
- Glyph 45: Possibly distinct form with common usage—unify with TAI VIET VOWEL UA.
- Glyph 46: Possibly distinct form with common usage—unify with either TAI VIET LETTER HIGH PO or TAI VIET VOWEL UA.

2.13.3 /u/, /ɨ/, and /ɔ/

Phonemic value	*TAI VIET… (TayViet font)	*TAI VIET… (Tai Heritage)	Tai Don (Vietnam)	Jinping	
/um/	◌ꪴꪄ	◌ꪴꪄ	ꪹ◌ꪄ	◌ꪄ or ◌ ⟶ Glyph 47	
/uŋ/	◌ꪴꪇ	◌ꪇ	ꪹ◌ꪇ	ꪹ◌ꪇ or ꪹ◌	
/u/	◌ꪴ VOWEL U	◌ꪴ VOWEL U	◌ꪴ	◌ꪴ	
/ɨ/ or /ɯ/	◌ꪳ VOWEL UE	◌ꪳ VOWEL UE	◌ꪳ ◌ꪳ ◌ꪳ Glyph 48	◌ꪳ	◌ꪳ Glyph 49
/◌ɔ#/ (open syllable)	◌ꪮ MAI KHIT	◌ꪮ MAI KHIT	◌ꪮ	◌ꪮ	
/◌ɔ-/ (closed syllable)	◌ꪮ- LOW O	◌ꪮ- LOW O	◌ꪮ	◌ꪮ	

Table 2.15: /u/, /ɨ/, and /ɔ/

Tai Don uses TAI VIET VOWEL O to write /u/ in front of final /-m/ and final /-ŋ/. The stylistic variations for /ɨ/, all of which are labeled Glyph 48, are common—the same variations are found in Tai Dam writing. Finally, it uses TAI VIET LETTER LOW O to write /ɔ/ in all contexts. None of these practices introduces a new character.

Jinping adds a couple variations to how the write these vowels. For writing /u/ before final /-m/, it uses Glyph 47, a diminutive combining form of the TAI VIET LETTER LOW O, written over the upper left corner of the consonant. Since this glyph is a combining mark instead of a spacing character, it cannot be unified with TAI VIET LETTER LOW O.

In Jinping, the final consonant is optional for both /-um/ and /-uŋ/.

Glyph 49 occurred in one Jinping data sample. Although it bears a visual similarity to TAI VIET LETTER HIGH MO (/p4/ in Jinping), I believe it is probably related to the Tai Daeng symbols for /ɨ/:

◌ꪳ Tai Daeng /ɨː/

◌ꪳ Tai Daeng /ɨ/

Note that Tai Daeng has vowel length contrast, and thus has two contrasting symbols for /ɨ/. The Jinping symbol corresponds closely to the first Tai Daeng symbol shown here. It is not clear how frequently this glyph is used for /ɨ/. At this point it should be unified with TAI VIET VOWEL UE.

- Glyph 47: Distinct form with a unique combining class—disunify.
- Glyph 48: Similar forms with common usage—unify with TAI VIET VOWEL UE.
- Glyph 49: Distinct form with common usage—unify with TAI VIET VOWEL UE.

2.13.4   Short /a/-final consonant sequences

Phonemic value	*TAI VIET… (TayViet font)	*TAI VIET… (Tai Heritage)	Tai Don (Vietnam)	Jinping
/at/	꦳ꦵ	꦳ꦵ	ꦵꦶ See Glyph 26	ꦵꦶ
/am/	ꦵ VOWEL AM	ꦵ VOWEL AM	ꦵꦖ	ꦵꦖ or ꦵ
/aw/	(ꦵ꧀	ꦵꦵ꧀	ꦵꦵ꧀	(ꦵ(

Table 2.16: Short /a/-final consonant sequences

Glyph 26 was dealt with in Section 2.8, where it was shown that it needs to be encoded as a distinct character.

Tai Don and Jinping use the same combining mark for the final sequence /-am/ (TAI VIET VOWEL AM) that Tai Dam uses, but in Tai Don the symbol for /m/ (TAI VIET LETTER HIGH PO) may be retained in the spelling, while in Tai Dam it is omitted. Stylistically, the symbol consists of three to five components, which may be dots or short lines. Compare the styles of the TayViet and Tai Heritage fonts.

