

THE MON-KHMER SUBSTRATE IN CHAMIC: CHAMIC, BAHNARIC & KATUIC CONTACT

Paul Sidwell

Centre for Research in Computational Linguistics (Bangkok)
& Australian National University
<paulsidwell@yahoo.com>

1 Introduction

The publication of Thurgood's (1999, henceforth "Thurgood") comparative-historical study of the Chamic languages puts the field of historical Chamic studies on a very solid foundation, yet also raises many new and important questions in relation to the history of language contact in Mainland Southeast Asia. Recent advances in the phonological and lexical reconstruction of Bahnaric and Katuic (Sidwell 2002, Sidwell & Jacq 2003, Sidwell 2005)¹ have clarified the internal genetic classifications of these groups, allowing inferences to be made about their historical locations and contacts with Proto-Chamic (henceforth "PC"). Comparison with the PC lexicon strongly suggests an ancient Sprachbund involving especially strong PC influence upon North and Central Bahnaric, but rather less direct contact with Katuic. The source of a substantial proportion of the apparently MK stratum in PC remains a mystery, and may well have been an extinct language (or languages) spoken in areas now home to Katuic or Bahnaric speakers.

2 Champa & Proto-Chamic

The Chamic languages are a Malayo-Polynesian sub-grouping, with speakers located today in Vietnam, Cambodia, Hainan Island (China) and Sumatra (Aceh Province of Indonesia). It appears that more than 2000 years ago a people speaking a language closely related to Proto-Malayic began to colonise the central Vietnam coast, and by the middle of the first Millennium they had established a series of allied coastal polities that we refer to conventionally as Champa. For around a thousand years the Cham settlements were important religious, military and trading centre of the region, rivaling Cambodia and Vietnam in importance. However, Champa entered an eight century long period of intermittent decline in the face of pressure from Vietnam, beginning with sacking of the northernmost centre of Indrapura (located near present day Da Nang) in 982, ultimately terminating in a rump Cham state persisting at Phan Rang into the 1830s. As the Vietnamese advanced southward Cham speakers variously resettled into the highlands, along the Mekong, and abroad by sea. Today we find the ironic situation that Chamic languages are virtually unknown in the zone of their ancient glory but are more or less alive and well on it's geographical periphery.

¹ These references were unpublished but circulating in manuscript form at the time of the SEALSXII meeting. Taking advantage of the unfortunate delay in the publication of these proceedings I took the liberty of updating the text of this paper to reflect the reality of these publications, as well as useful comments I received from colleagues, particularly Graham Thurgood and Anthony Grant.

It is reasonably assumed that during the first Millennium the ethnic Cham peoples spoke a more or less mutually intelligible language reconstructable as PC. A version of PC was offered by Lee (1966) and Burnham (1976) based upon Highland Chamic languages. Thurgood revised and expanded the reconstruction, taking into account a more representative range of Chamic languages, including Acehnese, which Thurgood controversially includes within the family. This proto-language has extraordinary characteristics: although self-evidently related to Malayic, it contains hundreds of Mainland SEAsian loans (many more than does Malay), and underwent partial phonological restructuring towards Mainland SEAsian type. This creates a very special opportunity for comparative linguistics, including the possibility that the inferred ad-strata may be subjected to internal and comparative investigation, revealing their identities and periodisation of contact influence.

Which were these contact language(s)? —the principle candidate according to Thurgood is Bahnaric, and secondarily Katuic. This is indicated in part by the fact that of the 267 MK-PC lexical parallels identified by Thurgood, the overwhelming majority have excellent Bahnaric matches. For the purposes of this paper I will assume that prolonged and intense Bahnaric-Chamic linguistic contact is accepted as a given, the outstanding issue being to identify the direction of Bahnaric-Chamic borrowings. Understandably Thurgood suggests that:

The early arriving pre-Chamic peoples most likely landed south of Danang and thus probably encountered Bahnarics. Given the restructuring of the arriving Austronesian language that took place, these pre-Chamic people must have become socially dominant, with this dominance leading many most probably Bahnaric-speaking people to shift to Cham, [p.251]

This conclusion is dependent on showing that the substantial common Bahnaric-Chamic lexicon is borrowed from Bahnaric. None-the-less it is clear that these two groups have (and continue) to experience significant linguistic contact. However, it is not so clear just how extensive Chamic-Katuic contact may have been. Thurgood's appendices list some 62 Katuic parallels, which appear to be the principle motivation for Thurgood's confidently assertion that:

...early contact between Chamic and Katuic speakers needs to be recognized anyway to account for the presence of Katuic borrowings in Chamic and Chamic borrowings in Katuic. (p.241)

and

Of particular interest are the Pacoh forms as they indicate intense contact between ancient Chams and the Katuic. (p.240)

Reference is also made to Reid's (1994) claim of an apparent morphological strata in Katuic, which Thurgood presumes is due to Chamic influence. I don't wish to offer a detailed critique of Reid's position here, but I read that paper as arguing for an Austric (and hence relation at much greater time depth) explanation for the MK-An morphological

parallels (which were earlier raised by Schmidt 1905, 1906). Notwithstanding the higher proportion of Chamic loans into Pacoh (clearly a local phenomenon), Katuic morphology is derived via its MK heritage, and not borrowed via later An contacts.

