Dating the separation of Acehnese and Chamic by etymological analysis of the Aceh-Chamic lexicon[*]

Paul SIDWELL
Australian National University, Australia

Abstract

The genetic classification of Acehnese vis-à-vis Chamic is an issue whose solution may contribute important new insights for the history of Austronesian settlement and language contact in mainland SEAsia. This paper critically examines the hypothesis of Thurgood (1999) that Acehnese is a Chamic language that separated from the family from around the end of the 10th century CE. The Acehnese reflexes listed in Thurgood's Proto-Chamic (PC) lexicon are subjected to etymological analysis, looking for phonological and lexical indications of the place of Acehnese within Malayo-Chamic. Also the common borrowed lexicon is examined and its implications discussed. The results indicate that Acehnese does sub-group with Chamic, but that the separation probably occurred rather earlier than Thurgood suggests, I argue that it was well before the bulk of Mon-Khmer (MK) lexical influence on Proto-Chamic occurred.

1.   The problem of Chamic & Acehnese

The Chamic languages are a Western-Malayo-Polynesian (WMP) sub-grouping that, according to Blust's (1994:47) estimate separated from the ancestor of Malay around 200~300 BC. By the second century of the first Millennium the Chamic peoples, under the leadership of a Hindu elite, had begun to build one of the great classical civilisations of Southeast Asia. Champa flourished for the next 800 years, until Vietnam, freed of Chinese dominion, began to assert its own colonial ambitions. In 982 Vietnamese occupied the northernmost provinces of Champa, including the capital Indrapura, and a complex history of conflicts, territorial exchanges and occupations followed. In 1471 the southern Cham capital Vijaya was also lost to the Vietnamese, and many Chams took refuge in the highlands; some are known to have gone to Aceh. In the 1700s many sailed up the Mekong River to Cambodia—the 'Western Chams'. A rump Chamic State persisted on the

---

Vietnamese coast in the vicinity of Phan Rang until 1835, where a Cham speaking minority still lives today.

The language of first Millenium Champa is reconstructable by application of the Comparative Method to the ten or so modern Chamic tongues, giving us a phonology and lexicon of Proto-Chamic (Lee 1966, Burnham 1976, and particularly Thurgood 1999, henceforth 'Thurgood'). The so derived proto-language has a number of structural features more typical of MK than Austronesian: fixed final stress, reduction of initial syllables, a series of imploded stops, a large vowel inventory with phonemic vowel length, all linked historically with a significant lexicon derived by borrowing from MK languages, much of it probably due to language shift. The recent reconstruction of Proto Chamic by Thurgood has significantly improved our understanding of the phonetic and lexical evolution of Chamic, yet a number of important questions remain to be satisfactorily resolved, especially in relation to sequencing and dating the various internal Chamic splits, and identifying the specific contact languages that so heavily influenced its transformation.

According to Thurgood's model, for most of the first Millenium PC was a linguistic unity, undergoing transformation as the local Mon-Khmer speaking population were absorbed into Champa. The result of this process was the Mon-Khmerised PC represented by Thurgood's lexical reconstruction. Acehnese is assumed to have split off sometime around 982, and this can be treated as the highest branching split within the Chamic family. By implication, the PC lexicon should be well represented in Acehnese; in fact, theoretically, the reconstruction of PC lexicon should substantially depend upon its witness in Acehnese. However, this is not quite the case, as there are relatively few Acehnese reflexes of Chamic-MK borrowings among Thurgood's etymologies, and my own efforts to expand the set have met with little success. This fact alone suggests that if Acehnese is a Chamic language, it may have separated before much of MK contact effects on Chamic. Significant structural innovations common to Chamic and Acehnese may be largely the results of parallel internal structural drift.

The prehistory of Champa is a matter over which conflicting views have been offered. Shorto (1975) saw Chamic as a Mainland remnant of Austronesian that had evolved in situ. An extreme version of this hypothesis was offered by Collins (1996), who asserted that Austronesians originated on the dry Sunda Shelf during the last Ice Age, retreating upland as rising sea levels turned mountains into islands. However, in recent decades an alternative view has gained widespread acceptance among linguists, namely that the Austronesians colonised the Indonesian archipelago from Formosa, via the Philippines, from around 4000~3500[1] years ago (Blust 1994). Accepting the latter hypothesis we may posit that during the first Millenium BCE pre-Malay peoples spread from Borneo to Sumatra, the Malay Peninsula and the SEAsian mainland.

---

[1]Thurgood (1999:5) suggested that Austronsians expanded "out in to the Pacific some six or eight thousand years ago."

Thurgood, correlating reconstructed linguistic events with archaeological data, places the beginnings of Austronesian settlement in Indo-China from circa 600 BCE, suggesting that:

> It was some of these Austronesians-speaking people who, after extended contact with the MK people then living along the coast, would then become the speakers of Chamic. (p.16)

And speculates that:

> The early arriving pre-Chamic people most likely landed south of Danang and thus probably encountered Bahnarics. Given the major restructuring of the arriving Austronesian language that took place, these pre-Chamic people must have become socially dominant, with the dominance leading many most probably Bahnaric-speaking people to shift to Cham, but bringing with them many MK characteristics. (p.251)

We can reconcile this chronology with Blust's (1994) dating and classification of Malayic, Chamic and Acehnese languages into a sub-group he calls Malayo-Chamic (MC), by suggesting that an MC colonisation of the Asian mainland began around the middle of the first Millennium BCE, with a distinctive Chamic identy emerging a few hundred years later. However, where and how this process occurred is also a matter of differing views. While Thurgood (1999) has Acehnese and Chamic united on the mainland until the 10th Century, Dyen (2001:392) suggests that:

> If Achenhese is part of the Malayo-Chamic occupation of Sumatra, the settlements on the coast of Vietnam could have originated from there and, if they did, from northern Sumatra. Such a hypothesis would just as simply account for the "parallels" or agreements between Achehnese and Chamic, and would have the advantage of reducing the advance of Malayo-Chamic to a simple spreading from southern Sumatra, first to Acheh and then, as Achenhese, leaving the single colonisation from there to Champa as yet to be posited.

Dyen's scenario holds that Acehnese has always been spoken in Sumatra, not withstanding any later Chamic migrations. Given that Chinese sources unambiguously identify Champa (under the name *Lin-yi*) from 190 CE, Dyen's Aceh-Chamic split would be early indeed, perhaps even before the Common Era.

