A Phonology of the Bu Dai (Thu Lao) Language

KOSAKA, Ryuichi
Tokyo University of Foreign Studies, Graduate School

0. Introduction
1. Phonological Description
   1.1. Initials
   1.2. Vocalic Nuclei
   1.3. Finals
   1.4. Tones
2. Genetic Affiliations
3. Phonological Features
   3.1. Conservation of Proto-voiced Initials
   3.2. Initial Palatalization
   3.3. Disappearance of the Second Element of Initial Clusters
   3.4. Merger of Proto-short Vocalic Nuclei
   3.5. Shortening of Proto-long Vowels Accompanied by a Final
   3.6. The /-y/:/-i/: Distinction
   3.7. Final Mergers
   3.8. Vowel Shifts

0. Introduction

Devoicing of initial consonants is a phonetic phenomenon that is generally observed with languages distributed in an extensive area of East and Southeast Asia such as Chinese, Tai, Mon-khmer and Tibeto-Birman languages. A similar kind of consonantal devoicing (*b > p, *d > t, *g > k), in fact, is also experienced by all Germanic languages. However, this phonetic change is fundamentally different from the former in that it is a constituent of the so-called systematic 'First Consonant Shift' of Germanic languages whereas the former devoicing in Asian tonal languages causes a consonantal merger and most frequently results in a subsequent tonal split.

Bu Dai is probably the only language among the known Tai dialects that still conserves voice in all of the voiced initials of the proto-language. This is not, however, unique, in consideration of the Wú (吴) dialects of the Chinese language.

As the Bu Dai language, previously referred to as Thu Lao, has seldom been the object of detailed linguistic analysis, what we know about it is considerably limited. As a descriptive study of Thu Lao phonology, we can refer to one carried out by Zhang Junru (1987). According to Zhang, Thu Lao is mainly spoken in some
districts (Wenshan, Malipo, Yanshan, Maguan etc.) of Wenshan autonomous county, China. The Thu Lao language (also referred to by him as Wennma patois (文麻土語) of the Zhuang language) that he describes is more specifically a dialect spoken in Heimo (黑末) village, and has already lost all the original stop finals unlike the Bu Dai dialect of the present study, which conditioned its posterior vocalic developments in a rather diversified manner. This might leave the impression that the two dialects differ considerably, but in reality they do have some fundamental phonological features in common. As for tones, Zhang’s Thu Lao has six tones instead of the five attested in the present dialect, which is probably related to the mentioned loss of its plosive finals. The Bu Dai language I worked on is segmentally more similar to the Thu Lao appearing in Ngô Đức Thịnh and Chu Thái Sơn (1975).

In the beginning of 1996, I was able to undertake field research on this language in Vietnam and formulate tentative conclusions as to its phonological system as well as its genetic affiliations.

The autonym of the Bu Dai people is /bu⁴ daŋ⁴/ (/bu⁴/ meaning ‘father’ or sometimes ‘man’ as in combinations of /bu⁴ kʰuŋ⁴/ ‘son-in-law’ and /bu⁴ phi⁴/ ‘ghost, spirit’), coinciding precisely with the autonym (*daŋ⁹) used by a number of Tai peoples (ex. Tày distributed roughly to the north of Sông Hồng river from CT group, a good many dialects from SWT group, but excluding the NT group) scattered from Yunnan and Guangxi provinces of China in the north, down to the Kingdom of Thailand in the south, and to the north-western part of India passing through the Shan state of Burma in the West. The direction of the Bu Dai’s migration must have been from north to south judging from some vocabulary items of apparently Chinese origin {cf. /pʰuŋ⁴/ ‘to sew’ (<縫?), /lu⁴/ ‘road’ (<路?), /χǒn⁴/ ‘bowl’ (<碗?), /tʰəŋ⁴/ ‘to carry in hand’ (<提?) etc.}.

The Bu Dai data were provided by Mrs. Vàng Diệu Sinh, farmer, 23 years of age, a resident of Na Cáng hamlet, Xin Ma Cai village, Bạc Hà district, Lào Cai province, Vietnam.

On account of the informant being incapable of understanding sufficient Vietnamese, the author’s original Vietnamese questionnaire was translated into Nung—a sort of regional lingua franca—by a native speaker, to which the informant responded.

1) According to Wei Qingwen et al. (1980), Yongnan (邕南), Zuojiang (左江), Dejing (德靖) and Yanguang (頑庯) patois as well as the Wennma patois are referred to as southern dialects of the Zhuang language. These southern dialects are situated from southern Guangxi to southern Yunnan, immediately south of the area of the northern dialects of the Zhuang language.

2) For instance, the phonological features demonstrated in 3.1., 3.2., 3.3. and 3.6. of this article also apply to Zhang’s Thu Lao, and the vowel changes in 3.8. are, for the most part, likewise applicable.
1. Phonological Description

1.1. Initials

The Bu Dai language is fundamentally, as the other Tai dialects, a monosyllabic tonal language, the syllabic structure of which is C₁V(C₂)/T³ (where the appearance of C₂ is not obligatory). C₁ is the initial consonant, V is the vocalic nucleus and C₂ is the final consonant. T is a supra-segmental feature — the tone. The Bu Dai initial consonant system is as follows (Table 1).

<table>
<thead>
<tr>
<th>Bilabial</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>t</td>
<td>c</td>
<td>k</td>
<td>?</td>
</tr>
<tr>
<td>ph</td>
<td>th</td>
<td>ch</td>
<td>kh</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>d</td>
<td>J</td>
<td>g</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ο)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v</td>
<td>δ</td>
<td></td>
<td>γ</td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>n</td>
<td>ŋ</td>
<td>η</td>
<td></td>
</tr>
<tr>
<td></td>
<td>j</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(tsh~s)

The phonemes in parentheses are presumed to be of external origin {cu² '(< 金秀 ?) 'rust', tshy⁴~sy⁴ '<吃 ?) 'to eat', these two, for example, being capable of coming from a dialect of Western Mandarin spoken in Yunnan, Kweichow and some other neighboring provinces of China).  

