6 *Acehnese and the Aceh-Chamic Language Family*[1]

Paul Sidwell

1. Introduction

The starting point for this paper is the treatment of Acehnese as a Chamic language by Thurgood (1999) (henceforth 'Thurgood'). While many scholars (e.g. Niemann 1891, Cowan 1933, 1948, 1974, 1981, Shorto 1975, 1977, Collins 1969, 1975, Blust 1981, Durie 1990 and others) have noted that, although widely separated geographically (Aceh in northern Sumatra and Champa centred in Vietnam), Acehnese and Chamic form a genetic sub-grouping. Thurgood is explcit in treating Acehnese as a descendent of Proto-Chamic (PC), specifically as the first dialect to separate from a more or less united Chamic speech community, sometime late in the 1st millennium CE. However, scholarly views on the precise nature of the Aceh-Chamic relationship vary, with no clear consensus on the likely date of separation of the Aceh-Chamic speech community.

Thurgood's monograph length study has revealed the extent to which Chamic was relexified by borrowings, particularly from Mon-Khmer, from ancient through to modern times. Earlier studies, such as Headley (1976), had suggested that around 10% of the reconstructable Proto-Chamic vocabulary was borrowed from Mon-Khmer (MK), while Thurgood's work indicates that the real proportion is perhaps more than three times that, with around 40% of the Proto-Chamic basic lexicon replaced by borrowings of one source or another. Yet for many of these borrowings it is difficult to clearly identify a specific source, not withstanding their frequent co-occurrence in neighbouring Bahnaric languages. My comparative and distributional analyses indicate that the mass of lexicon shared between Chamic and Bahnaric (and to some extent Katuic), is almost entirely borrowed from Chamic into Bahnaric, which implies that they formed a language area at a somewhat later phase, rather than from the outset of Chamic settlement.

My hypothesis, presented in this paper, is that Chamic and to a lessor extent Acehnese, preserves a "substratumised" branch of Mon-Khmer[2] that is otherwise unattested and now extinct—presumably the result of a language shift. The substantial body of borrowed lexicon reconstructable to Proto-Chamic (according to Thurgood) is very difficult to etymologise, and it is clear that there is a very old stratum that has no source in any known languages. A much smaller proportion of this stratum is shared with

[2] Please forgive the echos of the late Paul Benedit's (1976) imaginative hypothesis for explaining certain lexical aspects of his "Austro-Thai" hypothesis.

Acehnese, so logically the separation of Aceh-Chamic occurred sometime during the substratumisation process. The pre-Acehnese must have moved away from the zone of language contact, contra Dyan's (2001) idea that Aceh-Chamic orignated in Sumatra with the Proto-Chams moving on to Indo-China. Clearly Aceh-Chamic originated with initial settlement on the Indo-Chinese coastline, followed by the splintering off of the Acehnese.

Well after the separation of Acehnese there were other phases of significant MK influence upon Chamic, principally by Khmer, Mon and Vietnamese. Probably much of it was associated with historical events that led to the decline of Champa and the differentiation of Chamic into Coastal and Highland branches. The earliest and later contact phases must have been quite separate, as we find no identifiable traces of the oldest loan stratum elsewhere beyond mainland Chamic and the Mon-Khmer languages of the Annamite Range that came under strong Chamic influence.

We may speculate that some great historical event, perhaps a great political conquest, saw a foreign population absorbed completely into the nascent Champa, leaving no direct ancestor elsewhere in Indo-China. Alternatively the substratum may simply have been the language of the autochrones of the Indo-Chinese coastal plains that were first encountered, and then absorbed, by pre-Aceh-Chamic settlers. My favoured speculation is that we might connect the more obscure lexical stratum in Chamic with the mysterious kingdom of Funan, an ally of early Champa that was ovetaken by the pre-Angkorian Khmer Chenla (Zhenla) around the middle of the first millennium. I dare not pretend to have positively identified the "language of Funan"—presumably the name refers only to the political centre that ruled over an ethnically complex region—but one can claim at least to have identified a specific line of investigation.

Finally, from a programmatic perspective, I suggest that it is appropriate to build upon the solid foundation of Thurgood's data and analyses by drawing in more extensive sources, especially Mon-Khmer, to rework the reconstruction of the respective phonologies and lexicons of Proto-Aceh-Chamic and Proto-Chamic. A more extensive etymological compilation and stratification of the lexicon offers prospects for revealing the history underlying the remarkable contact-driven change which occurred in the Aceh-Chamic languages. It is also significant that, if as I suggest, the Acehnese have constituted an independent society for the better part of 2000 years, there will be historical implications for migration and settlement that other disciplines may be able to shed some light upon.