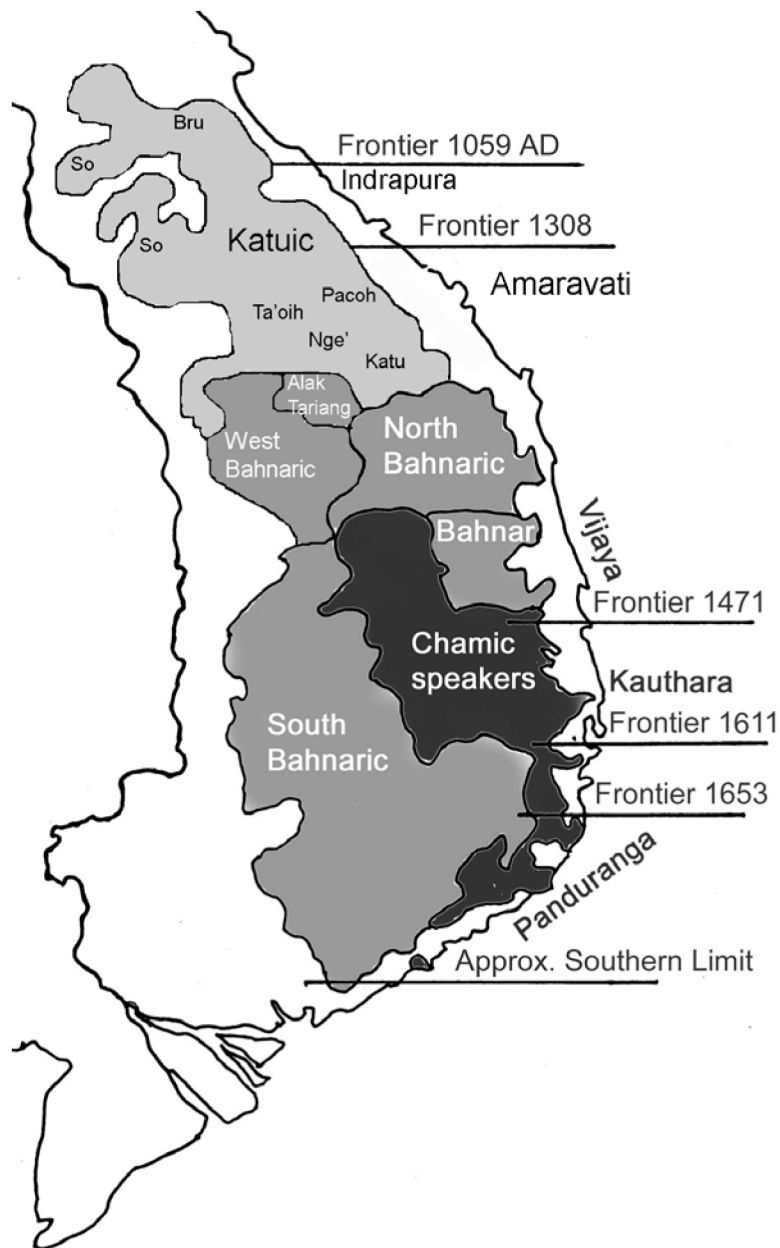


Figure 1: Map of approximate Bahnaric, Katuic and Chamic language distributions in Vietnam, and the historical Cham polities

I strongly suspect a geographical motivation strongly underlies ideas that Katuic speakers had an influence on ancient Cham, and I am sympathetic to the notion that we may gain some important insights by considering a broader synthesis of geographical and historical evidence in combination with linguistic data. It is tempting to place significance on the fact that the north-south distribution of Katuic and Bahnaric languages approximates

rather neatly the known extent of ancient Champa. The presently Katuic speaking areas lay adjacent to the old northern polities of Indrapura and Amaravati, while Bahnaric and remnant Chamic speaking areas are adjacent to Vijaya, Kauthara and Panduranga. Although the Katuic languages are certainly less Chamicised than the Bahnaric, this may simply be consistent with the lack of extensive Chamic resettlement into the Katuic zone, and the fact that the northern Cham polities were the first conquered by Vietnam.

Interestingly the main Highland Chamic intrusion (Ede and Jarai speaking) neatly corresponds in latitude to Kauthara, while the remnant coastal strip of Chamic speakers reflects the survival of an intact Panduranga into modern times.

We cannot assume that the modern distribution of Katuic and Bahnaric languages is unchanged from prehistoric times, but is rather the result of historical movements since the break-up of the respective proto-language speaking communities. Applying principles of dialect geography to the genetic classifications of these languages, we have a chance to suggest the homeland locations and migration paths of the daughter language speakers. These (necessarily tentative) results may then be correlated with what we know and presume concerning the geography of ancient Champa with what may be inferred from the distributions of loanwords. With this in mind I proceed to discuss the internal classifications of the Katuic and Bahnaric families.