3) Strictly speaking, there is an occasional so-called 'minor syllable' attached in front of this fundamental word structure. In terms of stress, it is always weaker than that of the following major syllable and its segmental realization varies slightly according to the vowel in the major syllable. Tonal, there seems to be a difference at least between 'high ('')' and 'middle (')'. Though, it is not yet clear whether or not there are minimal pairs involving this tonal difference in the minor syllable.

For animal names, we can attest the existence of a prefix-like minor syllable, and there seems to be a confirmable overall distinction between one for animals which have been familiar to man from the distant past (including those commonly domesticated) on the one hand and another for other types of animals (including insects) on the other.

\[ /hα^-mo/^ /'dog', /hα^-ya/^ /'water buffalo', /hα^-mu/^ /'cow', /hα^-prt/^ /'duck' \]

VS

\[ /mi^-thvo/^ /'headouse', /mi^-po/^ /'fish', /mi^-nv/^ /'rat', /mi^-han/^ /'goose' \]

These minor syllables for animals could probably be traced back to common words originally equipped with a complete major syllable structure, that is, /hκ/^ /'child' and /mι/^ /'mother, female', respectively.

In this article, we focused on major syllables and abstained from making a concrete analysis of minor syllables.
The phonetic realizations of initials /t/-, th-, d-, n-, l-/ are lamino-alveolar [t-], [t'-], [d-] (plosive), [n-] (nasal) and [l-] (lateral liquid), and those of fricative /θ-, δ-/ are authentically dental [θ-] and [ð-] instead of alveolar (Note that /θ, δ/ diachronically derive from *s and *z/r respectively).

The point of articulation of the 'palatal' series is in fact rather alveolo-palatal.

The initial /j-/ tends to be pronounced as quasi-fricative ([ʑ-]) when it is succeeded by a front vowel and /-a(-)/ (cf. /ja1/ [ʐɑː] 'to give' vs /ja1/ [ʐɑː] 'elder brother'), whereas the initial is more likely to be approximant in the other environments.

The notation of all the word forms appearing hereafter is phonological.

EX.

/p-, ph-, b-, m-/  
pik⁵ 'swing' VS phik⁵ 'taro',  po² 'fish' VS bo³ 'shoulder',  po² 'fish' VS mo² 'dog',  phvn¹ 'rain' VS bvn² 'to fly',  phvn¹ 'rain' VS mvn¹ 'round',  bvn² 'to fly' VS mvn¹ 'round',  pvt⁴ 'duck' VS tvt⁴ 'to far!',  pet³ 'eight' VS cet⁴ 'seven',  po² 'fish' VS ko³ 'young rice plants',  pvk³ 'to plant' VS ṭvk⁴ 'chest',  pha³ 'to walk' VS thol¹ 'to die',  phan¹ 'to dream' VS chan¹ 'arm',  phvn¹ 'rain' VS khvn¹ 'bitter',  bvn² 'to fly' VS dvn² 'earth',  bvk⁴ 'tomorrow' VS jvk⁴ 'rope',  bi¹ 'fat' VS gi⁴ 'where',  mo¹ 'to come' VS no¹ 'paddy field',  mo¹ 'to come' VS no¹ 'grass',  mo¹ 'to come' VS no¹ 'sesame'

/t-, th-, d-, n-/  
tvn² 'full' VS thvn¹ 'stone',  tvn² 'full' VS dvn² 'earth',  tvk¹ 'to fall' VS nyk⁴ 'bird',  thol¹ 'eye' VS do⁴ 'river',  thol¹ 'eye' VS no¹ 'paddy field',  do⁴ 'river' VS no² 'thick',  tvn² 'full' VS cvn² 'gold',  tv² 'animal classifier' VS kv² 'salt',  tvk¹ 'to fall' VS ṭvk⁴ 'chest',  thol¹ 'to die' VS cha⁵ 'to comb',  thol¹ 'eye' VS kho¹ 'leg',  dvn² 'earth' VS jvn¹ 'salty',  di² 'navel' VS gi³ 'where',  no² 'thick' VS ṭo⁵ 'grass',  nan⁴ 'water' VS nån¹ 'silver'

/c-, ch-, j-, ŋ-/  
cet⁴ 'seven' VS chet³ 'centipede',  cvw² 'heart' VS jvw⁴ 'name',  cvn² 'gold' VS īvt³ 'beard',  chvk⁴ 'six' VS jvk⁴ 'rope',  chvk⁴ 'six' VS īvt³ 'beard',  jvk⁴ 'rope',  VS īvt³ 'beard',  cvw² 'heart' VS kvw³ 'near',  cvn² 'gold' VS ṭvn⁵ 'thunder',  chan¹ 'arm' VS khan¹ 'to ask',  jvw⁴ 'name' VS gvw¹ 'who',  ṭo⁵ 'grass' VS no¹ 'sesame'
KOSAKA, Ryuichi: A Phonology of the Bu Dai (Thu Lao) Language

/k/-
koⁿ 'young rice plants' VS kho¹ 'leg', kvuⁿ 'near' VS gvoⁿ 'who',
kv² 'salt' VS n̄ȳi¹ 'snake', khun² 'to ascend' VS gun¹ 'person',
khan¹ 'to ask' VS n̄an¹ 'silver', gan¹ 'lazy' VS Ôan¹ 'silver',
ke⁵ 'to wring' VS Òe³ 'to hatch'

/?/-,

/ʔ/-, h/-
ʔvn³ 'thunder' VS hvn⁴ 'that', han³ 'goose' VS Ôan² 'three',
haⁿ 'to weep' VS khâi² 'egg', hoⁿ 'five' VS voⁿ 'sky',
hoⁿ 'five' VS yo⁴ 'to speak', han³ 'goose' VS Ôan¹ 'silver'

