To: Script Ad Hoc / UTC
From: Jim Brase
Re: Response to Eiso Chan "Comments on the Tai Don or Tai Khao encoding"
Date: May 20 2022

Response to Eiso Chan's "Comments on the Tai Don or Tai Khao encoding"

1.1   The 1955 unification was controversial and not universally accepted. Nevertheless, I know of no contrast in usage between the forms in question, and I think they can be regarded as presentation forms of the same underlying character. I.e. unify them.

1.2   Regarding the usage of U+AA86, U+AA87, U+AAA6, and U+AAA7: Eiso is correct. These characters do not represent sounds that are native to Tai Dam or Tai Don. They were included in the Tai Viet proposal at the explicit request of Tai committee from Vietnam that was partnering with me in the proposal, and who wished to use them for transliterating Vietnamese place names.

In Table 1.2.1, it is not clear to me where the glyphs in the "Submitted" column come from. Kushim did not include any glyphs equivalent to these four characters in L2/20-207.

I am surprised by the glyphs used in the Oswald font (see Table 1.2.1) for U+AA86 and U+AA87. Oswald's U+AA86 is a stylistic variant of U+AA97. (See the Tai Heritage Pro font.) His U+AA87 is a stylistic variation of U+AA85. (See "Các Mẫu Tự Thái Ở Miền Tây Bắc Việt Nam." http://www.evertype.com/standards/tai/viet-thai-samples.pdf.) There are many dialects of the Tai Viet script. It is difficult enough as it is to sort through all the glyph variants and determine where a glyph in one dialect conflicts with the usage of the same or similar glyph in another dialect. Font developers must use caution, and not introduce new usages for existing glyphs, lest they compound the problem.

Regarding the "Submitted" glyphs for U+AAA6 and U+AAA7 in Table 1.2.1, I suspect they were derived from U+0EA3 and U+0EAE.

Regarding character names for the voiceless velar consonants, the names used in the Tai Viet block (U+AA80..U+AA85) were derived from the Roman alphabet used in the Tai Don vocabulary by Điêu and Donaldson, 1970. That alphabet, in turn, was derived from Vietnamese. But Vietnamese and most Tai languages have only a two-way contrast between voiceless velar consonants, with K representing /k/, and KH representing either /kʰ/ or /x/, depending on the language and dialect. But Tai Don has a three-way contrast. So Điêu and Donaldson used K for /k/, KH for /kʰ/, and KHH for /x/.

2.   Syllable structure and labialization

Eiso includes a list of vowels that he found following labialized consonants. (Figure 2.4) Fippinger 1970 makes the following comment on the matter:

> "… there is considerable limitation on the number of vowels occurring after the three labialized consonants /kw, xw, ŋw/. The small number of words with labialized consonants in our data precludes a firm statement on the co-occurrence restriction, but it does seem certain that no back vowels occur following these three consonants. To date, /aa/ has been observed with all three, /ɛ/ with /kw/ and /xw, and /i/ with /xw/."

However, from looking through Baccam et. al., I also find /i/ after /kʷ/.

Also re Fig. 2.4, I had not previously seen labialization used with ꪫ, U + AAAD.

Re the spelling of /kiw/ vs /kʷi/ and /kɛw/ vs /kʷɛ/ (Fig 2.5): My description is based on Baccam, et al, 1989, as shown in the following sample entries:

ꪔ̂ꪫ	ꪹꪀꪮ	kíu	narrow
�ꪔꪫ	ꪵꪀꪮ	Keo	Vietnamese
ꪵꪔ̌ꪫ	ꪵꪁꪮ	kẻo	scissors
ꪀꪮ̂	ꪀꪮ̂	quí	tax
ꪵꪀꪮ	ꪵꪀꪮ̂	'que	cinnamon

Sample vocabulary words from Baccam, et. al.

This is the practice that I observed among Tai Dam immigrants living in the United States, which included the authors of Baccam, et. al. Among this group, I never observed the use of a final ꪫ to disambiguate between /kiw/ and /kʷi/. However, I have noted the use of final ꪫ in data received from the Tai community in Vietnam. This may represent a dialect difference, or a diachronic change in the spelling.

It should be noted that up until about 50 years ago, the Tai Viet script was almost exclusively hand written. In addition, there were very few dictionaries, if any. Thus one can expect some variation in both spelling and style of writing.

While we are on the subject of labialization, it should be noted that the velar nasals can also be labialized. E.g.

ꪹꪉꪱ /ŋʷa:⁵/ 'wild fig'

Regarding whether a consonant cluster can have a consonant other than ວ in the second position (see Figure 2.6 and Eiso's comments about it): Such structures do not exist in the Tai Dam or Tai Don phonology. However, Fippinger 1970 states:

> "Most words are monosyllabic. There are, however, a number of disyllabic words composed of an unstressed initial syllable of the pattern CV (V being a central vowel), followed by a stressed final syllable. $sə^1$-$laŋ^1$ <u>back</u> (body part)."

We have observed that some writers will write the vowel of the unstressed syllable, and some will omit that vowel. Omitting the vowel would create a consonant cluster in the written form. Also, there is the potential to create a written cluster by omitting an unstressed vowel in a multisyllable loan word.

3. Punctuation.

My knowledge of the use of these symbols is limited.

4. Suggestions.

1) My original intention was to include both Tai Dam and Tai Don characters under Tai Viet. However, I lacked access to Tai Don writers, and my boss pulled me off of the project before I could complete my research on the Tai Don.

2) The names are always tricky. Many names, spellings, and groupings have been used through the years. Linguistically, Tai Dam and Tai Don have different tone systems, which in Tai linguistics is considered adequate to distinguish them as separate languages. But my understanding is that the Vietnamese government does not distinguish between them, and groups both together as Tai Viet. Consequently, our intent was to let Tai Dam and Tai Don be used for the names of the languages, and use Tai Viet for the name of the combined script.

4) "Consider if it's suitable to use multi-columns form in the Tai Viet block." I've been retired for over 2 years, and my Unicode knowledge is getting a little rusty. I'm not familiar with "multi-columns form". But yes, if the Tai Don characters are unified with the existing Tai Viet block, I think some way of showing the glyph variations would be helpful.

References:

Baccam Don, Baccam Faluang, Baccam Hung, and Dorothy Fippinger. 1989. ꪝꪰꪚ ꪠꪽ ꪄꪮꪒ ꪄꪫꪱꪣ ꪼꪕ - ꪮꪰꪉꪁꪷꪒ; *'Păp Pe Khót 'Quám 'Tay - Ăng-kít; Tai Dam - English, English - Tai Dam Vocabulary Book.*

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

Fippinger Jay and Dorothy Fippinger. "Black Tai Phonemes with Reference to White Tai." *Anthropological Linguistics.* March 1970.