The final sequence /aw/ is written with a split digraph in all three dialects. But whereas Tai Dam and Tai Don use the digraph TAI VIET VOWEL UEA + TAI VIET VOWEL AA, Jinping uses TAI VIET VOWEL UEA + TAI VIET VOWEL UEA.

Neither the spelling for /am/ nor /aw/ require new characters or glyphs.

2.13.5 /aː/ tone 3 and /ɔ/ tone 4

Phonemic value	*TAI VIET… (TayViet font)	*TAI VIET… (Tai Heritage)	Tai Don (Vietnam)	Jinping
/aː#/ tone 3			⠲ Glyph 50	⠾ Glyph 51
/ɔ/ tone 4				⠢ Glyph 52

Table 2.17: /aː/ tone 3 and /ɔ/ tone 4

There are currently no equivalent characters for Glyph 50 and Glyph 52 in the Tai Viet character repertoire. These two need to be encoded as distinct characters.

It may be asked whether Glyph 46 can be unified with the existing COMBINING DOT BELOW, U + 0323. Given the fact that Glyph 46 may have to interact with Tai Viet line-breaking algorithms, and that there may be variations between dialects in how Glyph 46 is positioned, I do not think this would be wise.

Glyph 51 is a stylistic variation of Glyph 50.

- Glyph 50: Distinct form with contrasting usage—disunify.
- Glyph 51: Similar form with common usage—unify with Glyph 50.
- Glyph 52: Distinct form with contrasting usage—disunify.

2.14 Tone marks

Phonemic value	Various borrowings for Tai Viet tone marks		
tones 2 & 5	◌́ / ◌́ MAI EK	ℓ / e MAI NUENG	◌́ Glyph 53
tones 3 & 6	◌̌ / ◌̌ MAI THO	ꦶ/꧀ MAI SONG	◌̉ Glyph 54

Table 2.18: Tone marks

Tone marks are relatively new to Tai Viet writing, having been brought into the script only in the last 50 years. But rather than create arbitrary marks for writing tone, users of the script looked to other writing systems for the source of the marks they used.

The Tai Dam community in the U.S. chose to borrow the Mai Ek and Mai Tho from Lao.

In Vietnam, the Tai community considered two different systems of marks borrowed from Vietnamese. The first system employed the Vietnamese Sắc and Hỏi tone symbols. The Sắc, written with the COMBINING ACUTE ACCENT, marks tones 2 and 5 in unchecked syllables. The Hỏi, written with

the COMBINING HOOK ABOVE, marks tones 3 and 6 in unchecked syllables. See Glyph 53 and Glyph 54 in Table 2.18.

The second Vietnamese tone system explored by the Tai community in Vietnam employs the characters which are now called Mai Nueng and Mai Song. I believe they probably originated as stylized versions for the Latin letters 'e' and 'j', which have been used in Vietnamese telegraphy to mark the Sắc and Hỏi tones.

As it currently stands, the use of Mai Nueng and Mai Song seem to be the most popular choices in Vietnam. The Jinping have chosen the Sắc and Hỏi symbols, and the Tai community in the U.S. strongly prefers the Mai Ek and Mai Tho.

Two of the three have been encoded in the current Tai Viet character repertoire, as TAI VIET TONE MAI EK and TAI VIET TONE MAI THO, on the one had, and as TAI VIET TONE MAI NUENG and TAI VIET TONE MAI SONG on the other hand. It was not possible to unify these, since they have different combining classes (one set being made up of combining marks, and the other set of spacing characters).

However, it may be possible to consider Glyph 53 and Glyph 54 to be alternative glyphs for the MAI EK and MAI THO, respectively.

- Glyph 53: Distinct form with common usage—unify with TAI VIET TONE MAI EK.
- Glyph 54: Distinct form with common usage—unify with TAI VIET TONE MAI THO.

