

Thai Etymologies

Christian BAUER
Mahidol University

The etymological notes that follow are intended to be contributions to the lexicography of Thai. In this installment—there will be further contributions in the future—I discuss loan words from Mon and Khmer. The individual notes were written over a number of years and are still fragmentary in nature; some sections presented here, notably #9 and #10, will be continued in some other issue of *MKS*. I emphasize that these notes are provisional, and that much more textual analysis is needed. I decided to publish these notes in their fragmentary form now mainly because of the continuing discussion about the possible dating of the loss of voicing distinctions in Thai adding loans from Khmer and Mon, and also as a contribution to the sociolinguistic history of the Chao Phraya basin and northeastern Thailand.

It will be noticed that five of the eleven loans discussed here are grammatical markers, and one of the remaining six belongs, by any definition, to the core vocabulary. Until now discussions of early loans in Thai, mostly Khmer, centered around lexical loans, and structural borrowing was rarely taken into account; where it was, the analysis was based only on a comparison of the modern languages, Khmer and Thai, at the exclusion of Mon, although it is the latter that may have given rise to syntactic reanalysis in Thai. Where Old Khmer and non-epigraphic Thai data were adduced, only the grammatical functors were identified but not their functional range and exact syntactic environments.*

A discussion of syntactic and semantic calques (borrowing of structures) and loans (borrowing of elements) will tell us more about bilingualism and patterns of language shift in pre-modern Thailand and adjacent areas.

The following loans will be discussed here:

- | | |
|----------|-------------|
| 1. ชัน | /khǎn/ |
| 2. กระยา | /kraya/ |
| 3. ทรหด | /thɔːrəhòt/ |
| 4. ทะเย | /thayɛː/ |
| 5. ละ | /láʔ/ |
| 6. ฉะ | /chàʔ/ |

*For earlier studies on Khmer-Thai linguistic contacts see Ferlus 1985; Huffman 1973, 1986; Karnchana Nacasakul 1962, 1971; Shorto 1971; Uraisi Varasarin 1984, a thesis submitted to the Université Paris III in 1975. Abbreviations used are: ETh.-Epigraphic Thai; ModTh.-modern Thai; OM-Old Mon; MM-Middle Mon; SM-modern Spoken Mon; IA-Indo-Aryan.

7. เฉพาะ /chəphǎw/
 8. ลี /sǐi/

1. /khǎn/ ชัน

This word is attested for the first time in 8.2.14-15 (AD 1369); it is spelt *khanna* **ขนน**, to be reconstructed as */khan/, corresponding to modern Thai **ขัน** /khǎn/ 'vessel, cup, for offerings'. It is borrowed from Old Mon *khal* corresponding to modern spoken Mon /khǎw/. The loss of final /-l/ in Mon can be dated to have taken place before Early Middle Mon (EMM) in the late 15th century. Shorto reconstructs for OM a form /khǎw/, although a proto-Mon form */khal/ must be postulated given (1) Nyah Kur evidence where /a/ has been retained corresponding to the OM written inherent vowel *a*, (2) IA loans in OM, and (3) early Khmer-Mon contact words such as the relative clause-marker OKhm. *man* */man/, OM *man*, *min*, *mun*. The date of the shift in Mon of written inherent *a* */a/ to /ɔ/ (SM /ɔ/) cannot be dated, but given Nyah Kur evidence it must have existed in varieties of early Old Mon.

Michel Ferlus argues, in conversation (May 1990), that ETh. *khan* &c. is more likely to be borrowed from OKhm. *khan*, where short /a/ is noted either as *a* or *a*.

To this one may counterargue that (1) no variation in vowel length is attested for this particular word in Khmer, (2) no variation in vowel length is attested in Epigraphic Thai for this particular word, and (3), most importantly, modern Thai **รับ** /ráp/ 'to receive' cannot have been borrowed from Khmer given the different vocalism of its modern reflex, but rather from OM *rap* 'to ready'.