Thurgood, quite reasonably, seeks to associate his linguistic analysis with known historical events, principally the great Vietnamese "push to the south", and the subsequent multiple emigrations of Chamic peoples seeking refuge from conflict and foreign domination. He speculates that a series of emigrations, with possible intermediate settlement in Malaya, unfolded over hundreds of years until we find the first written record of a Cham voyage to Aceh in the late 1500s. Dyen (2001:391) cites Brown (1953):

> The Sejarah Malay [....] records that after the fall in 1471 of Vijaya, the then Cham capital, Prince Syah Pau Ling went to Acheh and started a line of Acheh kings, presumably in colonies earlier established.

For me the last few words of the quote beg the question—did Prince Syah Pau Ling take his people to an established Chamic colony, or to a sympathetic country with long standing ties of blood and/or common (Islamic) religion? Was it a reuniting of two societies that had been separated for perhaps as long as a millennium and a half? To begin to shed some light upon these questions we should revisit Thurgood's linguistic arguments for a late Chamic-Acehnese separation with a critical eye.

2.   Linguistic arguments for dating Aceh-Chamic unity/separation

The key components of Thurgood's results are the reconstruction of PC lexicon and phonology. Assuming that those results give a more or less fair representation of first Millennium CE Chamic (in lieu of a comprehensive reworking of the reconstruction), the challenge is to see if Acehnese is better explained as a descendent of late first Millennium PC, or an offshoot of a much earlier time that may even have received a wave Chamic influence in the wake of the collapse of 1471. A detailed etymological analysis will hopefully give an indication of the likely historical scenario.

Thurgood's PC lexicon is internally classified as follows: 285 items of direct Austronesian origin, 277 from Mon-Khmer into Proto-Chamic, and 179 words of unknown origin, 20 from Indic, 3 from Arabic, (plus some 173 items borrowed after the breakup of Proto Chamic). Below I examine the inherited Austronesian vocabulary and the borrowed strata separately.

2.1 *Aceh-Chamic lexicon of direct Austronesian origin*

Thurgood's PC lexicon of Austronesian origin is the least problematic stratum, since it is identified on the basis of well established Austronesian reconstructions. For the 285 etymologies Thurgood listed some 203 Acehnese comparisons were adduced. To this we can add at least eight more items with Austronesian etymologies misclassified by Thurgood (see Appendix 1a.). This stratum shows multiple non-trivial sound correspondences that demonstrate that Acehnese belongs to the MC subgroup of WMP. Blust (1994) identified three principal sound changes that define this group (exemplified with data drawn from Thurgood; note Malay words in authography):

1) PMP *R* > PCM *r*
   PMP *Rusuk* 'ribs', Malay *rusuk*, Aceh. *rusoʔ*, PC *rusuk*
   PMP *daRaq* 'blood', Malay *darah*, Aceh. *darah*, PC *darah*

2) PMP *w-* > PCM *Ø-*
   PMP *wahiR* 'water', Malay *air, ayer*, Aceh. *ʔiə*, PC *ʔiar*
   PMP *wakaR* 'root', Malay *akar*, Aceh. *ʔukhuə*, PC *ʔughaar*

3) PMP *q > PCM *h
   PMP *$baseq$ 'wet', Malay *basah*, Aceh. *basah*, PC *$basah$
   PMP *$qataj$ 'liver', Malay *hati*, Aceh. *ʔate*, PC *$hataj$

Note that word initially the Acehnese reflex is /ʔ/, requiring a sequence *q > *h > /ʔ/. This is also the occasional reflex in Malay, e.g. *abu* 'ashes' < PMP *$qabu$.

So it would appear reasonable that MC is well justified, but further evidence is needed to demonstrate subgrouping Acehnese with Chamic, in other words that a single community split from MC and established a distinct PC (or pre-Chamic) language before the separation that created Acehnese. Unrelated languages may well acquire common borrowings within a linguistic area, so for our purposes we need to examine the post-MC sound changes evident within the etymologically Austronesian lexicon (examples drawn from Thurgood):

1) Reflexes of PMP disyllables in Acehnese and Chamic are reduced to clusters where the first syllable is unstressed, and the second consonant is one of /r, l, h/. For example: PMP *$beli$ 'buy' > Malay *bĕli*, Aceh. *blɔə*, PC *$blɛj$; PMP *$puluq$ 'ten' > Malay *puloh*, Aceh. *siploh*, PC *$pluh$, PMP *$daqan$ 'branch' > Malay *dahan*, Aceh. *dhʉən*, PC *$dhaan$. This kind of change is widespread in Mainland SEAsia, consistent with the gross areal shift to final stress and sesquisyllabicity, and is widespread in WMP (including spoken Malay, not withstanding Malay authography[2]).

2) PMP *n- > /l/. Two examples showing /l/ in Acehnese are adduced: PMP *h-in-ipi 'to dream' > Malay *mimpi*, Aceh. *lumpɔə*, PC *$lumpɛj$; PMP *$nipis$ 'thin' > Malay *nipis*, Aceh. *lipeh*, PC *$lipih$. Blust 2000 challenges both of these comparisons: in the first it is not clear that etymological *n- is the source of /l/, and the second is conceivably a borrowing into Acehnese. A counter example exists in the word for coconut: PMP *$niuR$ > Malay *nyor*, Aceh. *bɔh ʔu*, PC *$lə ʔu$, where Acehnese and Chamic share the same loss of final (and blocking of diphthongisation, see (4)), but Acehnese has lost the initial lateral, rather than shifting it to /n/ which parallels the change discussed next under 3).

3) In three Austronesian etyma imploded stops have developed in Chamic, having /ʔ/ reflexes in Acehnese, e.g.: PMP *$buhuk$ 'hair', Proto-Malayic *$buʔuk$, Aceh. *ʔoʔ*, PC *$ɓuk$; PMP *$nahik$ 'climb', Malay *naik*, Aceh. *ʔeʔ*, PC *$ɗiʔ$; PMP *$hideRaq$ 'lie down' > Aceh. *ʔeh*, PC *$ɗih$ (this last set lacks a Malay reflex, and was mistakenly identified as MK by Thurgood). These etyma are few in number and rather problematic. Firstly, there is a counter example to the regularity of the 'hair' etymology, namely the reflex of PMP *$bahu$ 'stench' reflected as Aceh. *bɛə*, without glottal initial, the vowel clearly indicating that it is not borrowed from Malay *bau*. Thurgood reconstructs 12 PC words with initial *$ɓ$, and 11 with initial *$ɗ$. All are

---

[2]Drawn to my attention by David Gill in 2001 while at the Max Planck Institute (Leipzig).

borrowings from MK into Chamic except for the one example of *ɗ discussed above. *Not one of these borrowings has an Acehnese reflex.* For the few relevant Austronesian etyma it is clear that a fusion of glottal and oral stops has occurred in Chamic. The Acehnese reflexes could have derived from the PC imploded stops by lenition, or independently by simple loss of initial syllables. The latter would be more phonetically plausible, but would not explain why this odd development has occurred in these same two etyma and not others with similar structure.