2. Malayo-Chamic

Thurgood approvingly cites Blust (1994) identifying a Malayo-Chamic (MC) subgrouping within Proto-Malayo-Polynesian (PMP), which split into Malayic and Chamic branches (see Fig. 1, below) sometime in the first Millenium BCE.
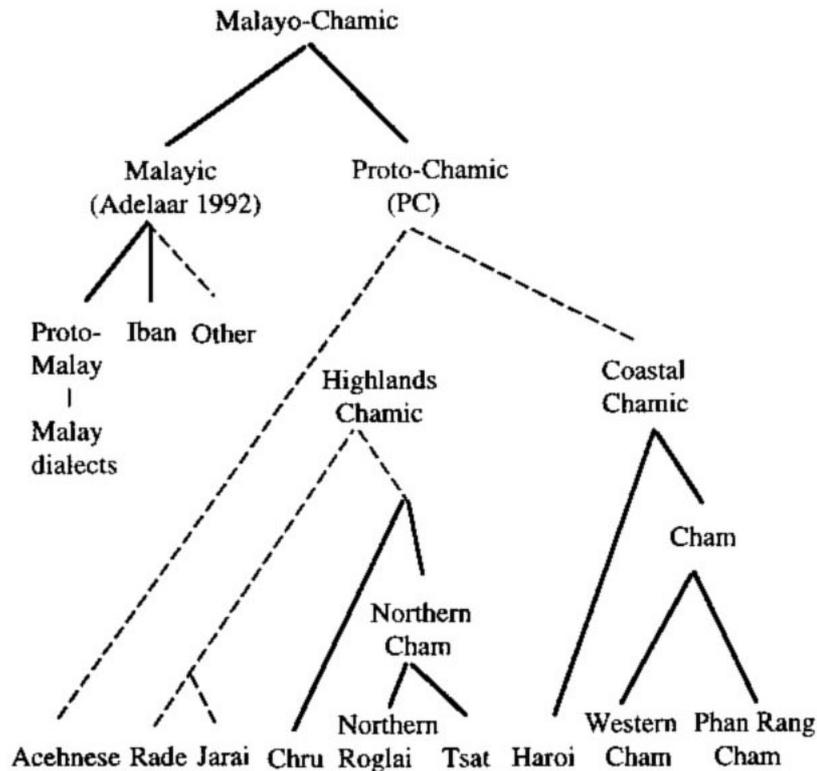
Figure 1: *Thurgood's Figure 6: the Malayo-Chamic Languages (p.36)*

Three principal sound changes that mark the formation of Proto-Malayo-Chamic (PMC) are discussed: 1) PMP **R* > PMC **r*, 2) PMP **w-* > PMC **Ø-,* 3) PMP **q* > PMC **h*:

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

2) PMP **q* > PMC **h*, e.g.:
PMP **qataj* 'liver', Malay *hati*, Aceh. *ʔate*, PC **hataj*
PMP **daqih* 'forehead', Malay *dahi*, Aceh. *dhɔə*, PC **ʔadhẽj*
PMP **baseq* 'wet', Malay *basah*, Aceh. *basah*, PC **basah*

3) PMP **w-* > PMC **Ø-*, e.g.:
PMP **waRiH* 'sun/day', Malay *hari*, Aceh. *ʔurɔə*, PC **hurɛj.*
PMP **wakaR* 'root', Malay *akar*, Aceh. *ʔukhɯə*, PC **ʔughaar*
PMP **wahiR* 'water', Malay *air, ayer*, Aceh. *ʔiə*, PC **ʔiar*

In the case of word initial PMP **q* the Acehense reflex is /ʔ/ which requires a sequence **q* > **h* > **ʔ*. This initial glottal stop is not usually written in transcription, as it is predictable, a phonotactic artifact. This is also the occasional reflex in Malay, e.g. *abu* 'ashes' < PMP **qabu*.