2. /kraya/ กระจาย

Attested in 3.1.5 (AD 1357) as *karayā* **กรยา**, Griswold/Prasert connect ETh. *karayā* with modern Thai /kraʔya/ **กระจาย** which they take as borrowing from Khmer /krəya/.¹ The reason for this interpretation is that in ETh. 3.1.5 it is part of a compound *karayā taiwāy*, the latter being Old Khmer for 'offering', derived from OKhm. *twāy* 'to offer'. However, they point out that in this context the compound is not to be translated as 'food' but rather as 'provisions, offerings'. No reference to OKhm. forms of modern Khmer /krəya/ is given.

The passage in 3.1.5/6 reads as

แต่งกรยาดงวายของฝากหมาก(ป)สามาโหวิ

which Griswold/Prasert translate as "[... vassals,] having provided ritual offerings and gifts of fruit and fish, come to do him homage" (p.94-5, *op.cit.*). In footnote 16, they comment "In Siamese **กระจาย** means food, provisions, articles or utensils (McFarland) though in Khmer *krayā* generally means food only" (*ibid.*).

¹A.B. Griswold and Prasert ๓ Nagara "The epigraphy of Mahādhammarāja I of Sukhodaya." JSS, 1973.Ixi.1:71-178.

In Khmer epigraphy this word is equally problematic: In some cases such as K.257.S.37 (AD 979), Coedès translates it as "cérémonie", or as "accessoires de l'oblation" (K 659.15, 22), in others, the majority, as 'food': "C'est le sens [*sc.* "cérémonie"] qu'a *kriyā* en sanskrit. Mais il semble que dans plusieurs expressions *kriyā* ait déjà le sens qu'a pris *krayā* en camb. mod. à savoir: "nourriture, ingrédients" et qui irait mieux ici après "donner."²

By contrast, I propose here to interpret ETh. *karayā* as a borrowing from OM (although ultimately an IA loan); whether one can reconstruct an ETh. form */kərya/ rather than */krəya/ as implied by Griswold/Prasert, that is, postulating for early Thai a syllabified cluster, remains to be seen.

The OKhm. form is attested 29 times in the following orthographies: *kriyā* (25), *kriyā* (1), *krayā* (3); a compound *krf lya taiwāy* is not attested in Khmer epigraphy. Attested are the compounds *kriyābhojana* (K.88), *kriyādyah* (K.270.S), and *kriyāsthāpaka* (K 958.N). In OM it occurs in a syllabified form as *kiryā*, *kirryā*, *karyā*, *kiryāh*, reconstructed by Shorto as */kərya/ and glossed as 'articles of personal adornment, parure, regalia, appurtenances, gear', and connected with Pali *kriyā*.³ With the exception of I.E.5 (AD 1095 [Shorto]/AD 1102 [Luce]) where we find *kriyā*, it is only *krayā* from EMM onwards (15th c.) that non-syllabified clusters /krə-/ occur in this particular word. In native Mon words the difference between C₁əC₂- and C₁C₂- (where C₁- is a stop and -C₂- a liquid) is phonemic; IA loans are naturalized in this way in OM, as for instance Skt. *pradhāna* 'leader' OM *pidhān*, *pardhān* (in the Nagayon glosses once *pradhān*) */pərdhan/; Skt. 'length, extent' OM *piyām* *paryām*, */pəryam/. Syllabified clusters in Khmer loans are also attested in ETh., such as 'to teach' ETh. **បរយណ** *parayana* in 3.1.74 (AD 1357), OKhm. *paryyan*, *paryyān*, *paryānna* derived from OKhm. *ryān*, *ryyān*, *ryyan*, *ryyan* */ri:ən/ 'to learn'.⁴ Thai orthography since Sukhothay—unlike regional varieties of **ธรรม** script—cannot distinguish between syllabified CəC- and non-syllabified clusters CC- (say, /kə-/ and /kr-/). Since the predominant form in Khmer is *kri*° it is unlikely that the ETh. form is derived from Khmer. Given the semantic proximity to Mon, it is unlikely to have been borrowed from Khmer; in that case, one could assume the existence of syllabified clusters in early Thai.

3. /thəwəhət/ ព្រហ្ម

This word, glossed by various lexicographers (McFarland, Haas, So, Royal Institute 1982/2525) as 'to be tough, unyielding, resolute; to have endurance', is not attested in Sukhothay epigraphy. The dictionary of the Royal Institute (1982 ed.) accepts it only as part of a compound.