2.15 Position of combining marks

Looking at the script across all of its dialects, there is considerable fluidity in where combining marks are placed when a syllable contains both an initial and final consonant. On a page from one hand-written Tai Don document that I have, there are 32 instances of combining vowels in syllables with two consonants. Of these, 23 occur over the initial consonant, and 9 occur over the final consonant. With one exception (see below), the position does not seem to be conditioned by context. In the word ꪝꪮ, /səj/, the vowel is positioned over the initial consonant. In the word ꪚꪮ, /dəj/, the same vowel is placed over the final consonant.

The one exception where the context does condition the positioning of the combining mark is found in Tai Dam, which has the vowel digraphs ꪲ◌ and ꪱ◌ (/ə/ and /e/, respectively). When a word has these digraphs, the combining portion of the vowel is always placed over the initial consonant.

In the Jinping region, there is a strong preference for placing the combining *vowel* mark over the final consonant, and the combining *tone* mark over the initial consonant. I have seen just one author from Vietnam, a deceased professor from Hanoi National University, who followed this practice.

2.16 Ideographs

The five most common ideographs are contained in the current repertoire. Jinping uses different glyphs for some of these, it may vary the meaning of some, and has additional ideographs for 'horse' and 'moon'. The ideographs used by the dialects of Lai Chau Province have not been adequately studied. More work in this area is needed.

3 Summary

I have examined 54 glyphs used by the various dialects of Tai Don and Jinping Dai, and recommend that 13 of these glyphs, listed in Table 3.1, be disunified. A chart showing of all the characters that would then be contained in the Tai Viet repertoire, along with all of the anticipated regional alternative glyphs for each character, is listed in Table 3.2.

My recommendations are based on a conservative approach which minimizes the number of characters encoded, but which consequently leaves some of the current characters with a large number of alternative glyph forms. These alternative forms would be made available to the user either through user-selectable features in a font universal font, or as fonts designed for the specific style of a given region.

Some readers may wish to treat the current Tai Viet repertoire as a phonemic encoding. Then, each regional form may be a alternate glyph for the corresponding character. I do not think this will work. It would create situations where a glyph such as ꪣ would be an alternate glyph for two different characters—the Tai Dam form of the TAI VIET LETTER HIGH PO on the one hand, and the Tai Don form of the TAI VIET LETTER HIGH MO on the other hand. This would result in an encoding that is interpreted by the font—just what Unicode is supposed to eliminate—and would produce great confusion.

Other readers may wish to argue for the disunification of other glyphs, or for the disunification of the entire Tai Don/Jinping script from Tai Viet. *All such contributions and advice are earnestly sought.*

An alternative approach would be to focus on adding only the characters needed for Jinping Dai, and see whether that adequately covered the classical Tai Don of Lai Chau Province.

I was surprised to find several features of Tai Daeng characters in the Jinping Dai dialect. Given that Jinping Dai is the most northern dialect, and Tai Daeng the most southern, I expected these two forms of the script to be the most different. I had once thought it would be sufficient to compare Tai Daeng writing only to Tai Dam, but given these latest findings, I think it would be desirable to compare it with the Tai Don and Jinping as well.

Table 3.1: Glyphs recommended for disunification

Glyph	Form	Description
Glyph 3	ꪀ	tone 4-6 /k/ in Central, Northern, and Jinping
Glyph 18	ꪁ	tone 4-6 /kʰ/ in Jinping; tone 4-6 /ɲ/ in Western
Glyph 20	ꪂ	tone 4-6 /ɲ/ in Central and Jinping
Glyph 22	ꪃ	tone 1-3 /ɲ/ in Western
Glyph 23	ꪄ	tone 1-3 /ɲ/ in Central and Jinping

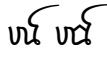|		
Glyph 25	ᖴ	tone 1-3 /d/ in all dialects
Glyph 26	ᖴ	final /-at/ in all dialects
Glyph 34	ᴡᖯ	tone 4-6 /p/ in Central and Jinping; alternative glyph ɯ is used in Western and Northern
Glyph 37	℧ᴡ	tone 1-3 /m/ in all dialects
Glyph 43	℧ᶹ	tone 1-3 /j/ in Jinping
Glyph 47	ℾ	/um/ in Jinping
Glyph 50	⦂	/aː/ tone 3 in Western, Central, and Northern; alternative glyph ⦂ is used in Jinping
Glyph 52	∘	/ɔ/ tone 4 in Jinping