3 The Katuic Family

There are approximately 15 Katuic languages spoken today, mostly in the area immediately north of the Bahnaric range and, by a huge number (perhaps a million?) Kui dialect speakers in a discontinuous spread from Laos, across into Isaan Thailand and northern Cambodia. The geographic centre of Katuic linguistic diversity lies in Laos in the hills east and north-east of Salavan, extending across into Vietnam.

It has been recognised consistently in comparative Katuic linguistics that Kui (also called Souei) sub-groups with Bru and So in a West-Katuic sub-branch (Thomas 1967 etc.). Phonologically the group is marked by a register split and vowel restructuring that is strongly parallel to that of Middle Khmer. Thus it is suggested that the Kui dialect chain, and the significant northern range of Bru and So must reflect historically late expansions of Katuic speakers under Khmer influence/dominance.

Although comparativists who have considered the issues using different methods (Thomas 1967, Efimov 1983, Diffloth 1982, Gainey 1985, Peiros, 1996, L-Tongkum 2000, Sidwell 2005) have offered various classifications, they have all agreed on the reality of the western sub-grouping and furthermore the identification of Katu (and dialects) as a distinct sub-group that branches from the highest node in the tree. On the basis of my recent historical phonological analysis (Sidwell 2005) I advocate the following classification into four more or less equidistant sub-groups:

Katuic

West Katuic:	<i>Kui, Souei, Bru, So</i>
Ta'oih:	<i>Ta'oih, Katang, Talan/Ong/Ir/nh Nge'/Kriang Chatong</i>
Pacoh:	<i>Pacoh</i>
Katu:	<i>Katu, Kantu, Phuong, Triw, Dakkang</i>

Figure 2: *Classification of Katuic languages by Sidwell (2005)*

It seems likely that the PK home-range was located within the zone still occupied by Ta'oih, Pacoh and Katu speakers today, corresponding in latitude to the old Cham polity of Amaravati. It is therefore possible that Katuic speakers were not living adjacent to Indrapura during most of the first Millennium. The lexicostatistical analysis of 50 Katuic wordlists by Miller & Miller (1996) found Kui and Bru dialects to be rather close, sharing between 60% and 75% of basic vocabulary (depending on the various lists compared), and consistently less than 60% with other Katuic languages. This would suggest a West Katuic dispersal from late in the first Millennium, first northward into the Indrapura hinterland, and then south-west over the Mekong.

4 The Bahnaric family

Bahnaric is an MK family of more than 30 languages spoken in Vietnam, Cambodia and Laos. Most comparative-historical studies of the Bahnaric languages have been focussed on the smaller, self-evident subgroups, and even purported Proto-Bahnaric studies (such as my own PhD thesis of 1999) have not yet considered the totality of the family. Foundational works such as Thomas (1966), the Proto-North-Bahnaric of Smith (1972) and the Encyclopaedia Britannica entry of Diffloth (1974) distinguished a basic north-south split, with the most important discussion concerning whether Bahnar groups with the northern or southern divisions (especially Gregerson, Smith & Thomas 1976).

I recently presented a classification of the Bahnaric languages based upon historical phonology (Sidwell 2002). This scheme divides the family into 3 main branches, as follows:

Bahnaric

North:	<i>Sedang, Hre, Rengao, Jeh, Halang etc.</i>
West:	<i>Loven, Lawi, Cheng, Oi, Laveh/Brao, Nhaheun etc.</i>
Central:	
West Central:	<i>Taliang/ Kasseng</i>
North Central:	<i>Alak</i>
East Central:	<i>Cua</i>
South Central:	<i>Tampuon, Bahnar, South Bahnaric (Stieng, Chrau etc.)</i>

Figure 3: *Classification of Bahnaric languages (Sidwell 2003)*

Both West Bahnaric (WB) and North Bahnaric show a degree of geographical and linguistic unity indicative of relative stability since the initial break-up of Proto-Bahnaric. By contrast there is geographically discontinuous distribution and greater internal diversity among Central Bahnaric (CB), implying several significant dispersal events.

Highland Chamic speakers of Jarai and Ede separate the South Bahnaric (SB) from the Bahnar, interrupting what once must have been a continuous community. On lexicostatistical grounds one may estimate that the linguistic separation of Bahnar and SB began up to two thousand years before present. According to my analysis the SB languages share about two thirds (or a little less) of basic lexicon with each other, but consistently less than half of basic lexicon with other Bahnaric languages. This suggests that SB began to disperse in the latter half of the first Millennium, having already become distinct from Bahnar etc. significantly earlier.

An even earlier separation from CB is reflected in the positions of Alak and Taliang, spoken in Laos adjacent to WB and Katuic communities. They share less than half their basic lexicons with each other and the other Bahnaric languages. Furthermore Alak and Taliang lack any significant stratum of Chamic loans, instead showing various Katuic and Khmer borrowings, in fact many of the same borrowings as found in WB languages.