²G. Coedes, *Inscriptions du Cambodge*, Paris, EFEO, 1952 (Vol.iv, p.148).

³Some final *-h* spellings, like in epigraphic Khmer, especially Middle Khmer, indicate in some contexts lengthening of the preceding vowel.

⁴The syllabified cluster in Khmer may, in fact, mask an earlier nasal infix, or may be explained as a form contaminated by Mon.

This is a Mon word, borrowed into Thai not earlier than the 12th century and not later than the 15th century AD.

In OM *dirhat* /dərɬət/ is a noun, meaning 'strength' magnitude', derived from the verb /tɬət/ 'to be strong, powerful, severe' by affixation with [-r-]. When we look at syllabified complex initials CəC(C)- in OM we witness a voicing process, affecting the first member of the sequence, already attested in the 6th century; but this voicing process affected different classes of initials at different times.

Initials of clusters consisting of two different voiceless stops C in a sequence CəNC- (where N is a nasal), such as */təmpəh/ 'seven', were voiced by the 6th century, > /dəmpəh/ whereas initials of clusters consisting of identical voiceless stops C₁- in a sequence C₁əC₂C₁-, such as OM *tirtūy* ~ *tirtuy* /tərtuy/ 'perfection' were voiced only after the 13th century and before the late 15th century. As far as complex initials of the form C₁C₂h- are concerned, we witness a period of transition in the 12th century:⁵

- (i.a.) *tirhin kirsah* 'applause' ~ *thin ksah* 'to praise'
 (ii.a) *dumhic* 'right conduct' ~ *thic* 'good, fair'
 (i.b) *cirhāy* 'beauty, splendour' ~ *chāy* 'beautiful' (Skt./P.chāyā)
 (ii.b) *jirhān* 'loving-kindness' ~ *chān* 'to love'

Modern Thai has a number of words with syllabified clusters pronounced /tɬə: rə-/ apart from Skt. loans of the form *dura*⁶, 18 in the dictionary of the Royal Institute (1982 ed.); these are Khmer loans attested in early Thai /ka:p/ poetry, and here prosodic rules may indeed govern the pronunciation.⁶ But this is not the case with /tɬə: rəhət/ which occurs in colloquial contexts in modern Thai.⁷

4. /thayɛ:/ ทະเย

Like the preceding word, /tha:yɛ:/ ทະเย is not attested in inscriptions; it is first attested in the early 19th century epic *Khun Chang Khun Phaen*, and glossed for the first time in Bradley's 1873 dictionary, there spelt ทเย, as "to sing [Mon style]". So (1965) contains the gloss "a variety of musical compositions, especially ทเย and ทเยนพอก often played as overtures" (p.447), suggesting that it was well integrated into Thai musical culture. It is also attested in the dictionary of the Royal Institute (1982), glossed as "old Thai songs," also to accompany Mahori performances.⁸ McFarland and Haas do not list it.

⁵See Bauer 1982, which contains an internal reconstruction of Mon consonantism in the chapter on phonology (pp.1-241); pp.81-90 deal specifically with the problem of voicing in the contexts mentioned here. Ferlus 1983 [1984] discusses Mon consonantism briefly on pp.77-82, but not the development of complex initials.

⁶More textual studies are needed, both in Thai and Khmer, to show whether Khmer rhotacized initials may have had two different realizations, say, (i) /dər-/ and (ii) /drə-/.

⁷If it had been borrowed prior to the 12th c., it would have been borrowed as */tɬə: rəhət/, or **/tɬə: rəhət/; the latter form would have been retained in modern Thai.

⁸Mon music was performed at royally sponsored cremation ceremonies until the third reign; I doubt that this was introduced as an innovation by the Chakri monarchs. They may have relied on a continuing tradition. I should also point out that the relevant passage in *Khun Chang Khun*

This word is a loan from Middle Mon, and cannot have been borrowed before the 13th century; it must have been borrowed prior to the 19th century, and is not spelling pronunciation calqued on a Mon spelling. This suggests that after the 13th century Thai must have had a voicing distinction. The Middle Mon form for 'to sing' is /dəyeh/.