Table 3.2: Cross-dialect character table

An asterisk (*) indicates characters that are not used in the Tai Dam language, but are part of the current Tai Viet character repertoire. The alternate glyph forms for each character are labeled beneath the glyph. There are several additional alternate glyphs which were not discussed in the above text. I do not consider them candidates for disunification, but font developers will want to consider them for inclusion in regional fonts. These have been marked with a double asterisk (**).

Name:	Existing Tai Viet (Tai Dam based)			Western Tai Don		Central Tai Don		Northern Tai Don		Jinping Dai	
Existing Unicode character name, or glyph number of proposed character	TayViet glyph	Tai Heritage glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)
TAI VIET LETTER LOW KO	ꪀ	ꪀ	/k1/	ꪀ	/k1/	ꪀ	/k1/	ꪀ	/k1/	ꪀ	/k1/
TAI VIET LETTER HIGH KO	ꪁ	ꪁ	/k4/	ꪁ	/k4/	ꪁ	/x4/	—	–	—	–
New char based on GLYPH 3 (*TAI DON LETTER HIGH KO)	—	—	–	—	–	ꪂ	/k4/	ꪂ	/k4/	ꪂ	/k4/
TAI VIET LETTER LOW KHO	ꪂ	ꪂ	/kʰ1/*	ꪂ	/kʰ1/	ꪂ	/kʰ1/	§ (Glyph 10)	/kʰ1/	ꪂ (Glyph 12)	/kʰ1/
TAI VIET LETTER HIGH KHO	ꪃ	ꪃ	/kʰ4/*	—	–	—	–	ꪃ	/kʰ4/	—	–
TAI VIET LETTER LOW KHHO	ꪄ	ꪄ	/x1/	ꪄ	/x1/	ꪄ	/x1/	§ (Glyph 7)	/x1/	ꪄ (Glyph 9)	/x1/
TAI VIET LETTER HIGH KHHO	ꪅ	ꪅ	/x4/	ꪅ (Glyph 4)	/x4/	—	–	ꪅ	/x4/	ꪅ (Glyph 5)	/x4/
TAI VIET LETTER LOW GO	ꪆ	ꪆ	/g1/*	GO and RO are not part of the normal Tai phonology, but were added to the script for use in transcription of Vietnamese names.						—	–
TAI VIET LETTER HIGH GO	ꪇ	ꪇ	/g4/*							—	–

Name:	Existing Tai Viet (Tai Dam based)			Western Tai Don		Central Tai Don		Northern Tai Don		Jinping Dai	
Existing Unicode character name, or glyph number of proposed character	**TayViet glyph**	**Tai Heritage glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**
TAI VIET LETTER LOW NGO	ขึ	ขึ	/ŋ1/	ข6**	/ŋ1/	ข6**	/ŋ1/	ข6**	/ŋ1/	ขb**	/ŋ1/
TAI VIET LETTER HIGH NGO	၆	6	/ŋ4/	6	/ŋ4/	6	/ŋ4/	6	/ŋ4/	6	/ŋ4/
TAI VIET LETTER LOW CO	ป	ง/ป	/tɕ1/	ป	/tɕ1/	ป	/tɕ1/	ป	/tɕ1/	ป	/tɕ1/
TAI VIET LETTER HIGH CO	ๅ	ๆ	/tɕ4/	ᄁ Glyph 13	/tɕ4/	ᄳ Glyph 14	/tɕ4/	ᄁ Glyph 13	/tɕ4/	ᄳ Glyph 14	/tɕ4/
TAI VIET LETTER LOW CHO	ผึ	ผึ	/tɕʰ1/	ผึ	/tɕʰ1/	ผึ	/tɕʰ1/	ผึ	/tɕʰ1/	ผึ	/tɕʰ1/
TAI VIET LETTER HIGH CHO	ยึ	ยึ	/tɕʰ4/	ยึ	/tɕʰ4/	ยึ	/tɕʰ4/	ขึ Glyph 15	/tɕʰ4/	ยึ	/tɕʰ4/
TAI VIET LETTER LOW SO	x	x	/s1/	พ Glyph 16	/s1/	พ Glyph 16	/s1/	พ Glyph 16	/s1/	พ Glyph 16	/s1/
TAI VIET LETTER HIGH SO	ถ	ถ	/s4/	พ Glyph 17	/s4/	พ Glyph 17	/s4/	พ Glyph 17	/s4/	พ Glyph 17	/s4/
TAI VIET LETTER LOW NYO	ฃฎ	ขฎ	/ɲ1/	—	–	—	–	—	–	—	–
New char based on GLYPH 22 (*TAI DON LETTER LOW NYO)	—	—	–	ขฎ	/ɲ1/	—	–	—	–	—	–
New char based on GLYPH 23 (*JINPING LETTER LOW NYO)	—	—	–	—	–	ขฎ	/ɲ1/	ขพ3 Glyph 24	/ɲ1/	ขฎ	/ɲ1/