This suggests the modeling of a relative chronology as follows:

1. In the earliest phase WB, NB and CB became established in approximately their present locations north of, and perhaps somewhat overlapping, the nowadays Ede speaking area, with a wedge of CB (pre-Alak, pre-Taliang) speakers between the western and northern groups.
2. In the second phase Chamic influence extends inland over the NB and the main CB speaking areas. WB and the northern wedge of CB speakers avoid Chamic influence, forming linkages with Katuic.
3. Chamic influence in the Central Highlands builds, and at some point a major dispersal of CB speakers southward and eastward was triggered.

I speculate that the phases 1 and 2 correlate with the emergence during the first Millennium of rivalries between Cham polities and Cambodia (Funan?). It is also possible that for some time into the first Millennium Katuic and Bahnaric speakers were mainly located in the hinterlands of Chamic Amaravati, Vijaya and Kauthara, and the extent of linguistic contacts and influences were conditioned by the specific circumstances of each of these polities (internal government, economy, external relations etc.). One outcome of these early conditions/events was the emergence of an areal-linguistic divide that ran through the Highlands, between the zones of Chamic and Cambodian influence (corresponding to CB, NB and Highland Chamic versus Katuic, WB and outlying CB speaking zones).

5 Analysis of lexical comparisons and conclusions

Thurgood's search for the MK sources of Chamic lexicon focused specifically upon Bahnaric and Katuic, mindful of their known geographical proximity to Chamic. It is also a practical reality that broad lexical resources for comparative MK research, such as

etymological dictionaries, were not readily available.² Consequently this preliminary exploration of the question missed an important consideration – without knowing the wider MK distributions and phonological forms of the relevant MK etyma it is not possible to reliably assess the origin and direction of borrowing of specific forms.

Focusing specifically on the Katuic issue, I extracted the 62 Chamic etymologies where Thurgood adduced a Katuic comparison, and I subjected these to a broader etymological analysis, so far as my limited dataset would allow. The results are presented in the Appendix, where the items have been organised into categories reflecting which MK form shows the closest matches to PC.

My assumption is that, although a complete treatment of Thurgood's PC etymologies is desirable, an indication of the likely extent of ancient Chamic-Katuic contact should be revealed by such an investigation. I am specifically concerned to distinguish comparisons of the following types:

- Unique PC-Katuic isoglosses, being the best evidence of specific contact.
- Widely distributed MK forms which may have been borrowed from more or less any MK source.
- Forms with specific phonological features marking them as likely borrowed from an identifiable language.

The results are quite striking, the most important summarised as follows:

- For only one item out of 62 was it possible to identify a unique Chamic-Katuic isogloss (item 1.a in the Appendix).
- 21 items are widely distributed MK etyma with no specific phonological features identifying a likely specific source of borrowing into Chamic.
- For 9 items that show a good Katuic match, there is an equally good Bahnaric match, and the restricted distributions indicate borrowing in from Chamic.
- 8 items resemble various MK words, but are not shown to be of MK origin.
- 31 items show equal or better matches with other MK branches, including Bahnaric, Vietic, Khmer, Mon, Northern MK. Allowing for the likelihood that a significant proportion of the Bahnaric and Katuic comparisons reflect inward loans from Chamic, there is no single standout MK language identifiable as the major donor to Proto-Chamic.

It would appear that there is little empirical evidence that would unambiguously “indicate intense contact between ancient Chams and the Katuic” (Thurgood p.240). And on the basis of this brief survey one may well speculate that the contact with Katuic such as it was (some dozens of words only?), was decidedly one-way, with borrowing overwhelmingly from Chamic. Even the hundreds of Bahnaric-PC comparisons appear to be mostly borrowed from Chamic, occurring after important splits among Bahnaric speakers.

This raises a greater riddle – what were the sources of the hundreds of clearly MK loans into PC (not to mention hundreds of words of unknown origin) that helped to

² Fortunately this situation is beginning to change, for example I can report that in 2006 Harry Shorto's *A Mon-Khmer Comparative Dictionary* was published by Pacific Linguistics, Canberra.

condition its dramatic lexical and phonological restructuring? I simply do not yet know. However, the reconstructed linguistic geography discussed above raises some tantalising possibilities, since during early stages of Chamic settlement it is quite likely that much of the Indo-Chinese hinterland now or recently occupied by Bahnaric and Katuic speakers was inhabited by speakers of other MK languages. I can only suggest that much more work needs to be done if we are to properly reconstruct and understand the relexification and restructuring of the ancient Chamic language.