4)   Acehnese and Chamic share a loss of final *r, which blocked the diphthonisation of open syllable *u. E.g.: PMP *ikuR 'tail' > Malay *ĕkor*, Aceh. *ʔiku*, PC *ʔiku*; PMP *niuR 'coconut' > Malay *nyor*, Aceh. *bɔh ʔu*, PC *lə ʔu.

5)   PMC high vowels *i, *u diphthongised in open syllables, e.g.: PMP *beli 'buy' > Malay *bĕli*, Aceh. *blɔə*, PC *blɛj*; PMP *balu 'widowed' > Malay *balu*, Aceh. *balɛə*, PC *balɔw*. Thurgood reconstructs the Acehnese /ɔə, ɛə/ deriving from PC *ɛj, *ɔw (respectively) by dissimilation of vocalic onsets followed by neutralisation of final glides. Remarkably a parallel development occurred in Moken/Moklen—according to Larish (1999:396-398) PMP *i, *u diphthongised to *ɛj, *ɔw, then both merged to *əj. While this kind of change is otherwise rare or unknown in WMP languages, it is common in MK, Cf. Khmer *dəj 'hand' < *tii.[3]

6)   PMP *a, *e (ə) in stressed final syllables shifted to *aa, *a in Aceh-Chamic, with subsequent diphthongisation of *aa to /ɯə/ in Acehnese. E.g.: PMP *qudaŋ 'shrimp' > Malay *hudang/udang*, Aceh. *ʔudɯaŋ*, PC *hudaaŋ*; PMP *halem 'night' > Malay *malam*, Aceh. *malam*, PC *malam*. This chain shift of *a, *e also occurred variously in Malayic[4] and Moken/Moklen (Larish 1999).

The above six phonological developments found in the Austronesian lexicon of Acehnese and Chamic can be divided into two types:

(a)   developments 1), 5) and 6) are systematic, occurring in all lexicon where the phonological conditions are met, and in sufficient numbers of tokens to convincingly rule out wholesale borrowing into Acehnese. On the other hand, they are also not absolutely unique to Aceh-Chamic.

---

[3]Note that this example of diphthongisation in Khmer is not related any devoicing of the initial consonant and is unrelated to the Middle Khmer register split.

[4]I would like to thank Bill Foley for drawing this fact to my attention at the 2004 SEALS meeting in Bangkok.

(b) developments 2), 3), 4) are restricted to a very small number of tokens, and in light of the fact that 2) and 3) have counter examples demonstrating that they are not exceptionless sound-changes in Achenese, they may be explained by borrowing. Development 4) is the strongest example (if the two words in question are not borrowed into Acehnese), as it involves the blocking of a shared sound-change (5)).

In light of the above it is clear that Acehnese and Chamic do share some non-trivial developments in the phonology of their Austronesian lexicon. The systematic developments 1), 5), 6) and the complex development 4) are reconstructable to PC—so the question is to which period of Chamic history they belong. To help complete the picture we need to consider the analysis of borrowed (mostly MK) lexicon that is reconstructed for PC and also found in Acehnese.

2.2 Aceh-Chamic lexicon of MK and other origin

As noted above, Thurgood's PC lexicon includes 277 items borrowed from Mon-Khmer, 179 of unknown origin (plus 20 from Indic and three from Arabic).

It is striking that a very high proportion of the PC lexicon (and an even higher proportion in the case of modern Chamic tongues) consists of borrowing. By contrast relatively little of the Acehnese lexicon can be identified with these loans. This is pointed out by Dyen (2001), who states that only 44 items of Thurgood's PC lexicon of borrowed origin have an Acehnese cognate. Furthmore Dyen noted that a number of these sets include a Malay member, and in light of longstanding contact with Malay (ic) throughout Chamic history, this could indicate anything from a Pre-Proto Malayic Mon-Khmer borrowing to a post-PC diffusion. Thus Dyen's summation that "twenty-eight entries, perhaps better reduced to twenty-six, then appear to constitute the basis of the hypothesis that Acehnese is a Chamic immigrant" (p.393).

In respect of Dyen's count, it appears that only the PC lexicon that Thurgood identifies as being of MK origin was considered. We should also examine the strata characterised as of "uncertain" origin, in which case the count is 56 (allowing for some faulty comparisons). Even so, it is still apparent that this specific external contribution to Acehnese is in the order of a tenth or so of its contribution to Chamic. This strongly suggests that Acehnese and Chamic have been apart for a much longer time than they were ever together, although a late first Millennium CE separation remains a possibility if the great mass of MK borrowing into Chamic could be related to later history (such as the 13th century Khmer occupation of Champa). To deal with this problem we need to examine in detail Thurgood's 64 Aceh-Chamic comparisons of non-Austronesian origin, attempting a more thorough etymologisation and stratification. This data is presented in the Appendix to this paper, and informs the following discussion.

Of the 56 relevant etyma 30 are identified as certainly borrowed from MK, including at least six that were apparently borrowed separately. These include some etyma that show Acehnese innovations such as the diphthingisation of *aa > /ɯə/, the raising of *a > /ə/, and the loss of *-r. Such changes are indicative of earlier rather than later borrowing, although the analysis is complicated by the absolute low numbers of examples, which is in itself indicative of an early Aceh-Chamic separation.

The list also includes 18 items of unknown origin that show regular phonological correspondences between Acehnese and Chamic. As a subset of the much larger listing of Chamic words lacking any wider etymology these are very problematic. A couple of these include the *aa > /ɯə/ change in Acehnese. In the absence of any specific indication, our default hypothesis must be that these reflect common borrowings from an unknown language, identified as MK by Thurgood on structural grounds.

The remainder of the comparisons include two items of Indic origin (possibly borrowed via Malay), three comparisons that show phonological irregularities/inconsistencies that indicate separate borrowing, and three items borrowed from Moken/Moklen (MM). The Indic and irregular items have no sub-grouping consequences, but the MM isoglosses are perhaps significant.