/θ/-, ð/-
θan¹ 'cloth' VS ðv⁴ 'to buy', ʔvn³ 'sour' VS thvn¹ 'stone',
ðan¹ 'branch' VS van¹ 'tooth', duⁿ 'to know' VS yu¹ 'neck',
ðiⁿ 'long' VS ðiⁿ 'navel', ðin¹ 'sword' VS jin¹ 'sated'

/v/-, v/-
vø⁴ 'sky' VS yo⁴ 'to speak', vø⁴ 'sky' VS bo³ 'shoulder',
vø⁴ 'sky' VS mo¹ 'to come', yon¹ 'day' VS goⁿ 'outside',
vø⁴ 'to speak' VS n̄o¹ 'sesame'

/l/-, j/-
ln¹ 'boat' VS jv³ 'to stay', lan⁴ 'hawk' VS nan⁴ 'water',
ja⁴ 'to give' VS ja¹ 'elder brother', jo² 'medicine' VS n̄oⁿ 'grass'

1.2. Vocalic Nuclei

Bu Dai has seven vowels, as follows:

i /[-iː] / [-iː] (after a palatal initial) ~ [-iː] (in the other environments)
e /[-eː] / [-eː] (after a palatal initial) ~ [-eː] (in the other environments)

a /[-aː] / [-aː] ~ [-aː] (in the other environments)

u /[-uː] / [-oː]

EX. gi² 'where', ŋin¹ 'to hear', ke⁵ 'to wring', cet⁴ 'seven', khâi¹ 'to sell',
kai² 'chicken', kv² 'salt', kvw² 'I', ñu¹ 'neck', kuji² 'far',
yo⁴ 'to speak', yo¹ 'water buffalo', yon¹ 'day'

1.3. Finals

There are seven final consonants in Bu Dai, as follows:
-t [-t'], -k [-k'], -n [-n], -ŋ [-ŋ], -w [-ʊ], -j [-ɛ] ~ [-i], -w [-u̯]

Final /-u/ can occur only in sequence with the vowel /-x-/.

It is to be pointed out that in Bu Dai, as with many other Tai dialects, original vowel length conditions tonal realizations of the word when the syllable ends in a stop final.

The following table (Table 2) demonstrates the possible Bu Dai combinations of V (vocalic nucleus) and C₂ (final consonant including the case of zero final).

### Table 2. Possible V+C₂ combinations in Bu Dai

<table>
<thead>
<tr>
<th>C₂</th>
<th>-t</th>
<th>-k</th>
<th>-n</th>
<th>-ŋ</th>
<th>-w</th>
<th>-j</th>
<th>-u</th>
<th>-o</th>
</tr>
</thead>
<tbody>
<tr>
<td>-i(-)</td>
<td>jit³</td>
<td>pik³</td>
<td>ŋin¹</td>
<td>liŋ²</td>
<td>ciw²</td>
<td>—</td>
<td>—</td>
<td>khi³</td>
</tr>
<tr>
<td>‘to stretch’</td>
<td>‘wing’</td>
<td>‘to hear’</td>
<td>‘yellow’</td>
<td>‘Vietnamese’</td>
<td>—</td>
<td>—</td>
<td>‘to ride’</td>
<td></td>
</tr>
<tr>
<td>-e(-)</td>
<td>cet⁴</td>
<td>bek³</td>
<td>len⁴</td>
<td>deŋ²</td>
<td>ȵew⁴</td>
<td>—</td>
<td>—</td>
<td>ke⁵</td>
</tr>
<tr>
<td>‘seven’</td>
<td>‘to carry on ↲’</td>
<td>‘to run’</td>
<td>‘red’</td>
<td>‘to play’</td>
<td>—</td>
<td>—</td>
<td>‘wring’</td>
<td></td>
</tr>
<tr>
<td>-a(-)</td>
<td>?at³</td>
<td>dok⁴</td>
<td>nan⁴</td>
<td>thəŋ¹</td>
<td>khaw¹</td>
<td>maj⁴</td>
<td>—</td>
<td>kha¹</td>
</tr>
<tr>
<td>‘to bathe’</td>
<td>‘deep’</td>
<td>‘water’</td>
<td>‘tail’</td>
<td>‘white’</td>
<td>‘tree’</td>
<td>—</td>
<td>—</td>
<td>‘sell’</td>
</tr>
<tr>
<td>-r(-)</td>
<td>pry⁴</td>
<td>ɣyk⁴</td>
<td>phvŋ¹</td>
<td>phvŋ¹</td>
<td>krv⁵</td>
<td>kvj⁵</td>
<td>kvr⁵</td>
<td>kχ²</td>
</tr>
<tr>
<td>‘duck’</td>
<td>‘ripe’</td>
<td>‘rain’</td>
<td>‘to sew’</td>
<td>‘nine’</td>
<td>‘crab’</td>
<td>‘near’</td>
<td>‘salt’</td>
<td></td>
</tr>
<tr>
<td>-u(-)</td>
<td>hvt⁴</td>
<td>khut³</td>
<td>nuk³</td>
<td>ɗuŋ¹</td>
<td>—</td>
<td>kuj⁵</td>
<td>—</td>
<td>yu¹</td>
</tr>
<tr>
<td>‘mushroom’</td>
<td>‘deaf’</td>
<td>‘person’</td>
<td>‘cage’</td>
<td>—</td>
<td>‘banana’</td>
<td>—</td>
<td>‘neck’</td>
<td></td>
</tr>
<tr>
<td>-o</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>yo⁴</td>
</tr>
<tr>
<td>‘to speak’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ɔ(-)</td>
<td>yɔt³</td>
<td>yɔk³</td>
<td>mɔn²</td>
<td>mɔŋ⁵</td>
<td>—</td>
<td>nɔj⁴</td>
<td>—</td>
<td>yɔ⁵</td>
</tr>
<tr>
<td>‘to swim’</td>
<td>‘to go out’</td>
<td>‘pillow’</td>
<td>‘to teach’</td>
<td>—</td>
<td>‘a little’</td>
<td>—</td>
<td>‘water buffalo’</td>
<td></td>
</tr>
</tbody>
</table>

As a general rule, velar initials do not appear in words of truly Tai origin when the vowel is a front vowel (/-iø, -i, -øə, -e/-), being impeded by the forthcoming phonological rule 3.2. Among the words with front vowels cited above, those with velar initials are /ŋin¹, khi³, ke⁵/, of which /ke⁵/ must be of loan origin based on comparative evidence, whereas /khi³/ must derive from a proto-form with a labialized initial (*khwii¹), and /ŋin¹/ from a proto-form with a palatal initial (*nín⁰). Thus, no example in fact contradicts the mentioned rule of 3.2.