References

- Burnham, Eugene C. 1976. *The Place of Haroi in the Chamic Languages*. M.A. thesis, University of Texas at Arlington.
- Diffloth, Gérard. 1974. Austro-Asiatic Languages. Encyclopaedia Britannica. Chicago/London/Toronto/Geneva, *Encyclopaedia Britannica* Inc. pp 480-484.
- Diffloth, Gérard. 1982. Registres, dévoisement, timbres vocaliques: leur histoire en Katouïque. *Mon-Khmer Studies* 11:47-82.
- Efimov, Aleksandr. 1983. *A phonological reconstruction of Proto Katuic* (in Russian). PhD thesis, Institute of Far Eastern Studies, Moscow.
- Gainey, Jerry. 1985. *A comparative study of Kui, Bruu and So phonology from a genetic point of view*. Bangkok, Thailand: Chulalongkorn University thesis.
- Gregerson, Kenneth J., Kenneth D. Smith, and David D. Thomas. 1976. The place of Bahnar within Bahnaric. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds.), *Austroasiatic studies, part I*. Oceanic Linguistics Special Publication, 13. Honolulu: University of Hawaii Press. pp.371-406.
- Jacq, Pascale. & Paul Sidwell. 2000. *A Comparative West Bahnaric Dictionary*. München, Lincom Europa.
- Lee, E.W. 1966. *Proto-Chamic Phonologic Word and Vocabulary*. Ph.D. dissertation, Indiana University.
- L-Thongkum, Therapan. 2000. *Languages of the Tribes in Xekong Province Southern Laos: a foundation for research and development*. Bangkok, Chornalonghorn University.
- Miller, John and Carolyn Miller. 1996. Lexical comparison of Katuic Mon-Khmer languages with special focus on So-Bru groups in Northeast Thailand. *Mon-Khmer Studies* 26:255-290.
- Peiros, Iliia. 1996. *Katuic comparative dictionary*. Canberra, Pacific Linguistics C-132.
- Reid, Lawrence A. 1994. Morphological evidence for Austric. *Oceanic Linguistics* 33(2):323-344.
- Schmidt, Pater Wilhelm. 1906. Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens. *Archiv für Anthropologie*, Braunschweig, 5:59-109.
- Schmidt, Wilhelm. 1905. Grundzüge einer Lautlehre der Mon-Khmer-Sprachen. *Denkschrift der Akademie der Wissenschaften*, Wien, Philologisch-Historische Klasse 51:1-233
- Sidwell, Paul & Pascale Jacq. 2003. *A Handbook of Comparative Bahnaric: Volume 1, West Bahnaric*. Canberra, Pacific Linguistics 551.
- Sidwell, Paul. 2000. *Proto South Bahnaric: a reconstruction of a Mon-Khmer language of Indo-China*. Canberra, Pacific Linguistics 501.
- Sidwell, Paul. 2002. Genetic classification of the Bahnaric languages: a comprehensive review. *Mon-Khmer Studies* 32:1-24.

- Sidwell, Paul. 2005. *The Katuic Languages: classification, reconstruction and comparative lexicon*. Munich, Lincom Europa.
- Smith, Kenneth, D. 1972. *A phonological reconstruction of Proto-North-Bahnaric*. Dallas, Language Data Series, Summer Institute of Linguistics.
- Thomas David. 1966. Mon-Khmer subgroupings in Vietnam. In Norman Herbert Zide, ed. *Studies in Comparative Austroasiatic Linguistics*. Indo-Iranian Monographs, V. The Hague, pp 194-202.
- Thomas, Dorothy. 1967. *A phonological reconstruction of Proto East Katuic*. MA thesis, University of North Dakota.
- Thurgood, Graham, 1999. *From Ancient Cham to Modern Dialects: two thousand years of language contact and change*. Honolulu, University of Hawai'i Press.

APPENDIX: Etymological analysis of 62 PC-MK comparisons with Katuic citations drawn from Thurgood (1999)

Note: Vetic reconstructions are from Ferlus (pers. com.) other MK reconstructions are my own unless otherwise marked. Abbreviations are as follows:

An	Austronesian	PMK	Proto-Mon-Khmer
MK	Mon-Khmer	PMP	Proto-Malayo-Polynesian
PB	Prot-Bahnaric	PSB	Proto-South Bahnaric
PC	Proto-Chamic	PV	ProtoViet-Muong
PK	Proto-Katuic	SB	South Bahnaric

1) Unique PC-Katuic matches

- a) PC ***baal** ‘mend; to patch’
Thurgood gives a PK form *_p_l ‘to patch’ (from Thomas 1967), no broader etymology apparent.

2) PC-Bahnaric matches

- a) PC ***caɗiŋ** ‘finger’
Cf. Bahnar **ɗeŋ**, Rengao **hədiin** lack wider MK etymology, suggesting borrowing from Chamic. PK ***ɗoŋ** ‘limb of animal/plant’ is not obviously related.
- b) PC ***ɗaŋ** ‘lie supine; be on back’
Bahnar **pəɗaŋ** ‘lie face up’ is isolated in MK and therefore likely borrowed from Chamic. It is not obvious that Katu **ləlaŋ** ‘lie on back’ can be related.
- c) PC ***kɔɔʔ** ‘white’
Bahnar **kɔɔk**, Köho (SB) **kɔɔ** ‘white fur’ etc. appear to be borrowed from Chamic, other MK suggest PMK ***klɔɔk**.
- d) PC ***lɔɔk** ‘peel’
Cf. PK ***lok** ~ ***lɔɔk**, Bahnar **plɔʔ**, PSB ***plook**. The distribution and irregularities among the MK forms suggests borrowing from Chamic and/or other source. Katuic forms such as Pacoh **lɔt** ‘peel back edges’ are not related.