Larish (1999) presents a reconstruction of Proto-Moken/Moklen (PMM), including a detailed discussion of its sub-grouping within Malayo-Polynesian, MM dialects are today spoken in the island chain that sweeps down the western coastline of the Isthmus of Kra. PMM innovations, such as PMP *q > *k, PMP *$ɟ/g$ > *j indicate that it represents a sub-group intermediate between PMP and PMC (Larish 1999:326-327) and thus an earlier Austronesian colonisation of Mainland SEAsia than the first Malayo-Chamic dispersal. Larish reconstructs the PMM homeland as the region of the Isthmus of Kra and the Malay Peninsula, with first arrivals on the East coast of the peninsula. He dates the internal divergence of MM to between 1300 and 1700 CE, correlating with the Thai occupation of the peninsula.

MM and Chamic share some important structural innovations, the most important of these being 1) the diphthongisation of high vowels in open syllables (unparalleled elsewhere in WMP), and 2) phonemic length across their vowel systems. In relation to 1), PMP *-i, *-u diphthongised and merged as *$əj$ in PMM, plus at least eight words in which PMP *-i yielded PMM *uj. If we accept Larish's suggestion that the *uj forms are old borrowings from Acehnese *$ɔj$ (before the shift to /ɔə/ in Aceh. occurred) we can account for both the PMM and Aceh-Chamic reflexes of PMP *-i, *-u by positing an initial common change of *-i, *-u to *$ɛj$, *$ɔw$, with various subsequent independent phonological changes. As for 2), phonemic vowel length first developed with the manner described under 2.1 (6) above, followed by accommodation of MK (and later Thai) borrowings. Larish argues for an ancient language area that he calls the Early Mainland Austronesian Complex (EMAC). However, I have serious reservations about the EMAC hypothesis, as the borrowed lexicon in MM is mostly different to that of Aceh.-Chamic, and I

rather suspect that the common phonological developments outlined above arose independently as a consequence of a shift to fixed word-final stress.

2.3 Post Chamic break-up borrowings into Acehnese?

Thurgood (p.54) lists nine items that he identifies as borrowed into Chamic languages and Acehnese from MK after the breakup of PC, a sub-set of the 173 post-PC borrowings he lists. This is a very problematic data-set—Thurgood's justification for this identification is the restriction of reflexes to Highlands Chamic, yet he admits some with Acehnese members, suggesting that "The presence of these forms in Acehnese shows that the Acehnese speakers left Champa not only after the breakup of PC but also after these words were borrowed from MK sources." Hence it is a crucial part of his argument for a late (10th century CE or later) separation of Acehnese.

I find this analysis unsatisfactory, as it does not stand up well to etymological scrutiny—only three of the words can be treated as potentially common Aceh-Chamic borrowings from MK, and they lack clear indications of when borrowing occurred.

Firstly, two of the words in question are arguably not borrowings from MK: post-PC *pruac* (?), Aceh. *pruət* 'stomach' Cf. Malayic *pərut*, Meningkabaw *paruj?*, post-PC *cum,* Aceh. *com* 'kiss, smell' Cf. Malay *chium* < PMP *cijum* (Bahnaric reflexes are phonologically marked as borrowings because the regular shift *c > /s/ did not occur).

Secondly several of the words are evidently borrowed independently: post-PC *?əmraak,* Aceh. *muıra?* 'peacock' Cf. Malay *měrak*—Monic forms have initial *m*, other MK groups indicate *braak* (except Aslian borrowings via Malay), and the variation in vowel length indicates multiple borrowings from Monic as length was lost in Mon; post-PC *dεl,* Aceh. *duıə* 'shallow'—MK forms such as Bahnar *daal* suggest PMK *-daal*, but Chamic must have borrowed from a language that had already raised *aa*; post-PC *dhual/r* 'dust; fog, mist', Aceh. *dhoj* 'ash, dust earth'—the semantics do not reconcile as Highland's Chamic forms mean 'fog, mist', Thurgood mistakenly compared to a Katuic form that mean 'to smoke/grill over fire'[5]; Post-PC *patuh,* Aceh. *buıruıtoh* 'explode'—Chamic compares to Khmer *phtùh*, Aceh. campares to Mon *batoh.*

Three are certainly MK borrowings into Chamic, either during Aceh-Chamic unity or later Chamic loans into Aceh: post-PC *cagəw,* Aceh. *cagεə* 'bear'—Bahnaric, Katuic and Vietic suggest PEMK *ɟkaw,* the best phonetic match is with Bahnar *cəgəw*; post-PC *kəmuan,* Aceh. *kuımuən* 'nephew'—most MK reflexes indicate /ɔɔ/ main-vowel, some sub-groups such as West Bahnaric and West Katuic have diphthonged vowels that developed secondarily; post-PC *grɔh,* Aceh. *dloh* 'to bark'—Katuic reflexes suggest *grɔɔh*, Cf. also Khmer *groh* 'ill mannered'.

---

[5]That error goes back to Peiros (1996), who read 'smoke' as a noun rather than a verb.

On balance it is apparent that while Acehnese and Chamic do subgroup, they share at most only a few dozens of MK borrowings, while Chamic without Acehnese went on to acquire hundreds of MK loans. This suggests a more modest period of Aceh-Chamic unity than posited by Thurgood, followed by a long period of largely independent contact with various MK languages. Also there may well have been some new Chamic influence on Acehnese after the attested post-1471 migrations, which may also explain some of Thurgood's comparisons.

3.   The sources of MK loans into early Chamic

A real mystery surrounds the question of from which languages did early Chamic (particularly after the split from Achenese) borrow so heavily. Thurgood adduces many Bahnaric and Katuic comparisons to support his historical analysis, relying on these two families for his MK comparisons. In 2001 I began to examine Thurgood's data in detail and discussed my (then) conclusions at the May 2002 SEAsian Linguistic Society (SEALS) meeting at DeKalb, reporting that:

- for more than 200 etyma Bahnaric-Chamic comparisons are readily identified, and there is generally close phonological agreement between Thurgood's PC and Bahnaric forms.
- There is no evidence of significant Chamic vocabulary in West Bahnaric, suggesting that Bahnaric had already split into distinct branches once contact with Chamic began.
- Of Thurgood's 63 comparisons which include Katuic members, in only 6 do the Katuic forms show better agreement with Chamic than does Bahnaric.
- In some cases Vietic forms show better phonological agreements with Proto-Chamic, compared to Katuic or Bahnaric. This may indicate some Vietic contact, or simply be indicative of their wider MK etymology.