SM /həyèh/ 'to sing' corresponds to an OM form *jiñjeh* ~ *jiñjeh* ~ *janjeh* to be reconstructed as */jəñjeh/. By 1462 the shift from OM /jəñj-/ initials to MM /dəy-/ was already completed; the change also affected voiceless complex initials, as in OM *circūn* 'staff' /cərcun/ > MM *dacun* /dəcun/, SM /həcun/. This means that modern Thai /thaʔye:/ cannot have been borrowed before the first quarter of the 13th century when dissimilation of initials (OM /jəñj-/ > MM /dəy-/) and medio-cluster reduction had not taken place yet. By the time it was borrowed into Thai the voicing distinction there must have been retained. The word cannot have been borrowed in the 19th century because by 1847—at the time of Haswell's compilation of the first Mon dictionary—we must presume the velarization of what have become SM /hə-/ initials (from LMM */xə- ~ ʔə-/), in which case Thai would have 𑜀𑜢𑜤𑜰 or 𑜀𑜢𑜤𑜰.⁹ It cannot have been borrowed through a spelling because (1) we would expect the final *-h* /-h/ in Mon to have been retained in Thai, but marked by virāma, say, 𑜀𑜢𑜤𑜰, and (2) the vowel /ɛ:/ corresponds to a spoken form in Mon (lowering of [e] before final laryngeals).¹⁰

5. /láʔ/ 𑜀𑜢

The precise origin of Thai 𑜀𑜢 cannot be determined except that it is either Khmer or Mon; in both languages it has become obsolete. Its paucity in Pre-Angkor Khmer, and the fact that Old Khmer is likely to have borrowed a limited number of Mon particles during the 9th to 11th centuries AD—a second subordinating marker OKhm. *man* (in addition to its own *tel* and *syāñ* &c.), the clause-/sentence-final *ra*, among others (the list of contact-words is actually larger)—suggests that it may ultimately be an OM word, borrowed by both Khmer and Thai, if one discounts Khmer-Mon cognacy (there are cognates in Sre, Biat and Vietnamese). This seems to be supported by early spellings: Mon final *-h* /-h/ is spelt, in this instance, by *visarga* (reflected in the Thai spelling) whereas Old Khmer has *-h* spelling for *-h* in this case.

This loan is most problematic; phonologically, my case rests on analogy: Modern Central Thai /phráʔ/ 𑜀𑜢 is a late Old Khmer borrowing /βrah/ ~ /brah/ (Old Khmer /vrah/, spelt *vrah*). Central Thai subsequently shifted *voiced stops, in initial position, to aspirated voiceless ones, Northern Thai to unaspirated voiceless stops. What happened to the laryngeal final in the target-language (Thai)

Phaen (in the episode of receiving gifts on behalf of Phali Ngam, verse 649-56) lines from Mon songs are inserted into the Thai text in the original Mon, and that the Thai spellings of Mon words indicate the imitation of spoken Mon at the time (where *-lʔ/* rhymes had not shifted yet to *-ɛʔ/*).

⁹The literary Mon form, historically representing the intermediate stage between Middle Mon and spoken Mon, is LM *gayeh*, *gayih*, possibly reflecting this initial velarization. This hypothesis of velarization of initials in pre-modern Mon was first proposed by Shorto in teaching in the late 1970s.

¹⁰Problematic here is the fact that Mon /keh/ 'to engrave, write with a stylus [on palmleaf]' corresponds to Thai 𑜀𑜢; one would then expect MM /dəyeh/ 'to sing' to correspond to Thai 𑜀𑜢𑜤𑜰.

(Old Khmer /vrah/ , spelt *vrah*). Central Thai subsequently shifted *voiced stops, in initial position, to aspirated voiceless ones, Northern Thai to unaspirated voiceless stops. What happened to the laryngeal final in the target-language (Thai) we do not know. Proto-Tai, as reconstructed by Fang Kuei Li, does not have any laryngeal finals (*-p, *-t, *-k, *-m, *-n, *-ŋ, making allowance for a possible *-l to accommodate the Saek evidence); so what is the phonological origin of Thai ๕? Phoneticians have questioned the existence of a glottal stop in Thai; yet, final /-h/ in MK languages appears to correspond consistently to final /-ʔ/ in Thai or an open syllable with a lengthened vowel.