3) PC-Katuic-Bahnaric matches

- a) PC ***?awaak** ‘spoon; ladle’
Bahnar **?waak** and Katu **waaʔ** appear to match, but other MK forms show front vowel reflexes corresponding to Khmer form **wæək** ‘ladle’, suggesting that Bahnar, Katu may have borrowed from Chamic, cf. Cebuano Bisayan (Austronesian) **luwag** ‘ladle’.
- b) PC ***?ale** ‘medium bamboo’
PK ***?aleɛ** ‘type of bamboo’, Bahnar dialects ***?aleɛ**, ***leɛ**.
South Bahnaric ***glee** ‘small bamboo’, Khmer **khləj** and Viet. **tre** suggest PMK

gliʔ. Note also North Bahnaric **pələe** ‘type of bamboo’.

On balance the source of *ʔaləe forms is not clear, although the ʔa presyllable is a typically Katuic feature.

- c) PC ***biluay** ‘gourd’
Cf. PB ***pluuɟ**, PK *ʔ**aluuj** - the latter lacks initial labial, but could be secondary in Katuic. Not found elsewhere in MK.
- d) PC ***cuah** ‘sand’
Cf. Bahnar **cwəh**, PNB ***cuas**, PK ***cuah** are clearly borrowed from Chamic on phonological criteria; other MK suggest PMK ***ksaac** ~ ***skaac**.
- e) PC ***haŋɔ** ‘pine’
Cf. NB and Bahnar **həŋɔɔ** ‘pine’, PK. ***sŋɔɔ** ‘pine’.
Katuic initial /s/ suggests the proto-form.
- f) PC ***jaaŋ** ‘spirit; god’
Bahnaric and Katuic suggest ***jaaŋ** ‘spirit’, but An forms (Malay, Javanese etc. **hyang** ‘deity’) suggests origin in Western Austronesian.
- g) PC ***raw** ‘wash’
Cf. PB ***raaw**, PK *ʔ**araaw** versus Khmu’ **raa**, Viet. **rũa**.
- h) PC ***padiəŋ** ‘carry; transport’
Peiros PK ***ptiaŋ** ‘haul, transport’ on the basis of Bru **pətiəŋ** and Pacoh **pətiəŋ**; also SB forms such as Mnung **pədiəŋ**.
These data suggest borrowing from Chamic, with subsequent regular devoicing in Katuic.
- i) PC ***picah** ‘broken; break’
Cf. Malay **pecah** ‘scattered’. Restricted distribution of MK forms: (WestKatuic) Bru **pəcàh**, Kui **kəcàh** (hence Peiros’ PK ***pəc[ə/a]h**~***kəc[ə/a]h** ‘crack’), Bahnar **pəcah**, **cah**, SB e.g. Stieng **cah** ‘break’.
The lack of regular c > s shift indicates borrowing into Bahnar and SB indicates borrowing from Chamic.

4) PC-Katuic-Vietic matches

- a) PC *ʔ**eh** ‘excrement; defecate’
Cf. PK *ʔ**eh**, PV *ʔ**eh**; other MK reflect a different root: *ʔ**ic** ~ *ʔ**ac**.
- b) PC *ʔ**juɔp** ‘smoke tobacco’
PV *ʔ**juɔp** ‘to smoke’ (attested in Arem and Sáčh), Katuic; Bru **juòp**, Pacoh **ʔjuɔp** ‘to suck’.

5) PC-Katuic-Khmer matches

- a) PC ***rəh** ‘cut’
Cf. PK ***rəh** ‘cut/slash’, Khmer **rəh** ‘to cut with short, careful strokes’.

6) PC-Katuic-Northern MK matches

- a) PC ***klaj** ‘penis’
Cf. PK. ***klæ** ‘penis’, Kammu-Yuan **tlé?**, Riang-Lang **kle?** etc.

7) PC-Katic-Vietic-Northern MK matches

- a) PC ***krɔŋ** ‘river’
Cf. PB **kroŋ**, PV ***krɔŋ**, Old Mon **kruŋ**, Riang-Lang [**om**] **_kroŋ**, Kammu-Yuan **krɔŋ** etc., PK ***kruuŋ** ‘forest’.
Vietic and Khmuic showing best phonetic match to Chamic.

8) PC-Bahnaric-Vietic matches

- a) PC ***dəŋ** ‘stand; stop’
PV ***təŋ?** > Viet. **đứng** [d̪iŋ] ‘get up’ could be the direct source of Chamic and Bahnaric forms (e.g. Bahnar **dəŋ** ‘stand, stop’) if we can reconcile the timing of the initial voicing change in Vietnamese.
- b) PC ***kaaŋ** ‘chin; jaw’
Cf. PB ***kaaŋ**, PV ***kaaŋ?**. Note also similar widely distributed Kadai forms such as Thai **k^haaŋ**.
PK ***tbaaŋ** (mentioned by Thurgood) is not related.
- c) PC ***cim** ‘bird’
Only SB and Vietic suggest ***cim** with short vowel, while other MK attest a long vowel (e.g. PK ***ceem**, Bahnar **sɛem** and others).