At that time I assumed that contact with Bahnaric was the most likely source of MK vocabulary in Chamic, and argued that the theory of close Katuic-Chamic contact (including the suggestion of morphological borrowings) put forward by Thurgood lacked empirical support. Subsequent special attention to the Katuic lexicon reveals several isoglosses that suggest at least some very limited borrowing from Katuic into Mainland Chamic (such as *ʔəlɛɛ* 'bamboo', *hual* 'steam, vapour'), as well as several Chamic words uniquely borrowed into Katuic (e.g. *ŋin* 'wind', *kəw* 'I'?).

However, my efforts generally to etymologise the Chamic lexicon have revealed a much broader picture, and it is now evident that Khmer and even Mon are likely sources of many MK loans into Chamic, e.g.:

PC *madəh 'awaken'   Khmer dah 'awaken'
PC *padaw 'warm, hot'   Khmer kdaw 'hot'
PC *ɲam 'vegetables, food,   Khmer ɲam 'eat'
        legumes'
PC *sənaa 'crossbow'   Khmer snaa 'gun' (other MK have
                                                final glottal)
PC *cih 'write, draw'   Khmer cih 'learned'

PC *ʔɔʔ 'vomit'   Mon haʔɔʔ 'vomit'
PC *klah 'lose (v.)'   Mon klah 'cleared away'
PC *klɛʔ 'steal'   Mon klɛʔ 'lose, disappear'
PC *pɔk 'open (v.)'   Mon/Nyahkur pɔk 'open'

But many PC forms with MK parallels cannot be identified with a specific sub-group, as they are widely attested forms and/or composed of segments that have undergone few if any changes, such as:

PC *ʔaak 'crow'
PC *ʔadaa 'duck'
PC *ʔahaa, *haa 'open (mouth to say something)'
PC *ʔaw 'clothing, shirt'
PC *klan 'python'
PC *prɔɔk 'squirrel'

Of the roughly 200 Bahnaric comparisons in Thurgood's PC lexicon about half are restricted to Chamic and Bahnaric languages in immediate contact with Chamic, specifically Bahnar, South Bahanric and North Bahnaric. They are almost entirely missing from West Bahnaric and the Central Bahnaric languages spoken in Laos: Kasseng/Taliang and Alak. So far I have counted only three West Bahnaric-Chamic comparisons that are not also shared with Khmer, listed below (see also Sidwell and Jacq 2003):

Proto Chamic *tapaj 'rice wine, alcohol', *tapɛj 'yeast'; Malay tapai
Cf. West Bahnaric *tʔpɛɛ 'alcohol', South Bahnric *dpee 'yeast'

Proto Chamic *ʔantɔw 'ghost, corpse'; Malay hantu < PMP *qanitu
Cf. West Bahnaric *ʔətəw 'grave', Bahnar *ʔətəw 'ghost, corpse'.

Proto Chamic *raa 'person'; Thurgood suggests this is reduced form
        of Malayic *uraŋ
Cf. West Bahnaric *raa 'person classifier'.

The first two of these have significant cultural associations and they may have diffused into West Bahnaric without direct Chamic contact. The last item is of ambiguous significance—Thurgood suggested deriving it from Malayic *ʔuraŋ, but the radical phonetic simplification is not explained, so it remains an isolated curiosity.

I take this distribution of borrowings as a clear indication that Bahnaric-Chamic contact occurred after the separation of West Bahnaric and after the separation of Kasseng/Taliang and Alak from Central Bahnaric. The latter of these splits probably occurred sometime during the first Millenium, and judging by the Khmer and Katuic loans in Kasseng/Taliang and Alak I estimate that it was before Angkor, perhaps during the Chenla period or somewhat earlier. Then consider the following: if the Chamic-Bahnaric isoglosses are borrowed from Bahnaric, then by chance alone there should be a significant number of them in West Bahnaric, yet this is not the case. Furthermore, most of the isoglosses are shared with North Bahnaric, which must have split from Central Bahnaric before the separation of Kasseng/Taliang and Alak. The inescapable conclusion is that the direction of borrowing between Proto Chamic and Bahnaric was principally *into Bahnaric*, in a language area that included Bahnar, North Bahnaric and South Bahnaric. Logically then this lexicon must have been acquired by Chamic from another language or languages, before Bahnaric came under strong Chamic influence.

This unknown language(s), typologically resembling MK at least in terms of word structure, made a substantial lexical contribution to Chamic, and via Chamic the various Bahnaric languages spoken by communities that fell under Chamic control. Significantly we don't find loans from this mystery language(s) in Khmer, even though Angkor covered a vast territory at its peak, nor in Mon, as Dvâravatî extended from the Isthmus of Kra to the site of modern Vientiane. My suggestion is that this contact occurred, and the community of speakers was already gone, absorbed into Champa, before the rise of Angkor or Dvâravatî. These people must have been located on the lands east of the Annamite range, and not in the Mekong valley. And since the Acehnese lexicon shows only modest evidence of influence by the mystery language, we again have indication that Acehnese separated from Chamic at an early, rather than later, date. It is also possible that even early Chamic was spread geographically, and that the dialect ancestral to Acehnese was located away from the mystery substrate.

Returning to the question of which historical events may explain the Aceh-Chamic split. Although Thurgood places great weight on the history of Vietnamese-Chamic conflict from 982 the previous centuries were not always stable nor peaceful for Champa. From my reading of Majumdar's (1985) account of Cham history I am struck by the repeated attempts of China to subjugate Champa, and in particular the dramatic events of the fifth century. According to Majumdar (1985:29-31) in response to Chamic raids on Chinese administered territory, "In the year 420 the Chinese inflicted a crushing defeat upon the Chams", and again, "in 446 AD the Chinese army under T'an Ho-Tche invaded Champa. Terror now seized Yang Mai II and he proposed a humiliating peace to the Chinese emperor". The settlement reportedly included the loss of 100,000 pounds of gold to China. Events such as these are arguably just as likely to stimulate an exodus of refugees as the later conflicts with Vietnam. Furthermore we know that the fifth century battles often involved the use of great armadas of ships, so the capacity to transfer large populations great distances to safety were well developed. An emigration of people seeking to escape fifth century Chinese attacks is no less hypothetical than Thurgood's

theory, but may better explain why the Acehnese language acquired far fewer loans than the rest of Chamic.