The loss of initial *w- is interesting as there appears to be a trace of it in the labial quality in the Aceh-Chamic minor-syllable[3] vowel, which shifted to /u/. At this point I caution the reader that I am approaching the topic of Austronesian historical phonology as an outsider, but it seems logical to me that the syllable *wa- must have been present at the PMC level, since a simple *a would not have unconditionally shifted to [u] in Aceh-Chamic, any more than a secondary *u would have unconditionally shifted back to [a] in Malayic. In the case of PMP *wahiR 'water' an earlier regular loss of *h resulted in a change of syllable structure that eliminated the minor-syllable, creating a diphthong, so there was no eligible vowel to labialise (note that Aceh-Chamic metathesised the resultant diphthong). Strikingly the 'sun/day' etymon shows special evidence of connection with Malayic—sharing the otherwise uniquely Malayic addition of an initial [h]. If it was a simple loan from Malay(ic) we would not expect the [u] vowel, so we are left to suggest some kind of contamination was caused by a knowledge of Malay(ic) among Aceh-Chamic speakers.

The above changes are not uniquely restricted to MC among MP: *q > [h] also occurred in Balinese, Javanese, Sundanese and Batak, and the merger of *R and *r and the loss of *w also occurred in Batak and Balinese. In these circumstances Blust's phonological arguments for MC also suggest that parallel changes elsewhere in MP were independent, and we may wonder why their occurrence in Malayic and Aceh-Chamic is not similarly coincidental, particularly in the light of the necessarily independent development of Aceh-Chamic *ʔu- < *wa-.

To the phonological data we can add the innovations among the numerals. Thurgood (p36-39) provides a detailed discussion of these, showing how Malayic and Aceh-Chamic replaced the PMP forms for 'seven', 'eight' and 'nine' with new words, the latter two based upon subtractive formulations. Thurgood concedes that the innovated 'eight' and 'nine' forms also occur in Maloh and Rejang, although Blust (1992) cautions that this "may be due to borrowing". One may also wonder whether the ancient Aceh-Chamic also acquired the new numeral forms by borrowing from Malayic.

My brief review of the Malayo-Chamic hypothesis leaves me with the strong impression that it does not demonstrate a very neat process of separation and branching such as we might like to see in a phylogentic model—instead it suggests a much messier (yet perhaps more realistic) dialect chain that saw prolonged contact and mutual influences, as sub-groupings emerged and population movements occurred. This is quite a normal thing in the real world, but we are still at a loss to understand the specific historical consequences this may have had for the place of Aceh-Chamic vis-à-vis Malayic, and the version of Malayo-Chamic I am relying upon in this paper. For now I do not wish to argue for any particular alternative to Blust's MC, as I am concerned with the Aceh-Chamic hypothesis in particular, but it is clear that the issue deserves further examination.

---

[3] The term 'minor-syllable' is used by Mon-Khmerists to designate the initial syllable within the typically MK phonological word pattern that maximally permits only iambic structures, with strong restrictions on which segments may occur in the initial syllable.

3. Aceh-Chamic
3.1 Phonological Innovations
We now turn to the issue of the relationship of Acehnese to Chamic. Restricting matters to the etymologically Austronesian material, Thurgood states that in Chamic and Acehnese the following changes occurred:

1) PMP *n- > *l;
2) PMP *-r > *Ø;
3) PMP *-i, *-u > *-ɛj, *-ɔw, and later to [-ɔə, -ɛə] in Acehnese;
4) PMP stressed *a, *e (ə) > *aa, *a
5) Unstressed PMP initial syllables are reduced to clusters according to the same underlying patterning;
6) Imploded stops developed in some PMP etyma, reflected as /ʔ/ in Acehnese;

We will now discuss each of these in detail.

1) 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/, it is at least as likely the source of the nasal in the [mp] cluster, which case the /l/ is unexplained. The shift of *n- > *l in the 'thin' etymon is phonologically straightforward, although it may have been borrowed into Acehnese from Moklen/Moken (if not Chamic), which also shifted PMP *n- > *l, cf. MoklenLmp *lipih* 'thin (things)', MoklenKY *lipuj* 'to dream'. Other apparent loans from Moklen/Moken are discussed below. An important counter example to this proposed sound change exists in the etymon for 'coconut': PMP *niuR > Malay *nyiur*, Aceh. *bɔh ʔu*, PC *ləʔu*, where Acehnese and Chamic share the same loss of final and blocking of diphthongisation, but Acehnese has lost the initial lateral, rather than shifting it to /n/ (or potentially to [ɗ] if we accept the arguments concerning implosives, see below). There are at least two examples of this change which lack Acehnese forms: PC *lanah 'pus' < PMP *nanaq; PC *lasɛj 'rice (cooked)' cf. Malay *nasi*. The limited comparisons we have seem to establish the general rule of PMP *n- > *l in Chamic, but we have only one reasonable example in Acehnese, and it is far from clear how it acquired the form, so it may be actually be a post-Aceh-Chamic change.