9) PC-Khmer matches

- a) PC ***buhaj** ‘otter’
PMK ***bhe?** > OKhmer **bhee** suggests the closest match to Chamic; Vieto-Katuic forms have medial **-s-** rather than **-h-**, reflecting a secondary development.
- b) PC ***padaw** ‘warm, hot’
Mon, Khmer, Katuic, Bahnaric forms suggest PMK ***ktuu?**. Voicing/glottalisation shift in OKhmer **ktau** > **kdau** proves Khmer as source for Chamic.
- c) PC ***kɔŋ** ‘bracelet’
Cf. PB ***koŋ**, PK ***kɔŋ**, Khmer **kɔŋ/kɔŋ**.
- d) PC ***laŋi** ‘basket, winnowing’
Khmer **lɔŋy** appears to be the source of Chamic and various Bahnaric forms, e.g. Sre **ləŋii**, Rengao **rəŋii** and others.
Thurgood noted Peiros’ PK ***ʔərie**~***kərie** ‘k.o. basket’, but may rather be related to Khasi **khri** ‘small basket’.

10) PC-Northern MK matches

- a) PC ***kacuh** ‘spit’
Cf. PB ***kcoh** ‘spit’, PK. ***kcoh** ‘spit’, Mưong **chủ**, Kammu-Yuan **kəcúh**, Riang-Lang **təkchu**.
Northern MK forms agree more consistently with PChamic.
- b) PC ***ruah** ‘choose’
Cf. Bahnar **rəjh**, PSB ***ruəh**, PK ***rəəs**, OKhmer **res /rəəh/**, PWaic **rəs**.
MK forms suggest PMK ***rəəs**. Chamic vowel ***ua** may be explained by diphthongisation of [ɔ] or similar vowel attested in Waic and other NMK. SB clearly back-borrowed from Chamic, since the diphthong cannot be derived from Bahnaric vocalism.
- c) PC ***plum** ‘land leech’
Cf. PB ***pləəm**, PK ***pləəm**, PV ***pleem**, PASlian ***pləm**, Mon **kləm**, Khasi **thliem**, Palaung **plum**, PWaic ***pli/om**.
The short back-vowel reconstructed by Thurgood is consistent with various Northern MK forms.

11) PC-Mon matches

- a) PC ***[kulit]** **səʔ** ‘lungs; placenta’
Mon **kəsəʔ** provides a unique phonological match, NMK forms such as Lawa **hməuʔ** may also be related by infixation.
PB ***səh**, PK ***səh** etc. cannot be connected to Chamic.

12) PChamic matches broadly distributed MK forms with no specific indication of source in PC:

- a) PC ***ʔaha**, ***ha** ‘open (mouth to say sthg.)’
Mon, Khmer, West Katuic, Bahnaric, Vietic all suggest PMK ***haʔ**. Note also Viet. **hả** suggests ***hah**.
- b) PC ***ʔəʔ** ‘vomit’
Corresponding MK forms in Mon, Bahnaric, Palaungic, Nicobaric suggest such a PMK root.
PK ***həək** ‘choke on food’ is a different etymon.
- c) PC ***həʔaap** ‘yawn’
Wide distribution in MK suggests PMK ***sʔaap** ‘yawn’, specific Chamic source not obvious (cf. PMP ***uap** ‘yawn’).
- d) PC ***kalaəŋ** ‘hawk; bird of prey’
All MK branches show forms indicating PMK ***klaəŋ** or similar.
- e) PC ***klaas** ‘escape’
Widespread MK comparisons, e.g. PB ***klaas**, Bru (West-Katuic) **rəklah**, Mon **pləh**, Khasi **khlad** ‘to separate’, Kammu-Yuan **kláh** ‘to exceed’, etc.
Etymon is clearly MK in origin