Integrating the above insights we can to propose the following tentative timeline:

500~300 BC	Separation of Malayo-Chamic—mainland settlement and begining of contact with mainland languages.
100-200 CE	Establishment of Champa kingdom.
100-1000 CE	Mainland Chamic linguistic unity—contact with unidentified as well as Bahnaric and Katuic languages.
400s CE ?	Separation of Acehnese from Chamic.
1000 CE onwards	Beginning break-up of Mainland Chamic languages - Tsat > Hainan etc. Conflicts with Vietnam, Angkor etc. lead to further MK borrowings.
1400s onwards	Last and greatest phase of Chamic dispersal/diaspora.

The substantial problem remains of explaining the borrowings of unknown origin, and the successful resolution of this issue could shed important new light on the first Millennium history of Aceh and Champa.

ABBREVIATIONS

An	Austronesian
CE	Common Era
Demp.	Dempwolff (1938)
MC	Malyo-Chamic
MK	Mon-Khmer
MM	Moken/Moklen
PC	Proto-Chamic
PHF	Proto-Hesperonesian-Formosian
PMK	Proto-Mon-Khmer
PMM	Proto- Moken/Moklen
PS	Paul Sidwell
WMP	Western-Malayo-Polynesian

REFERENCES

Blust, Robert A. 1994. "The Austronesian settlement of mainland southest Asia." In *Papers from the Second Annual Meeting of the Southeast Asian Linguistics Society 1992*. Karen Adams and Thomas Hudak, eds., Tempe Arizona, Arizona State University. pp. 25-83.

Blust, Robert A. 2000. "Review of Thurgood 1999." *OL* 39:435-445.

Burnham, Eugene C. 1976. "The place of Haroi in the Chamic languages." M.A. thesis, University of Texas at Arlington.

Collins, William. 1996. "The Chams of Cambodia." In *Interdisciplinary Research on Ethnic Groups in Cambodia: Report on Research to United Nations Development Programme.* pp. 15-107, Center for Advanced Study, Phnom Penh.

Dempwolff, Otto. 1938. *Austronesisches Wörterverzeichnis*. Berlin, Dietrich Reimer.

Dyen, Isidore. 2001. "Review of Thurgood 1999." *AL* 43(3):390-394.

Larish, Michael. 1999. "The position of Moken and Moklen within the Austronesian language family." Ph.D. dissertation, University of Hawaii.

Lee, E.W. 1966. "Proto-Chamic phonologic word and vocabulary." Unpublished Ph.D. dissertation, Indiana University.

Majumdar, R. C. 1985 (preprint of 1927 edition). *Champa: History and Culture of an Indian Colonial Kingdom in the Far East 2nd-16th Century A.D.* New Delhi, Gyan Publishing House.

Shorto, Harry. ms (no date). *Mon-Khmer Etymological Dictionary*.

Sidwell, Paul. 2002. "Classification of the Bahnaric languages: A comprehensive review." *MKS* 32:1-24.

Sidwell, Paul and Pascale Jacq. 2003. *A Handbook of Comparative Bahnaric: Volume 1, West Bahanric*. Canberra, Pacific Linguistics 501.

Thurgood, Graham. 1999. *From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change.* Oceanic Linguistics special Publications No. 28. Honolulu, University of Hawaii Press.

Zorc, David. 1995. "A glossary of Austronesian reconstructions." In Darrell Tryen I ed., *Comparative Austronesian Dictionary*, part 1, fascicle 2, pp. 905-1197. Berlin and New York, Mouton de Gruyter.

Received: 26 January 2005   Department of Linguistics
                                      Research School of Pacific & Asian Studies
                                       Australian National University, 0200, ACT,
                                                        AUSTRALIA
                                             <paul.sidwell@anu.edu.au>

Appendix

Thurgood's Acehnese-PC lexicon of MK or uncertain origin with etymological commentary.

Note: Proto-Mon-Khmer (PMK) reconstructions are Harry Shorto's (Shorto ms.) unless marked as my own (PS). Austronesian reconstructions are from Zorc (1995) unless mared otherwise.

1)   Words with PMP etymology:

1a) Words inherited via Proto-Malayo-Chamic

1.1)   Aceh. *bantaj* 'pillow', PC **bantal*
◊   PMP **bandal*, Malay *bantal*. Also borrowed into Aslian from Malay.

1.2)   Aceh. *cabɯəŋ* 'branch', PC **caɓaaŋ*
◊   PMP **cabaŋ*, Malay *chabang*. Chamic forms have unexpected imploded labial, yet there is no evidence of origin in MK. Has been borrowed into Aslian from Malay.

1.3)   Aceh. *picah* 'broken; break', PC **picah*
◊ PAn **pĕcéq*, Malay *pĕcah*. Borrowed into Katuic, Bahnar and South Bahnaric. In the latter two the presence of a palatal stop, rather than a fricative clearly indicates inward borrowing.

1.4)   Aceh. *truəŋ* 'eggplant', PC **trɔŋ*
◊   PHN **terúŋ*, Malay *tĕroŋ*. Borrowed into Bahnaric from Chamic. There is a look-alike in Proto-Monic **trʔɔɔŋ* > Mon *hdɔɔŋ* (Diffloth 1984:88-89, who speculates that Malayo-Polynesian is the source in Monic).

1.5)   Aceh. *krɯəŋ* 'shellfish', PC **kraaŋ*
◊   PHF **kaRaŋ* 'landcrab', Malay *kĕrang.*

1.6)   Aceh. *mat* 'hold', PC **mat* 'take; fetch, get'
◊   PAn (Demp.) **amet* 'take with hand', Malay reflex not apparent.

1.7)   Aceh. *pʌt* 'pick, pluck', PC **pɛt*
◊   PAn **pit* 'press', Malay *apit, pepet,* Javanese *pipit.*

1.8)   Aceh. *pula* 'to plant', PC **pula*
◊   PMP **múla*, Malay reflex not apparent.

1b) Isoglosses with Moken/Moklen:

1b.1) Aceh. *paʔɛɛ* 'gecko', PC **pak-kee*
◊   Cf. proto-Moken/Moklen **tɔkɛɛʔ*, Cf. Malay *tĕkek* < PMP **tekik*

1b.2)   Aceh. *ʔiəʔ* 'urinate', PC **maʔiãk.*
◊   Pre-Moklen **niʔiək* < PMM **niʔ̪aak* < PMP **[ ]iSeq*

1b.3)   Aceh. *lhon* 'naked', PC **[ma][sa]lun*
◊   Cf. proto-Moken/Moklen **ɲuɪlɔn.* No wider etymology apparent.