Bu Dai /-eø/ (without a final) could be referred to as a vowel basically for words of external origin {cf. /eθ⁴/ ‘to wash (hands)’ (< βα), /eθ²/ ‘to have’ (cf. Proto-Tai *mii⁰) etc.}, for, diachronically speaking, Bu Dai *-eø and *-ɔø became /-i/ (cf. *mee⁰ > /mi⁴/ ‘mother’) and /-u/ (cf. *yɔo⁵ > /yu¹/ ‘neck’) respectively (cf. *-ee > /-e/-, *-ɔo-

---

4) Zhang’s Thu Lao diphthongized original *-eø and *-ɔø- to /-iø/- and /-uø/-, the tendency of which is, as already described, observed phonetically in the Bu Dai of my survey as well.
> /-ə-/. This might also indicate that proto-Bu Dai could have already had an articulatorily narrower variant for *-ɛɛɛ and *-ɛɛɛ than was the case for *-ɛɛɛ and *-ɛɛɛ.

/-ɛ/- (with a final) seems capable of appearing only when the initial is palatalized as in the case of /ceč/ ‘seven’ (cf. 3.4. to appear). The actual pronunciation of /ceč/ is realized as something like [cɛčʔ] and the appearance of this medial-like [-l] must somehow be related to the occurrence of palatalization in initials (see 3.2., below).

An important feature of the pronunciation of /-ə-/ is that, when followed by a final, it is phonetically longer than the other vowels followed by a final, being realized as [-q:]. Furthermore, when accompanied by a sonorant final, /-ə-/ is more likely to be pronounced even longer, just as long as [-q:].

The occurrence of /-ə/ (without a final) is rather restricted, one of which origins is involved with a diachronic change of medial -w- as will be mentioned in 3.3. below {cf. ɤɛ ‘water buffalo’ vs ɤɔ ‘to speak’, kʰɔ ‘fast’ (<快?) vs kʰɔ ‘leg’}

The paucity of examples in vocalic column /-ə/ is explained in 3.4. below.

1.4. Tones

The Bu Dai of my survey has five tones altogether, phonetic contours of which represent a well balanced geometrical symmetry.

The tonal values of the five tones can be described as follows.

1. mid high falling
   EX. kʰvɛ̃ ‘horn’, kʰɛn ‘to ask’, mɛ̃ ‘hand’
2. low level (occasionally accompanied by a slight rise at the end)
   EX. kʰvɛ̃ ‘I’, hən ‘sweet’, mɛ̃ ‘pig’
3. high level (or sometimes high rising)
   EX. kʰvɛ̃ ‘knee’, hən ‘goose’, mɛ̃ ‘new’
4. mid level
   EX. kʰɛt̜ ‘to bite’, gɛ̃ ‘lazy’, ɛt̜ ‘to buy’
5. mid low rising
   EX. kʰvɛ̃ ‘rice’, kʰɛn ‘to be upside down’, ɤɛ̃ ‘clothes’

The correlation of the contemporary tones with the proto-initial and tonal categories is as below (Table 3).

<table>
<thead>
<tr>
<th>Table 3. Bu Dai Tonal Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>proto-tone</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>proto-initial category</td>
</tr>
<tr>
<td>H</td>
</tr>
<tr>
<td>H’ M</td>
</tr>
<tr>
<td>L</td>
</tr>
</tbody>
</table>

The most salient feature about Bu Dai tones is that the proto-voiceless initials other than those of the M category (proto-voiceless unaspirated plosives) are divided
into two groups — H and H’ — according to their tonal behavior, of which the former includes voiceless aspirated plosive initials and the latter, non-plosive voiceless initials. In fact, a similar type of initial sub-categorization is also observed in Kam (Thurgood, 1988).

2. Genetic Affiliations

In order to make a conclusion about the genetic affiliations of Bu Dai, let us refer to the following comparative list (Table 4). The reconstructed forms are cited from Li. The phoneme /y/ of Lungming and of Li’s reconstructions is retranscribed as /j/.

<table>
<thead>
<tr>
<th>Table 4. Comparison of Bu Dai with Other Tai Dialects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gloss.</td>
</tr>
<tr>
<td>——</td>
</tr>
<tr>
<td>‘ghost, spirit’</td>
</tr>
<tr>
<td>‘thorn’</td>
</tr>
<tr>
<td>‘day before yesterday’</td>
</tr>
<tr>
<td>‘to weep’</td>
</tr>
<tr>
<td>‘to come’</td>
</tr>
<tr>
<td>‘rice’</td>
</tr>
<tr>
<td>‘bitter’</td>
</tr>
<tr>
<td>‘ear’</td>
</tr>
</tbody>
</table>

For the last four words, only the Saeck language shows a tone corresponding to that appearing in the other three dialects. As to the initial of the word for ‘to come’, it is based on the fact that the tone appears as 1 and not as 2 that we know the Bu Dai initial goes back to (voiced) *m-. For the last three, the voicelessness of the Bu Dai plosive initials immediately demonstrates that the original initials were voiceless (cf. 3.1.). The proto-initial category in these four Bu Dai words coincides with Lungming and Siamese, and contrasts with Saeck. Thus, in terms of genetic affinity, we can tentatively assert that Bu Dai is a non-NT dialect based on the lexicophonological features shown above.

One of the arguments for distinguishing between SWT and CT, on the other hand, is the existence of some characteristic phonological features in the latter that are not generally attested in the former (Table 5).