The second problem arises with early occurrences of Thai ๕. In the Sukhothay inscriptions it is attested only 38.2, in line 12 as an autonomous verb meaning 'to set free, abandon' (an Old Khmer loan), elsewhere (lines 11, 17, 18, 19), but only when combined with a co-ordinating clitic, the reflex of which is modern Thai ๕๓. There appears to be uncertainty as far as the tone is concerned since three different spellings occur in the same inscription; this may have something to do with the position of this clitic (stressed/unstressed) and other factors: ๕๓๓๓, ๕๓๓๓, ๕๓๓๓. It is even questionable that we deal here with two segmentable elements, ๕๓ and ๕.

However, I discuss the origin of modern Thai ๕ for the following reasons: (1) Syntactically it is clitics that relate to co-ordination in Khmer and Mon, (2) semantically they are quantifiers in all three languages, (3) distributionally all three are enclitics. Obviously, this is little to go on; but the speculation that ๕ may well turn out to be a Mon loan is fully justified.

By way of conclusion I shall illustrate its use in Old Khmer and Old Mon. In Old Khmer /lah/ may occur both as a verbal and a nominal enclitic:

[...] *cuh lah tlen lah* (K 940.10, AD 578-677 [undated])
'whether going downriver or whether going upriver'

[...] *cau lah kantai lah* (K 451.N.5-6, AD 680)
'be they grandchildren or women'

Old Mon:

gumr̥ lah paṇḍit lah bumn̥ah lah [...] (I.H.39-41)
'whether monks, scholar or Brahmins'.

6. /chàʔ/ ๕๓

๕๓ occurs together with the deictics *นั้น* and *นี้*, combined to *๕๓นั้น* and *๕๓นี้*, /chaʔ nán/ ~ /chaʔ níʔ/, translated, in context, as 'so, then, consequently'. It occurs in other expressions such as *เพราะ๕๓นั้น* /phrəʔ chàʔ nán/ 'that is why, therefore, consequently'. The latter is a blend-form, consisting of Khmer /pruəh/ 'because', and what I claim to be Mon /chaʔ/. Again this form is not attested in Old Mon but only in Middle Mon (AD 1480), classified by Shorto as a full noun which confirms its usage in Thai; the deictics require a preceding noun. Shorto translates it as 'whoever, whatever'; we can note one instance given in DMI:

cha deh ma ca twān [...] (XII.B.47)
 --- /pprN/SUB/to eat/ village/to be new
 'whoever rules the new village'

Halliday's MED (1922) glosses Mon *cha* as "fold, as only, whatever as follows", as shown in Figure 1.

The contexts given by Halliday, and to be further examined in future textual analyses of early and contemporary Mon texts, confirm Shorto's classification of Mon *cha* as a noun; when preceding verbal constructions it has to be linked by the subordinating clitic /ma? ~ mə-/. So far, I cannot explain the change of word-order with numerals. Combinations with demonstratives as they occur in Thai exist in Mon, such as /cha? nə?/ 'only this'. It is not attested in Epigraphic Thai.

<i>bā cha</i>	'two-fold'	(two.---)
<i>cha jaku mway</i>	'only himself'	(---/body/one)
<i>cha ma pru pren</i>	'whatever is arranged'	(---/SUB/to arrange/id.)
<i>cha gah hwa' señ</i>	'moreover'	(---/CL/NEG/[NEG]to be)
<i>cha gah</i>	'but'	(---/to say)

Figure 1: The distribution of LM *cha*; source Halliday 1922.

7. /chaphóʔ/ เฉพาะ

Thai /chaphóʔ/ means 'especially, peculiar', occurring also in expressions such as /dory chaphóʔ/ โดยเฉพาะ 'especially'. When not adverbial it occurs pre-nominally. It is derived from Khmer *chboh*, attested in Middle Khmer. Saveros Pou (*BEFEO* 60 [1973]: 163-70) glosses it in Middle Khmer as "vers, en présence de" (IMA 17.6); this particular passage is problematic since she interprets the published text version's *jhmoh* as a misreading for *chboh*. But a directional sense of 'towards, to' seems likely if other passages, where it occurs several times and providing unambiguous readings, are examined, such as *chboh dau* (IMA 38.134) and *chboh do* (IMA 3.A.82). In the attested cases *chboh* is not pre-nominal unless linked by a particle (IMA 17.6 *chboh nū*).