Jim Brase   May 5, 2008

Name:	Existing Tai Viet (Tai Dam based)			Western Tai Don		Central Tai Don		Northern Tai Don		Jinping Dai	
Existing Unicode character name, or glyph number of proposed character	TayViet glyph	Tai Heritage glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)
TAI VIET LETTER HIGH NYO	[glyph]	[glyph]	/ɲ4/	[glyph]	/kʰ4/	[glyph]	/kʰ4/	—	–	—	–
New char based on GLYPH 18 (*TAI DON LETTER HIGH NYO)	—	—	–	[glyph]	/ɲ4/	—	–	—	–	[glyph]	/kʰ4/
New char based on GLYPH 20 (*JINPING LETTER HIGH NYO)	—	—	–	—	–	[glyph]	/ɲ4/	[glyph] Glyph 21	/ɲ4/	[glyph]	/ɲ4/
TAI VIET LETTER LO DO	[glyph]	[glyph]	/d1/	—	–	—	–	—	–	—	–
New char based on GLYPH 25 (*TAI DON LETTER LOW DO)	—	—	–	[glyph]	/d1/	[glyph]	/d1/	[glyph]	/d1/	[glyph]	/d1/
TAI VIET LETTER HIGH DO	[glyph]	[glyph]	/d4/	[glyph]	/d4/	[glyph]	/d4/	[glyph]	/d4/	[glyph]	/d4/
TAI VIET LETTER LOW TO	[glyph]	[glyph]	/t1/	[glyph]	/t1/	[glyph]	/t1/	[glyph]**	/t1/	[glyph]**	/t1/
TAI VIET LETTER HIGH TO	[glyph]	[glyph]	/t4/	[glyph] / [glyph]	/t4/	[glyph]	/t4/	[glyph] Glyph 2	/t4/	[glyph] Glyph 2	/t4/
TAI VIET LETTER LOW THO	[glyph]	[glyph]	/tʰ1/	[glyph]	/tʰ1/	[glyph]	/tʰ1/	[glyph]	/tʰ1/	[glyph] Glyph 27	/tʰ1/
TAI VIET LETTER HIGH THO	[glyph]	[glyph]	/tʰ4/	[glyph] Glyph 28	/tʰ4/	[glyph] Glyph 28	/tʰ4/	[glyph] Glyph 28	/tʰ4/	[glyph] Glyph 29	/tʰ4/
TAI VIET LETTER LOW NO	[glyph]	[glyph]	/n1/	[glyph]	/n1/	[glyph]	/n1/	[glyph]	/n1/	[glyph]	/n1/
TAI VIET LETTER HIGH NO	[glyph]	[glyph]	/n4/	[glyph]	/n4/	[glyph]	/n4/	[glyph]	/n4/	[glyph]	/n4/