1c) Words of Indic origin, possibly borrowed via Malay:

1c.1)   Aceh. *gapɯəh* 'cotton', PC **kapaas*
◊   Cf. Sanskrit *karpāsa*, PHN **kapes*, Malay *kapas*. Also widespread in
        MK, Shorto reconstructing PMK **kpaas.* It seems that this word
        diffused from India quite early.

1c.2)   Aceh. *piɲan* 'plate', PC **piɲan* 'bowl, dish'
◊   Ultimately from Persian—Dempwolf reconstructed PMP **paŋgan*, Cf.
        Malay *pinggan* 'plate'.

2) Words of MK origin:

2a) MK words possibly borrowed into Proto-Aceh-Chamic

2a.1)   Aceh. *ʔaʔaʔ* 'crow', PC **ʔaak*
◊   Imitative form widespread in MK, suggesting PMK **kʔaak, *kʔaik*
        (Shorto). Note that Acehnese eschews the regular /ɯə/ reflex of **aa* in
        favour of /a/.

2a.2)   Aceh. *ʔawɯəʔ* 'spoon, ladle', PC **ʔawaak*
◊   There are various WMP reflexes, e.g.: Cebuano Bisayan *luwag* 'ladle'.
        Reflexes in Aslian and EMK suggest PMK **[ ]waik*, Shorto remarks
        "Vocalism against IN > MK".

2a.3)   Aceh. *ʔɛʔ* 'defecate', PC **ʔɛh*
◊   Katuic and Vietic **ʔeh* 'excrement', Khmer *cəʔih* 'stink of urine'
        suggest PEMK **ʔeh* (PS).

2a.4)   Aceh. *cicaʔ* 'gecko', PC **cicaʔ*
◊   Cf. Malay *chichak* < PAn **kʼəkʼak* (Demp.); Old Khmer *cik cak*, Mon
        hɔcɛk suggest PMK **cɔcak, *cɔcɔk.* Imitative word - problematic
        for the present purposes.

2a.5)   Aceh. *cʌʔ* 'great-grandchild', PC **cicẽʔ*
◊   Cf. Malay *chichit* 'id.'; Khmer, Bahnaric, Katuic, Khmuc and
        Palaungic forms suggest PMK **ceʔ* 'great-grandchild'. However, only
        some Palaungic forms preserve both the initial palatal stop and final
        glottal stop intact, yet Northern MK is surely an unlikely source for
        Chamic?

2a.7)   Aceh. *cɔh* 'peck (bird); strike (snake)', PC **cɔh*
◊   Khmer, Katuic, Bahnaric, Khmu and Aslian indicate PMK **[ʔ]cɔh.*

2a.8)   Aceh. *dʌŋ* 'stand; stop', PC **dəŋ*
◊   Widespread MK reflexes indicate **ɟə[ə]ŋ* 'stand', with Austronesian parallels, e.g. PAn **d'əŋ* (Demp.) in the light of reflexes such as Javanese *jĕng*, which themselves resemble MK forms (e.g. Old Khmer *ɟeŋ*). The best Chamic-MK match I have found is Vietnamese *đứng* 'stand', which suggests Proto-Vietic **təŋ*. Could it be a second Millennium diffusion from Vietnamese?

2a.9)   Aceh. *gulam* 'carry on shoulder', PC **gulam*
◊   Cf. NMK, Aslian suggest PMK **klam* or **kləm* 'carry on shoulder'.

2a.10) Aceh. *klɔ* 'dumb', PC **k-am-lɔ*
◊   Reflexes with and without *-m-* infix occur in Bahnaric and Katuic. Cf Khmer *kamlaw* 'ignoramus'

2a.11) Aceh. *kɯaŋ* 'chin, jaw', PC **kaaŋ*
◊   Bahnaric, Katuic and Vietic suggest PMK **kaaŋʔ*, also widespread reflexes in Tai languages such as Thai *khaaŋ* 'chin'.

2a.12) Aceh. *klɯaŋ* 'kite', PC **kalaaŋ*
◊   Widespread MK reflexes suggest PMK **kəlaaŋ*, with the first syllable possibly a prefix. Has been borrowed into WMP languages, e.g. Malay *hĕlang*.

2a.13) Aceh. *ʔulɯə lhan* 'boa', PC **klan*
◊   MK reflexes suggest PMK **tlan*, with the initial cluster shifting to /kl/ in Mon, Bahnaric and Northern MK.

2a.14) Aceh. *kɔh* 'cut off', PC **kɔh*
◊   Bahnaric, Katuic, Northern MK, Aslian and Nicobarese suggest PMK **kɔh.*

2a.15) Aceh. *kruaŋ* 'river', PC **krɔɔŋ*
◊   Bahnaric, Northern MK, and Vietic suggest PMK **krɔɔŋʔ* (PS) with widespread reflexes in Tai languages such as Lao *khɔɔŋ*; Katuic and Mon suggest PMK **-ruuŋ.*

2a.16) Aceh. *miəŋ* 'cheek', PC **miaŋ*
◊   Cf. Vietnamese *miệng* 'mouth', yet Muong *mẹnh* and other MK forms suggest PMK **miəɲ* (Shorto). A second Millennium diffusion from Vietnamese?

2a.17) Aceh. *pɔh* 'beat, hit, thrash', PC **pɔh* 'strike, pound'
◊   Cf. Khmer, Katuic and Northern MK suggest PMK **pɔh* (PS) or **puəh* (Shorto)*.*

2a.18) Aceh. *ruɯtɯəʔ* 'bean, pea', PC **rətaak*
◊   Old Khmer *santek,* Mod Khmer *sandaek,* Pearic *rətaak,* Alak *htaak.* Also borrowed into WMP, e.g. Moklen *suɯtak, lətak,* Iban *retak.* Diffused from Khmer?

2a.20) Aceh. *sapaj* 'upper arm', PC **sapal* 'arm'
◊   Katuic, Aslian suggest **sapaal* (PS), but the lack of wider etymology makes this difficult to interpret.

2a.21) Aceh. *sɔh* 'empty', PC **sɔh*
◊   Khmer, Katuic, Bahnaric suggest PMK **sɔh* (PS) or **suəh* (Shorto).