---

5) Lungming is a CT dialect spoken in southern Guangxi in China. The data are from Gedney (1991).
6) Siamese phonological notation follows Haas (1964). Linguistically, Siamese belongs to the SWT branch of the Tai languages.
7) Saeck is a NT dialect spoken in some villages on both sides of Mekhong River around the city of Nakhon Phanom (Thailand) and the city of Tha Khek (Laos). The data are from my fieldnotes collected in Nakhon Phanom province.
Table 5. Comparison of Bu Dai with Other Tai Dialects

<table>
<thead>
<tr>
<th>Gloss.</th>
<th>Bu Dai</th>
<th>Lungming</th>
<th>Siamese</th>
<th>Saek</th>
</tr>
</thead>
<tbody>
<tr>
<td>'to laugh'</td>
<td>khv$^1$</td>
<td>luu$^1$</td>
<td>hủa</td>
<td>ruaw$^2$</td>
</tr>
<tr>
<td>'to look for'</td>
<td>kho$^1$</td>
<td>laa$^4$</td>
<td>hảa</td>
<td>raa$^2$</td>
</tr>
<tr>
<td>'stone'</td>
<td>thvn$^1$</td>
<td>thin$^1$</td>
<td>hín</td>
<td>riń$^2$</td>
</tr>
<tr>
<td>'tail'</td>
<td>thaj$^1$</td>
<td>thaaŋ$^1$</td>
<td>hảaŋ</td>
<td>ruŋ$^2$</td>
</tr>
<tr>
<td>'to vomit'</td>
<td>ḍək$^4$</td>
<td>laak$^5$ (laak)</td>
<td>râak</td>
<td>raak$^3$</td>
</tr>
<tr>
<td>'long'</td>
<td>ści$^1$</td>
<td>lej$^4$ (ści)</td>
<td>rii</td>
<td>rąj$^6$</td>
</tr>
<tr>
<td>'root'</td>
<td>ḍak$^4$</td>
<td>laak$^2$ (laak)</td>
<td>râak</td>
<td>raak$^5$</td>
</tr>
<tr>
<td>'boat'</td>
<td>lv$^1$</td>
<td>luur$^4$ (ści)</td>
<td>rua</td>
<td>ruą$^1$</td>
</tr>
<tr>
<td>'far'</td>
<td>kuj$^2$</td>
<td>kwaj$^1$</td>
<td>klaj</td>
<td>trąj$^1$</td>
</tr>
<tr>
<td>'right (side)'</td>
<td>ɕo$^2$</td>
<td>saa$^1$</td>
<td>khwaası</td>
<td>khuą$^4$ (L)</td>
</tr>
<tr>
<td>'cow'</td>
<td>mu$^1$</td>
<td>moo$^4$</td>
<td>ȵua</td>
<td>bọɔ$^4$ &lt; Viet. bò</td>
</tr>
<tr>
<td>'navel'</td>
<td>ḍi$^1$</td>
<td>nej$^3$</td>
<td>-duu</td>
<td>duas$^1$</td>
</tr>
<tr>
<td>'deep'</td>
<td>ḍak$^4$</td>
<td>nak$^3$</td>
<td>lǔk</td>
<td>lak$^6$</td>
</tr>
</tbody>
</table>

The forms in parentheses are those of Lungchow (Li, 1977), which belongs to the Central branch of the Tai dialects, as does Lungming.

Proto-Tai initials reconstructed for 'to laugh', 'to look for' and 'stone', 'tail' are *xr- and *thr- respectively, the distinction of which was generally annihilated in SWT dialects but generally maintained in CT dialects.

Referring to 'to vomit', 'long' (proto-Tai *r-) and 'root', 'boat' (proto-Tai *dr-), although the original distinction is still maintained tonally in Saek, it is generally lost in SWT and NT dialects. In CT dialects, this initial distinction is maintained segmentally in most cases. Lungming happens to be a rather exceptional case probably due to posterior merger of the former (*r) with the latter (l < *dr). Here Lungchow examples are additionally cited to be referred to instead.

In the case of 'far', many other CT forms contain a labialized element /-w/- or /-u/-, in contrast to SWT and NT dialects (Li, 1977).

Li presents the possibility that the proto-Tai initial for 'right (side)' is *khrw-, taking CT reflexes into consideration (Note that PCT *s- regularly corresponds to /θ-/ in Bu Dai).

For 'cow', dialects to be assigned to the CT branch generally show a form that could be traced back to proto-CT *mos$^6$ (Gedney, 1981).

For 'navel', both Bu Dai and Lungchow indicate original form *ʔdiʔ in common, disagreeing with SWT and NT forms in both vowel and tone.

The proto-CT initial for 'deep' reconstructed on the basis of Bu Dai and Lungming initials is *ʔd-, coinciding with Sarawit's (1973) proto-CT reconstruction (*ʔd-) and contradicting with her proto-SWT (*l-).

Another characteristic of CT dialects concerns the initials of the following words that appear with aspiration, which also applies to Bu Dai as follows (Table 6). Proto-
CT initial aspiration for these words is supported by Bu Dai words for ‘eye’ and ‘to die’ showing tone 1 (Li reconstructs for these words *thr- (‘eye’ / ‘to die’) and *phr- (‘to expose to the sun’)).

<table>
<thead>
<tr>
<th>Gloss.</th>
<th>Bu Dai</th>
<th>Lungming</th>
<th>Siamese</th>
<th>Saeck</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘eye’</td>
<td>tho¹</td>
<td>thaa¹</td>
<td>taa</td>
<td>praa¹</td>
</tr>
<tr>
<td>‘to die’</td>
<td>thə¹</td>
<td>thaaj¹</td>
<td>taaj</td>
<td>praa¹</td>
</tr>
<tr>
<td>‘to expose to the sun’</td>
<td>thək³</td>
<td>phjaak²</td>
<td>taák</td>
<td>praa⁶</td>
</tr>
</tbody>
</table>

Considering the features enumerated above, we could probably conclude that Bu Dai belongs to the CT branch⁸, despite the fact that the autonym *daj⁰ is more typically reminiscent of the SWT branch, and the fact that there is a general tendency to drop a clusters’ second element (cf. 3.3. below) is also a feature more frequently observed in SWT than in CT dialects. This proximity of CT to SWT would not at all be problematic if we take into consideration Haudricourt’s (1967) modified version of Kadai (=Tai-Kadai) classification.