Jim Brase May 5, 2008

Name:	Existing Tai Viet (Tai Dam based)			Western Tai Don		Central Tai Don		Northern Tai Don		Jinping Dai	
Existing Unicode character name, or glyph number of proposed character	TayViet glyph	Tai Heritage glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)
TAI VIET LETTER LOW BO	ꪚ	√	/b1/	√	/b1/	√	/b1/	√	/b1/	√	/b1/
TAI VIET LETTER HIGH BO	ꪛ	w	/b4/	w	/b4/	w	/b4/	w	/b4/	w	/b4/
TAI VIET LETTER LOW PO	ꪜ	√	/p1/	∫ Glyph 31	/p1/	√ Glyph 32	/p1/	∫ Glyph 31	/p1/	∫ Glyph 31	/p1/
TAI VIET LETTER HIGH PO	ꪝ	w	/p4/	ω Glyph 36	/m4/	w	/m4/	ω Glyph 36	/m4/	w	/m4/
New char based on GLYPH 33 (*TAI DON LETTER HIGH PO)	—	—	–	ω Glyph 33	/p4/	ω Glyph 34	/p4/	ω Glyph 33	/p4/	ω Glyph 34	/p4/
TAI VIET LETTER LOW PHO	ꪞ	ꪺ	/pʰ1/	ꪺ	/f1/	ꪺ	/f1/	ꪺ	/f1/	ꪺ	/f1/
TAI VIET LETTER HIGH PHO	ꪟ	ꪽ	/pʰ4/	ꪽ	/pʰ4/	ꪽ	/pʰ4/	ꪽ	/pʰ4/	ꪽ**	/pʰ4/
TAI VIET LETTER LOW FO	ꪠ	ꪺ	/f1/	ꪺ	/pʰ1/	ꪺ	/pʰ1/	ꪺ	/pʰ1/	ꪺ	/pʰ1/
TAI VIET LETTER HIGH FO	ꪡ	ꪼ	/f4/	ꪼ	/f4/	ꪼ	/f4/	ꪼ	/f4/	ꪼ	/f4/
TAI VIET LETTER LOW MO	ꪢ	ꪢ	/m1/	—	–	—	–	—	–	—	–
New char based on GLYPH 36 (*TAI DON LOW LETTER LOW MO)	—	—	–	ꪤw	/m1/	ꪤw	/m1/	ꪤw	/m1/	ꪤw	/m1/
TAI VIET LETTER HIGH MO	ꪣ	ꪣ	/m4/	—	–	—	–	—	–	—	–

Name:	Existing Tai Viet (Tai Dam based)			Western Tai Don		Central Tai Don		Northern Tai Don		Jinping Dai	
Existing Unicode character name, or glyph number of proposed character	**TayViet glyph**	**Tai Heritage glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**
TAI VIET LETTER LOW YO	ญ	√	/j1/	ꪤ ꪤ Glyph 38, Glyph 39	/j1/	ꪤ ꪤ Glyph 39, Glyph 40	/j1/	—	–	—	–
New char based on GLYPH 42 (*JINPING LETTER LOW YO)	—	—	–	—	–	—	–	ꪤ ꪤ Glyph 41, Glyph 42	/j1/	ꪤ	/j1/
TAI VIET LETTER HIGH YO	ญ	√	/j4/	√	/j4/	√	/j4/	√	/j4/	√	/j4/
TAI VIET LETTER LO RO	�	ກ	/r1/	GO and RO are not part of the normal Tai phonology, but were added to the script for use in transcription of Vietnamese names.				—	–		
TAI VIET LETTER HIGH RO	ຶ	ກ	/r4/					—	–		
TAI VIET LETTER LOW LO	ຖ	ຖ	/l1/	ຖ	/l1/	ຖ	/l1/	ຖ	/l1/	ຖ	/l1/
TAI VIET LETTER HIGH LO	ຽ	Y	/l4/	ŋ Glyph 30	/l4/	ŋ Glyph 30	/l4/	ŋ Glyph 30	/l4/	ŋ Glyph 30	/l4/
TAI VIET LETTER LOW VO	ຫ	ꪫ	/v1/	ꪫ	/v1/	ꪫ	/v1/	ꪫ	/v1/	ꪫ	/v1/
TAI VIET LETTER HIGH VO	ງ	ວ	/v4/	ວ	/v4/	ວ	/v4/	ວ	/v4/	ວ	/v4/
TAI VIET LETTER LO HO	ຫ	ຫ	/h1/	ຫ	/h1/	ຫ	/h1/	ຫ	/h1/	ຫ	/h1/
TAI VIET LETTER HIGH HO	ຫ	ຫ	/h4/	ຫ	/h4/	ຫ	/h4/	ຫ	/h4/	ຫ	/h4/