- f) PC ***səna** ‘crossbow’
Khmer, Katuic, SouthBahnaric and Vietic suggest PMK ***snaʔ**, with no specific indication of which may be the source for Chamic, although it is likely that SB borrowed from Khmer or Chamic.
- g) PC ***ruaj** ‘fly; bug; insect’
Katuic, Bahnaric, Khmuic, Palaungic, Viet-Muong and Aslian all suggest PMK ***ruəj**, so there no specific indication of the source for Chamic.
- h) PC ***sagər** ‘drum’
Mon, Khmer, Bahnaric and Palaungic suggest PMK ***sgər**, while Katuic forms support **sgir** or **sgir**.
- i) PC ***tram** ‘soak’
Khmer, Katuic and Bahnaric suggest PMK ***tram**.
- j) PC ***səh** ‘only; empty; free; leisure’
Khmer, Bahnaric, Katuic suggest PMK ***səh**.
- k) PC ***sula** ‘leaf’
All MK suggest PMK ***slaʔ**.
- l) PC ***kuah** ‘shave, scrape’
Widespread MK reflexes suggest PMK ***kuəs** ‘scrape, scratch’, however there are also similar Austronesian forms, e.g. Iban **kukus** ‘to scrape out’.
- m) PC ***luaj** ‘swim’
The root is widespread in MK, although vowel reflexes considerably making reconstruction problematic – no obvious source of /ua/ apparent.
- n) PC ***prək** ‘squirrel’
All MK branches suggest PMK ***prək**, leaving no specific source for Chamic apparent.
- o) PC ***pəh** ‘open’
Cf. Palaung **puh**, Bahnar **pəh**, Nancowry **fuáh**, PV ***pəh/bəh**, PK ***pəh/poh/pooh**, suggesting various possible PMK reconstructions, and no specific indication of source for Chamic.
- p) PC ***hək** ‘pour out; spill’
PK. ***hək** ‘pour’, Biat (SB) **huk** ‘pour out’, Palaung **thəʔ** ‘id.’, Khasi **thəh** ‘pour, spill’ suggest Mk etymology, but reconstruction is problematic and specific Chamic source not obvious.
- q) PC ***pər** ‘fly’
Bahnaric & Katuic suggest ***par**, other MK suggest ***pər** (e.g. PV ***pər**, Palaung ***pər** and others).
- r) PC ***haŋ** ‘hot; spicy’
Bahnaric, Vietic and Nicobar suggest PMK ***haŋ**.
- s) PC ***pah** ‘slap’
Cf. Khmer **pah** ‘hit, strike’, PV ***t-pah**, PK ***pah** ‘slap’, PSB ***pah** ‘hit’. Other

MK show back vocalism, e.g. Palaung **poh** ‘slap’. Malay **pepah** ‘strike, beat’ with /a/ < *ə suggests a competing An source for Chamic.

- t) PC ***bala** ‘tusk; ivory’
 Reflexes with /a/ main-vowel attested in North Aslian, Vietic (Arem & Sáčh only) and Bahnaric.
 However, SB, WB and other MK have forms with a back vowel suggesting PMK ***bluk** ~ ***bluək**.
- u) PC ***sapal** ‘arm’
 Viet. **vai** ‘shoulder’, PK **ʔapaal** ‘shoulder’, Temiar (Aslian) **pal** ‘arm’ are suggestive, but we don’t have indication of the direction of semantic shift ‘shoulder’ ↔ ‘arm’.

13) PC forms resemble, but do not match MK forms:

- a) PC ***buc** ‘uproot, pull up; scratch’
 Cf. PK ***pooc**, PV ***pɔɔc**, WKhmer **baoc** ‘pull out’.
 Voice quality and vowel differences mean that explaining a relation between Chamic and MK forms is problematic.
- b) PC ***gaj** ‘stick; wood’
 Thurgood compares PK ***gir** ‘rod’, but finals do not agree.
- c) PC ***gulac** ‘return; go home; again’
 Thurgood compares (Peiros)PK ***kə[lh]aajh** ‘compensate for, return’, but the comparison is without foundation. Likely MK parallels are not evident.
- d) PC ***guj** ‘carry on back’
 Cf. PK ***kuuj** ~ ***kuj** ‘carry on back/shoulder’, PB **kɔɔj** ‘id.’ (with some regular short vowel reflexes, e.g. Hrê **kuj** ‘carry on head’).
 Voicing of the Chamic initial is not explained by Katuic/Bahnaric forms.
- e) PC ***klah** ‘lose’
 Forms such as Bahnar **klah** ‘to part, split’, PK. ***klah** ‘split’ do not show a reasonable semantic match.
- f) PC ***tubuac** ‘beak’
 PK ***crɔɔh** and PV ***k-ɔɔh** are suggestive, but do not explain the Chamic laminal final.
- g) PC ***siyaam** ‘good; nice; pretty’
 Compared by Thurgood to Katuic and Bahnaric forms suggesting ***liam**, but these cannot be the source of Chamic.
- h) PC ***pə-pət** ‘to fan’
 Thurgood compares various MK forms, but none provide a reasonable phonological match (no MK form with –t final evident).

14) Areal forms possibly via Malay:

- a) PC ***kapaas** ‘cotton’
Similar forms are found throughout Eastern and Northern MK languages, apparently borrowed from Sanskrit **karpāsa**/Pali **kappāsa**.
- b) PC ***ribuu?** ‘storm’
Probably a variant of the An root reflected in Malay **ribut**, and loaned into Bahnar **həbut** ‘wind to blow hard’. PK ***rpuu?** likely borrowed in from Chamic.
- c) PC ***kaduŋ** ‘pocket; pouch’
Cf. Bahnar **kəduŋ** ‘pocket, pouch’, PK. ***kduŋ** ‘bag, pocket’, OKhmer ***kəndaŋ** ‘basket made of leaves’, Malay ***kandung** ‘pocket, pouch’.