3. Phonological Features

Bu Dai has typically undergone the following series of phonological changes⁹.

---

⁸) Bu Dai, notwithstanding, shows a non-Tai form for several lexical items as follows. The Pupeo form cited below is according to Hoàng Văn Ma et al. (1992). The proto-Kam/Sui (PKS) is based on G. Thurgood (1988). The Laha is taken from Solntseva et al. (1986).

EX. /tshə¹/ ‘to eat’, /te¹/ ‘to have’, /məj¹/ ‘year’ (cf. Pupeo. /mjaŋ¹/ ‘id.’, PKS /mpe¹/ ‘id.’, /məj¹/ ‘to love’ (cf. Laha. /maŋ¹/ ‘id.’)

⁹) Apart from initials, Tai and Kam-Sui languages tend to conserve the original Vowel + Final segments relatively well. However, Tai Ya (雅) and Lüchun (羅眷), both belonging to the SWT branch, could be cited as somewhat contradicting this tendency. In Tai Ya, *-uaj, *-uaj and *-aaj turned to /-ja:/ (the first and the second, probably via *-ooj and *-aaj respectively), while *-aaj turned to /-jau/. This phonological phenomenon is a type of reciprocal assimilation in terms of point of articulation and labiality (cf. The same kind of assimilation in terms of vowel hight is attested with French *-oi (>[*ûe]) > [wa]). In Lüchun, we have no finals except /-ŋ/ where original *-ŋ and part of *-n, *-p, *-t, *-k have merged historically. The Lüchun Vowel + Final development is diachronically quite complicated, and includes the problem of diphthongation. However, the vocalic front-back distinction seems to be maintained as a principle, in contrast to the sort of centralization of vowels that has occurred in Bu Dai.

The Tai Ya data are taken from Xing Gongwan (1989). The Lüchun data are cited from Luo Meizhen (1984). The asterisked forms here are proto-SWT.

EX. Tai Ya: *kluaŋ¹ > kjua¹ ‘banana’, *suaj¹ > suaj¹ ‘to wash’, *dwaŋ⁰ > jua¹ ‘cockpaw’, *khaŋ⁰ > k¹jua¹ ‘son-in-law’, *ræŋ⁰ > hja¹tu¹ ‘mark, true’, *dæŋ⁰ > j¹at¹ ‘mountain’

Lüchun: *hmat > ma³³ ‘house’, *pet > piq⁰ ‘duck’, *nok > nau⁴⁵ ‘bird’, *hnak > nav⁴³ ‘heary’, *hlek > le³³ ‘iron’
Unless otherwise noted, the proto-forms appearing below are proto-Central Tai as reconstructed by Li.

3.1. Conservation of Proto-voiced Initials

It is certainly striking that Bu Dai still maintains its original voiced series of initials. However, this fact in itself is nothing but a confirmation of the already presented hypothesis of existence of proto-voiced-voiceless distinction in initials, and a parallel conservatism, as was mentioned above, can be seen in the Wú (呉) dialects of the Chinese language.

EX. *phi$^i_0 > phi$^i 'ghost, spirit' VS *bi$^i_0 > bi$^i 'fat'
    *khun$^2_0 > khun$^i 'to ascend' VS *gun$^0_0 > gun$^1 'person'
    *kla$^2_0 > kwu$^5 'near' VS *grau$^0_0 > gyu$^1 'who'\textsuperscript{10}
    *swa$^2_0 > \theta_v^3 'clothes' VS *zuwu$^2_0 > \delta_v^4 'to buy'

In fact, initial merger did occur in Bu Dai through the deglottalization of proto-preglottalized initials and the vocalization (or despiration) of proto-voiceless sonorants, which is, as easily predictable, compensated for by a substitutional tonal distinction.

EX. *$\text{ʔ}daj$^0_0 > dvj$^2 'good' VS *daj$^0_0 > daj$^i 'Bu Dai'
    *hmaa$^0_0 > mo$^2 'dog' VS *maa$^0_0 > mo$^1 'to come'
    *hla$^j_0 > la$^2 'many' VS *ml/raaj$^0_0 > la$^1 'saliva'

The phonological rule of 3.6. (to appear) is to be conferred as regards the lack of vowel correspondence in the first example.

3.2. Initial Palatalization

When succeeded by proto-front vowels, palatalization generally occurred in velar initials.

EX. *kheen$^0_0 > *khjeen$^0_0 > *chaan$^0_0 > chan$^1 'arm'
    *khi$^h_0 > *khj\eta$^0_0 > ch\eta$^1 'ginger'
    *khi$^i_2 > *khji$^i_2 > chi$^5 'excrement' (cf. 'to ride')
    *gem$^0_0 > *gj\eta$^m_0 > j\eta$^1 'salty'

Interestingly enough, the following two examples have initials /$\text{j}$-/ and /$\text{y}$-/ respectively instead of the expected /$\text{l}$-/ , which might be relevant to the feature described

\textsuperscript{10} The author hypothetically presents the cited diachronic change by reference to the Siamese word for 'who' as a parallel example.
here in terms of medial generation.

EX.  *ʔim¹ > jin³ ‘sated’, *ʔ³ok > ʔ³ok³ ‘to go out’
But, *ʔ³ap > ʔat³ ‘to bathe’ *ʔ³un¹ > ʔun³ ‘warm’

The diachronic evolution of the two words above could have been something like
*ʔim¹ > *ʔ³im¹ > jin³ and *ʔ³ok > *ʔ³wok > ʔ³ok³.