Name:	Existing Tai Viet (Tai Dam based)			Western Tai Don		Central Tai Don		Northern Tai Don		Jinping Dai	
Existing Unicode character name, or glyph number of proposed character	TayViet glyph	Tai Heritage glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)
TAI VIET LETTER LOW O	ꪮ	Ɵ	/ʔ1/	Ɵ	/ʔ1/	Ɵ	/ʔ1/	Ɵ	/ʔ1/	Ɵ	/ʔ1/
TAI VIET LETTER HIGH O	ꪯ	ꪯ	/ʔ4/	ꪯ	/ʔ4/	ꪯ	/ʔ4/	ꪯ	/ʔ4/	ꪯ**	/ʔ4/
TAI VIET MAI KANG	◌̆	◌̆	/a/	◌̆	/a/	◌̆	/a/	◌̆	/a/	◌̆	/a/
TAI VIET VOWEL AA	ꪱ	ꪱ	/aː/	ꪱ / ꪺ	/aː/	ꪱ / ꪺ	/aː/	ꪱ / ꪺ	/aː/	ꪱ / ı	/aː/
New char based on GLYPH 49 (*TAI DON VOWEL LOW TONE AA)	—	—	–	◌᪷	/aː#/ tone 3	◌᪷	/aː#/ tone 3	◌᪷	/aː#/ tone 3	◌᪷ Glyph 51	/aː#/ tone 3
TAI VIET VOWEL I	◌̂	◌̂	/i/	◌̂	/i/	◌̂	/i/	◌̂	/i/	◌̂	/i/
TAI VIET VOWEL UE	◌̒	◌̒	/ɨ/	◌̒ / ◌̒ Glyph 48	/ɨ/	◌̒ / ◌̒ Glyph 48	/ɨ/	◌̒ / ◌̒ Glyph 48	/ɨ/	◌̒ / ◌̒ʷ Glyph 49	/ɨ/
TAI VIET VOWEL U	◌̦	◌̦	/u/	◌̦	/u/	◌̦	/u/	◌̦	/u/	◌̦	/u/
New char based on GLYPH 46 (*JINPING VOWEL UM)	—	—	–	—	–	—	–	—	–	◌̒	/um/
TAI VIET VOWEL E	ꪹ	ꪹ	/ɛ/	ꪹ	/ɛ/	ꪹ	/ɛ/	ꪹ	/ɛ/	ꪹ	/ɛ/
TAI VIET VOWEL O	ꪻ	ꪻ	/o/	ꪻ	/u/ before /m/ or /ŋ/	ꪻ	/u/ before /m/ or /ŋ/	ꪻ	/u/ before /m/ or /ŋ/	ꪻ	/u/ before /m/ or /ŋ/
TAI VIET MAI KHIT	◌̌	◌̀	/ɔ#/	◌̀	/ə/	◌̀	/ə/	◌̀	/ə/	◌̀	/ə/