2a.22) Aceh. *tuwʌ* 'forget', PC **wər*
◊   The wordshape *wVr* is widespread in MK associated with the notion 'turn, spin', with derived meanings 'dizzy, confused, forget', e.g. Khmu *wiir* 'forget'.

2a.23) Aceh. *sideh, hideh* 'there (far)', PC **dih* 'that, there'
◊   Bahnaric forms suggest PB **dih* (PS) 'outside; below; near'.

2a.24) Aceh. *luɯmɔ* 'cow', PC **ləmɔ*
◊   Northern MK suggests PMK **ln[b][o]ʔ* (Shorto), although Katuic, Bahnaric may have borrowed from Chamic. Also borrowed into other WMP e.g. Malay *lĕmbu.*

2a.25) Aceh. *suɯəŋ* 'hut, tent', PC **saaŋ* 'house'
◊   Cf. Khmer *saaŋ* 'to build', borrowed into Thai/Lao variously as 'roof, granary, warehouse'

2a.26) Aceh. *tijɯəp* 'chase', PC **tijaap*
◊   Tampuon *tijạap,* Khmu *ŋgjaap*: although Tampuon may have borrowed from Chamic, the Khmu suggests a deeper MK origin.

2b) Words of MK origin borrowed separately by Aceh and Chamic:

2b.1)   Aceh. *cicem* 'bird', PC **cim*
◊   Reflected in all branches of MK except Khmer—the vocalism is problematic although EMK languages typically indicate **cim* or **ceem.* Interestingly Moken/Moklen **cicum* shares with Aceh. the reduplicated initial, paralleled only in Nicobarese *cecoon.* It is clearly evident that Chamic and Acehnese borrowed from different sources.

2b.2)   Aceh. *suɯmɯŋɯɯp* 'yawn' PC **həʔaap*
◊   reflexes throughout MK (absent only from Aslian) suggest PMK **sʔaap* or **snʔaap* (Shorto). The Acehnese vocalism and presyllables are difficult to relate to the Chamic form.

2b.3) Aceh. *kʌŋ* 'strong', PC **khaŋ* 'hard, stiff, strong'
◊   Khmer, Mon suggest PMK **kr[ɔ/a]ŋ,* Katuic suggests **kəŋ,* Chamic and Bahnaric forms with initial aspirate may reflect Vietnamese **khăŋ.* Either way Aceh. and Chamic have borrowed this separately. PMM **kalaŋ* may also be an old MK borrowing.

2b.4) Aceh. *phʌ* 'to fly', PC **pər*
◊   PC regularly reflects an early borrowing of PMK **par.* However the aspirated initial of Aceh. cannot not be explained by Chamic phonology, and therefore have probably borrowed this separately (speculation: Nicobarese shifted **p-* > /f/, presumably via **ph-*).

2b.5) Aceh. *pluəʔ* 'to peel', PC **lɔɔk*
◊   Bahnar *lɔɔk,* South Bahnaric **plook,* Katuic **lok,* Khmer *lɔɔk* 'scratch, score into'.

2b.6) Aceh. *rhah* 'wash', PC **raw*
◊   Bahnaric, Katuic, Vietic suggest PMK **ʔəraaw,* while Khmuic, Vietic suggest **raa[h],* Therefore Aceh. and Chamic forms are different borrowings.

3)   Aceh-Chamic isoglosses lacking wider etymology:

3.1)   Aceh. *bɛʔ* 'don't', PC **bɛʔ*
◊   Apparently restricted in MK to Bahnar and North Bahnaric, so it is arguably borrowed from Chamic, its ultimate source obscure.

3.2)   Aceh. *cʌt* 'hill', PC *cət 'mountain'

3.3)   Aceh. *lhʌh* 'dismantle', PC **gləh* 'descend; sink; knock down; collapse, destroy' Speculation: Vietnamese *giở,* *trở* 'untie, unwrap, to change' potentially reflect a Vietic **gləh,* but more data is required to confirm.

3.4)   Aceh. *gɔm* 'to cover', PC **gəm*

3.5)   Aceh. *gɔp* 'stranger, others', PC **gəp*

3.6)   Aceh. *gʌ* 'handle', PC **gər*
◊   Borrowed into Bahnar and South Bahnaric.

3.7)   Aceh. *ʔurɛt* 'rope, vine' PC **hurɛt*

3.8)   Aceh. *khuəŋ* 'drought', PC **khɔɔŋ*

3.9)   Aceh. *hlɯah* 'escape', PC **klaas*
◊   Bahnar and Katuic suggest **klah,* but the limited distribution implies borrowing into MK from one of the Chamic languages that merged final **-s* and **-h* > /h/.

3.10)   Aceh. *ləə* 'much, many', PC **luu*
   ◊   Borrowed into Bahnar as *ləə*. Note: the *əə—uu* vowel correspondence is problematic.

3.11)   Aceh. *sɔm* 'hide, put away', PC **səm* 'to wrap'
   ◊   Speculation: PMK **kləm* potentially yields [*səm*] in Vietnamese, but I have not found it.

3.12)   Aceh. *takuə* 'neck', PC **takuaj*

3.13)   Aceh. *tamɛh* 'pillar, post', PC **tamɛh*

3.14)   Aceh. *troh* 'arrive', PC **truh*
   ◊   Borrowed into Bahnaric from Chamic.

3.15)   Aceh. *ʔubo* 'snail', PC **ʔabaw*
   ◊   Borrowed into Bahnaric from Chamic.

3.16)   Aceh. *ŋuj* 'to use', PC **ʔaŋuj*

3.17)   Aceh. *liəh* 'lick, taste', PC **lijah*
   ◊   Comparison with Khasi *jliah* 'lick' is suggestive, but geographically unlikely. Perhaps the PC form is an irregular reflex of PHF **d₃ílaq* 'lick' > Malay *lidah* by metathesis.

3.18)   Aceh. *tho* 'dry', PC **thu*

4)   Possibly related forms with unexplained discrepancies:

4.1)   Aceh. *ɟɯmpuŋ* 'rice straw', PC **puuŋ*
   ◊   Aceh. short vowel is irregular. Cham reflexes suggest **piaŋ* No wider etymology apparent.

4.2)   Aceh. *siə* 'flesh, meat', PC **ʔusar*
   ◊   The PC form has no apparent etymology; however Aceh. form may be related to Malayic **isiʔ*.

4.3)   Aceh. *tuwan* 'parent-in-law', PC **tamaha*
   ◊   PC continues PAn **tuqás* 'wife's father, mother's brother' (no Malay reflex apparent), Aceh. form unexplained.