3.3. Disappearance of the Second Element of Initial Clusters
The second element of an initial cluster is simply lost, as can be seen in the following examples:

EX.  *klwa⁰ > kv² ‘salt’ (cf. Lungming. kjuwu¹)
*phlaaj² > pha³ ‘to walk’ (cf. Lungming. phjaaj³)
*graan² > gon⁴ ‘lazy’ (cf. Lungming. laan⁶)
*bruk > bvk⁴ ‘tomorrow’ (cf. Lungming. cok⁵)
*khraj¹ > khaj³ ‘egg’ (cf. Lungming. laj²)
*ml/reep > met⁴ ‘lightning’ (cf. Lungming. meep⁵)
*xruu⁰ (>*khruu⁰) > khv¹ ‘ear’ (cf. Lungming. low⁷)
*hwi¹⁰ > hi² ‘comb’ (cf. Lungming. vej¹)

Exceptions:
*klaj² > chj⁵ ‘fever’ (cf. Lungming. laj³)
*xrok (>*khrok) > chrk⁴ ‘six’ (cf. Lungming. lok³)
*ml/raaj⁰ > la¹ ‘saliva’ (cf. Lungming. naaj⁴)
*ʔbl/ʔoak > dok³ ‘flower’ (cf. Lungming. mjook⁵)

In contrast to the Lungming examples, the simple disappearance of a cluster’s second element in Bu Dai is a feature which is not always shared with other CT dialects.

Proto-medial *-w- has either simply disappeared or transformed the succeeding vowel during the process of disappearing, as in the following examples:

EX.  *khwaam² > khɔ⁵ ‘to turn upside down’
*khwi¹ > kʰi³ ‘to ride’
*ʔwaaj⁰ > *ʔwɔ⁰ > ʔɔ⁵ ‘water buffalo’
*hwaan⁰ > *hwɔn⁰ > hɔn² ‘sweet’

An interesting comparison could be made between /kʰi³/ ‘to ride’ and /chi⁵/ ‘excrement’ (cf. Latin. qui- [*kwí-] > French. qui- [ki-] VS Latin. ci- [*ki-] > French. ci- [si-]).
3.4. Merger of Proto-short Vocalic Nuclei

Proto-short vocalic nuclei generally undergo centralization (>/γ), except for */-a/- in whatever environment, and */-i/-e preceded by a palatalized initial.

EX. *sip > θvγ^4 'ten', *tem^0 > tvn^2 'full', *phun^0 > phvn^1 'rain', *suk > θvγ^4 'ripe', *λγ^0 > λvγ^1 'to descend'
(*zuun^0 >) > δvγ^1 'day before yesterday' (cf. 3.5.)
(*ruan^0 > *ruuun^0 >) > δvγ^1 'house'
cf. *ram^0 > δan^1 'bran', *ñin^0 > ñin^1 'to hear', *cet > cet^4 'seven'

Exceptions:
*pliŋ^0 > piŋ^2 'leech', *khun^2 > khun^5 'to ascend'
*gun^0 > gun^1 'person', *θun^0 > than^1 'to arrive' (cf. PNT *dan^0)
*ŋun^1 > ŋun^3 'warm'

3.5. Shortening of Proto-long Vowels Accompanied by a Final

Following are some examples demonstrating the shortening of proto-long vowels accompanied by a final:

EX. *piik > pik^3 'wing', *kluuj^2 > kuj^5 'banana'
*pεt > pet^3 'eight', *rουŋ^2 > δoŋ^1 'to chirp'
But, *scoŋ^0 > ŋov^2 'two'

Based on tonal evidence, we can conclude that the words for 'wing' and 'eight' must have had a long vowel at the proto-stage.

The vocalic change of */-uun- > -i- in the two following words seems exceptional with reference to the other examples.

EX. (*plhuak > *plhuuuk > phik^3 'taro'
(*hluaŋ^0 > *hluuŋ^0 > liŋ^2 'yellow'
cf. (*juaŋ > *juuuk > jvγ^4 'rope'
(*ruan^0 > *ruuun^0 > δvγ^1 'house'

As for the reflex of */-aa-, despite being phonologically interpreted as /-a/- owing to its merger with */-a/- (excluding the cases of the final being */w/- and */j/-), its actual pronunciation is half long or long. It is not, therefore, treated here as having undergone vocalic shortening.

EX. *naak > nak^4 ([-q-]) 'otter' VS *hnak > nok^4 ([-q-]) 'heavy'
*hnaam^0 > nan^2 ([-q-] ~ [-q-]) 'thorn'
VS *nam^2 > nan^4 ([-q-] ~ [-q-]) 'water'
cf. *khaaw⁰ > khaw¹ ‘white’ VS *khaaw² > khaw⁵ ‘rice’
*khaaj⁰ > kha¹ ‘to sell’ VS *khraj¹ > khaj³ ‘egg’

3.6. The /-xj/ : /-aj/ Distinction

As we will explain later on, proto-CT *-aaj developed to /-a/ in Bu Dai. However, Bu Dai possesses the /-xj/ : /-aj/ distinction, for both of which there is comparative evidence indicating that they seem to originate from *-aj. We leave the question unexplained for the time being for lack of further linguistic evidence, though it might indicate some vocalic distinction at the proto-Tai level.

EX. /-xj/ — vj³ ‘fire’, θvj² ‘intestines’, prj² ‘to go’ nvj⁴ ‘this’, dvj² ‘good’
/-aj/ — haj⁵ ‘to weep’, kaj¹ ‘chicken’, khaj³ ‘egg’ maj⁴ ‘tree’,
    dj¹ ‘Bu Dai’

The form reconstructed by Li for PCT and PSWT in common concerning all the words above cited is *-aj except for ‘this’ and ‘good’, PSWT form for these two being *-ii. Li (1977, p.286) points out that Lungchow — a CT dialect — has /-ai/ for the two words in question, and Lungming — another CT dialect — has /naaj⁵/ ‘this’ and /naaj⁴/ ‘good’ retraceable to *-aj as well (though as for ‘this’, Shan (SWT) and Wuming (NT) also have -aj). Here we have another piece of phonological evidence affiliating Bu Dai to the CT branch. This vocalic distinction is represented as /-ei/ vs /-ai/ in Wenma patois of Zhang.