8. /sìʔ/ สึ

In Lao and Lawa, the 'future, irrealis, hypothetical' is marked by the preverbal particle /sìʔ/ (Lao) or /si/ (Lawa). A connection with Mon warrants further examination.

The modality 'future, irrealis' is marked in Old Mon through Middle Mon by the prefix [s-]; this prefix has been extensively discussed elsewhere.¹¹ Both the Lao and the Lawa forms can be explained as the strengthening of [s-] inflected forms, /s-/ > /sə-/ > /sìʔ-/, the dental fricative conditioning the front vowel. Indeed, if in OM the base has [s-], it is consistently spelt *sis-*. An alternative explanation would be to interpret Lao /sìʔ/ as a reanalyzed form of the latter. The

¹¹See Bauer 1991a and 1991b.

hypothesis is strengthened by the fact that both Lao and Lawa are spoken in former Mon contact-zones, and the chronology of /s-/ in Mon (being lost in the modern spoken varieties) supports this well.

If our assumption is correct, it is remarkable that Lawa has borrowed /si/ from Mon whereas Northern Thai has not; and given what we know about regional literary traditions—Lao texts being mostly derived from Northern Thai recensions, the latter likely to derive from Northern Mon traditions at Lamphun—that Lao has borrowed it directly from Mon.¹²

REFERENCES

- Bauer, C. 1982. *Morphology and syntax of spoken Mon*. PhD thesis, University of London (SOAS).
- Bauer, C. 1991a. Old Mon s-. In *Austroasiatic Languages—Essays in honour of H.L. Shorto*, edited by Jeremy H.C.S. Davidson. London: SOAS/Honolulu: University of Hawaii Press, pp. 241-9.
- Bauer, C. 1991b. Old Mon s-. Addenda. *MKS* 18-19:250-3.
- Coedès, G. 1952. *Inscriptions du Cambodge, IV*. Paris: EFEO.
- Ferlus, M. 1983 [1984]. Essai de phonétique historique du môn. *MKS* 12:1-90.
- Ferlus, M. 1985. Les emprunts mons en thai et en lao. In *Southeast Asian linguistic studies presented to André-G. Haudricourt*, edited by Suriya Ratanakul et al. Bangkok/Salaya: Mahidol University, 1985, p. 217-233.
- Griswold, A.B. and Prasert na Nagara. 1973. The epigraphy of Mahāddhammarāja I of Sukhodaya. *JSS* 61.1:71-178.
- Halliday, R. 1922. *A Mon-English dictionary*. Bangkok: The Siam Society.
- Huffman, F.E. 1973. Thai and Cambodian—A case of syntactic borrowing? *JAOS* 93:488-509.
- Huffman, F.E. 1986. Khmer loan words in Thai. In *Papers from a conference on Thai studies in honor of William J. Gedney*, edited by R. Bickner et al. Ann Arbor: University of Michigan, p.199-209.
- Nacaskul, Karnchana. 1962. *A study of cognate words in Thai and Cambodian*. MA thesis, University of London (SOAS).
- Nacaskul, Karnchana. 1971. *Parallelism in the use and construction of certain grammatical and lexical items in Cambodian and Thai: a typological comparative study*. PhD thesis, University of London (SOAS).
- Shorto, H.L. 1971. *A dictionary of the Mon inscriptions from the sixth to the sixteenth centuries*. London: Oxford University Press.
- Varasarin, Uraisi. 1984. *Les éléments khmers dans la formation de la langue siamoise*. Paris: SELAF/Louvain: Peeters.

March 1987
January 1990

Institute of Language and
Culture for Rural Development

¹²Sections of this article were presented to the 22d International Conference on Sino-Tibetan Languages and Linguistics at the University of Hawaii, Honolulu, 5-8 October 1989, as part of a joint paper by William J. Gedney and the present writer under the title 'Evidence for the date of the great sound changes in Thai'.