Jim Brase   May 5, 2008

Name:	Existing Tai Viet (Tai Dam based)			Western Tai Don		Central Tai Don		Northern Tai Don		Jinping Dai	
Existing Unicode character name, or glyph number of proposed character	**TayViet glyph**	**Tai Heritage glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**	**Glyph**	**usage (IPA)**
New char based on GLYPH 51 (*JINPING VOWEL HIGH TONE OH)	—	—	–	—	–	—	–	—	–	◌̇	/ɔ/ tone 4
TAI VIET VOWEL IA	◌	◌	/iə/	◌	/e/	◌	/e/	◌	/e/	◌	/e/
TAI VIET VOWEL UEA	ꪵ	ꪵ	/ɨə/	used only in the digraph for /aw/: ꪵ◌ꪮ						/aw/ = ꪵ◌ꪵ	
TAI VIET VOWEL UA	ꪊ	ꪊ	/ua/	ꪊ / ꪋ Glyph 44	/o/	ꪊ / ꪋ Glyph 44	/o/	ꪊ / ꪋ Glyph 44	/o/	ꪋ Glyph 45	/o/
TAI VIET VOWEL AUE	ꪎ	ꪎ	/əw/	ꪎ	/əw/	ꪎ	/əw/	ꪎ	/əw/	ꪎ	/əw/
TAI VIET VOWEL AY	ꪼ	ꪼ	/aj/	ꪼ	/aj/	ꪼ	/aj/	ꪼ	/aj/	ꪼ	/aj/
New char based on GLYPH 26 (*TAI DON VOWEL AT)	—	—		ꪶ	/at/	ꪶ	/at/	ꪶ	/at/	ꪶ	/at/
TAI VIET VOWEL AN	ꪯ	ꪯ	/an/	ꪯ	/an/	ꪯ	/an/	ꪯ	/an/	ꪯ	/an/
TAI VIET VOWEL AM	◌	◌	/am/	◌	/am/	◌	/am/	◌	/am/	◌	/am/
TAI VIET TONE MAI EK	◌	◌	tones 2,5	◌	tones 2,5	◌	tones 2,5	◌	tones 2,5	◌̇ Glyph 53	tones 2,5
TAI VIET TONE MAI NUENG	ꪷ	ꪷ	tones 2,5	ꪷ	tones 2,5	ꪷ	tones 2,5	ꪷ	tones 2,5	—	–
TAI VIET TONE MAI THO	◌	◌	tones 3,6	◌	tones 3,6	◌	tones 3,6	◌	tones 3,6	◌̊ Glyph 54	tones 3,6

Name:	Existing Tai Viet (Tai Dam based)			Western Tai Don		Central Tai Don		Northern Tai Don		Jinping Dai	
Existing Unicode character name, or glyph number of proposed character	TayViet glyph	Tai Heritage glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)	Glyph	usage (IPA)
TAI VIET TONE MAI SONG	)	)	tones 3,6	)	tones 3,6	)	tones 3,6	)	tones 3,6	—	–
*JINPING SYMBOL HORSE	—	—									
*JINPING SYMBOL MOON	—	—									
TAI VIET SYMBOL KON			/kon^4/								
TAI VIET SYMBOL NUENG			/niŋ^5/								
TAI VIET SYMBOL SAM			repeat symbol								
TAI VIET SYMBOL HO HOI			onset of poetry								
TAI VIET SYMBOL KOI KOI			end of stanza								

4   Bibliography

___. "Các mẫu tự Thái ở miền Tây Bắc Việt Nam." Internet:
http://www.evertype.com/standards/tai/viet-thai-samples.pdf

Cheuk, Adrian. April 2008. Personal Correspondence.

Điêu Chính Nhìm and Jean Donaldson. 1970. *Păp San Khhãm Pák Tãy-Keo-Eng* (Ngũ-Vụng Thái-Việt-Anh, Tai-Vietnamese-English Vocabulary). Saigon.

Fippinger, Jay and Dorothy. 1970. "Black Tai Phonemes with reference to White Tai." *Anthropological Linguistics* 12.3: 83-97.

Gedney, William J. 1989. "A Comparative Sketch of White, Black and Red Tai." *Selected Papers on Comparative Tai Studies,* Michigan Papers on South and Southeast Asia no. 29, Center for South and Southeast Asia Studies, The University of Michigan.

Lo Văn Mươi (ꪶꪜ ꪣꪳ ꪨꪮ). 1966. Ép Sư ꞌTáy Piên Peng. (ꪶꪚ ꪀꪫ ꪜ ꪚꪸ, Learning the Revised Tai Alphabet.)

Minot, Georges. 1940. "Dictonnaire Tẏy Blanc-Français." *Bulletin de l'Ecole d'Extrême-Oriente*, t. XL.