3.7. Final Mergers

In Bu Dai, the proto-finals *-p and *-m merged to *-t and *-n respectively, thus creating the following pairs of (quasi-) homophones.

EX. *tap > tat⁴ ‘liver’ VS *tat > tat⁴ ‘to cut’
    *lum⁰ > lvn¹ ‘to forget’ VS *lin² > lvn⁴ ‘tongue’

3.8. Vowel Shifts

As the number of possible V+C₂ segments has been reduced due to the phonological changes mentioned in 3.4. and 3.7., the following vocalic developments systematically occurred in order to ‘maintain’ the required level of phonological distinctiveness in the Bu Dai language.

α *-a- > -a- (phonetically half-long or long)
EX. (*tok > tvk⁴ ‘to fall’ →) *tak > tak⁴ ‘to dip up (water)’
    (*khom⁰ > khvn¹ ‘bitter’ →) *khan⁰ > khān¹ ‘to crow’
\[ \beta \ *-\text{aa} > -\text{o} \]
EX. \((^\text{da}^\text{j} > \text{daj} \quad \text{Bu Dai}' \to)\quad *\text{thraj}^0 > \text{tha}^1 \quad \text{‘to die’} \)
\((^\text{kharaj}^1 > \text{khoaj}^3 \quad \text{‘egg’} \to)\quad *\text{kaaj}^0 > \text{kha}^1 \quad \text{‘to sell’} \)

\[ \gamma \ *-\text{aa} > -\text{o} \]
EX. \((^\text{thraj}^0 > \text{tha}^1 \quad \text{‘to die’} \to)\quad *\text{thraj}^0 > \text{tho}^1 \quad \text{‘eye’} \)
\((^\text{khaaj}^0 > \text{kha}^1 \quad \text{‘to sell’} \to)\quad *\text{khaaj}^0 > \text{kho}^1 \quad \text{‘leg’} \)

\[ \delta \ *-\text{öö} > -\text{u} \]
EX. \((^\text{ʔbaa}^1 > \text{bo}^3 \quad \text{‘shoulder’} \to)\quad *\text{boo}^1 > \text{bu}^4 \quad \text{‘father’} \)
\((^\text{maa}^0 > \text{mo}^1 \quad \text{‘to come’} \to)\quad *\text{maa}^0 > \text{mu}^1 \quad \text{‘cow’} \)

\[ \varepsilon \ *-\text{uu} > -\text{y} \]
EX. \((^\text{moo}^0 > \text{mu}^1 \quad \text{‘cow’} \to)\quad *\text{muu}^0 > \text{my}^2 \quad \text{‘pig’} \)
\((^\text{tu}^1 > \text{tu}^3 \quad \text{‘a kind of wasp’} \to)\quad *\text{tuu}^0 > \text{ty}^2 \quad \text{‘animal classifier’} \)
\((^\text{yu}^0 > \text{yu}^1 \quad \text{‘neck’} \to)\quad *\text{xuru}^0 > \text{kv}^1 \quad \text{‘ear’} \)

A diachronic vowel change similar to the one having occurred in Bu Dai (characterized as labialization of *-aa (and delabialization of *-uu)) seem to have occurred, for instance, in French, Swedish, Biao\(^{11}\) and the Jiamao (加茂) dialect of the Hlai language as well. Kosaka (1996) is to be conferred as to the apparent discrepancy between the Jiamao initials and those of the Siamese.

Examples of French Nasalized Vowels:
\[ *\text{ã} > [\text{ã}], \quad *\text{ı} > [\text{õ}], \quad *\text{ũ} > [(\text{ɛę }) > \text{ę}], \quad (*\text{ı̈} > [\text{ɛ}], \quad *\text{ɛ} > [\text{ɛ}] \)\]

Examples of Swedish vowels:
\text{gä} [\text{goː}] \quad \text{‘to go’}, \quad \text{ko} [\text{kɔː}] \quad \text{‘cow’}, \quad \text{du} [\text{duː}] \quad \text{‘you’} \]

Examples of Biao vowels:
\text{θo}^1 (Siam. taa) \quad \text{‘eye’}, \quad \text{mu}^3 (Siam. māa) \quad \text{‘dog’}, \quad \text{jo}^6 (PKS *ʔrâ) \quad \text{‘paddy field’} \]

Examples of Jiamao vowels:
\text{pou}^4 (Siam. māa) \quad \text{‘dog’}, \quad \text{tou}^4 (Siam. plaa) \quad \text{‘fish’}, \quad \text{pai}^4 (Siam. mūu) \quad \text{‘pig’} \]

\(^{11}\) The data of the Biao (標) language, which is easily confused with Pupeo or Pubiao (普標) — one of the ‘Kadaï’ languages —, are taken from Zhang Junru (1989). It is a language spoken in several villages in Guangdong province in China, and is, linguistically, a member of the Kam-Tai languages. However, its detailed genetic location is somewhat unclear, except that a few vocabulary items characteristically resemble those of Lakkia.
It seems that in Biao vowels no systematic ‘compensatory’ vocalic shift has occurred in conformity with the diachronic change of *-a > -o,-u. It is true on the other hand that neither proto-short vowel merger (cf. 3.4.) nor final mergers (cf. 3.7.) is attested in Biao.

Finally, my sincere appreciation is offered to Prof. Hoàng Thị Châu, who assisted me in making arrangements to carry out my field research in northern Vietnam. Also my sincere thanks go to Prof. Đỗ Ca Sơn, who readily accepted the troublesome task of accompanying me as far as the steep mountainous areas near the the Sino-Vietnamese border in search for a Bu Dai informant. I cannot thank enough the informant who worked with me so patiently.

I would also like to express my sincere and heart-felt thanks to Prof. Tadahiko Shintani and Prof. Yasuyuki Sakamoto for their valuable advice and for letting me refer to their bibliographies, and to Prof. Yasuyuki Mitani, who encouraged me to work on this intriguing language in the first place.

References