2) PMP *-r > *Ø; this is a change that has occurred among other Mainland SEAsian languages, perhaps most importantly in Khmer (although other changes are also common, e.g.: /-r/ merged with /-n/ in Thai/Lao and with /-j/ in Vietnamese). In Aceh-Chamic the loss must have occurred after the diphthongisation of open syllable *u had ceased to operate, i.e.: PMP *ikuR 'tail' > Malay *ĕkor*, Aceh. *ʔiku*, PC *ʔiku*. Thurgood seems to be a little confused about the reconstruction of this final *-r, positing it in some proto-forms but not others, e.g. it is absent in his PC *ʔiku 'tail', but it is present in his *ʔular 'snake'. The change is common to both Acehnese and Chamic, so it properly belongs to the Proto-Aceh-Chamic level if it is not an independent change, although it must have occurred later, rather than earlier, in their unity.

3) PMP *-*i*, *-*u* > *-*ɛj*, *-*ɔw*, and later to [-ɔə, -ɛə] in Acehnese. 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. This is a significant change that did not occur in Malayic, although it did occur in some other MP languages, in particular Moklen/Moken. Thurgood (p.58-59) takes pains to point out that the outcome of the diphthongisation in Moklen/Moken is different to Chamic, and therefore he considers it to be unrelated. However, Larish (1999:395-402) discusses the reconstruction of the diphthongisation in Moklen/Moken in considerable detail, arguing for precisely the same initial path of development as Thurgood posits for Chamic, namely a sequence: PMP *-*i*, *-*u* > *-*i̯i*, *-*ʋu̯* > *-*ɛi̯*, *-*ɔu̯*, subsequently followed by dissimilations and mergers that ultimately yielded -*ɔj* ~ -*əj* and -*uj* in Moklen/Moken. The parallelism is remarkable, especially given the fact that Aceh-Chamic and Moklen/Moken do not sub-group genetically. What they have in common is their geographical location on the Asian Mainland, with the influence (to a greater or lesser extent) of Mon-Khmer languages (and others). Thus, while this kind of diphthongisation is otherwise rare or unknown in MP languages, it is common in MK, Cf. Khmer *dəj* 'hand' < *tii*.[4] Perhaps, given their apparent geographical separation, it was simply that under mainland influence the shift to fixed final stress set these processes on track, following parallel paths for reasons that are closed tied to universal phonetic processes. In that case Thurgood is correct to conclude that the diphthongisation in Moklen/Moken is genetically unrelated to that in Chamic, but the same argumentation works against the conclusion that Acehnese and Chamic must have derived these diphthongs together as one proto-language. The strongest evidence that they likely did is in the reflexes of words with final *ur* rhymes. As discussed above, the common loss of final *-r* must have occurred after the diphthongisation process had ceased to be productive, and therefore occurred before the separation of Aceh-Chamic, assuming that the loss was not itself also independent.

4) PMP *a, *e* (ə) >*aa, *a* in Aceh-Chamic, with later diphthongisation of *aa to /ɯə/ in Acehnese closed syllables. E.g.: PMP *qudaŋ* 'shrimp' > Malay *hudang/udang*, Aceh. *ʔudɯəŋ*, PC *hudaaŋ*; PMP *halem* 'night' > Malay *malam*, Aceh. *malam*, PC *malam*. The same shift occurred in Moklem/Moken (Larish 1999), and the lowering of PMP *e* (ə) > /a/ was the normal result in most Malayic dialects (Adelaar 1992). Much ink has been spilled discussing the issue of the long /aa/ in Acehnese and Chamic. Writers such as Shorto (1975) and Cowan (1983) saw in it evidence of a much older, perhaps ProtoAN length distinction, an idea that has not survived closer examination. Clearly we are seeing an areal drift, again connected to some extent with the shift to final stress, and reinforced by contact with languages that already have length as an important component of their phonologies. It is apparent that the lengthening of PMP *a* > *aa* must have completed before PMP *e* (ə) > *a* to have prevented their merger. This clearly places these shifts before the separation of Aceh-Chamic, and we should probably treat them as a common inheritance in Aceh-Chamic.

---

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

5) Thurgood reconstructs PC word-initial consonant clusters of the types Cr/Cl/Ch, some of which are derived from reduction of initial syllables of AN disyllabic words, while others occur in borrowed vocabulary—Thurgood refers to them as "primary clusters". The former are attested as clusters in Acehnese and all Chamic languages, so their formation belongs to the earliest stage of the proto-language. Not all AN disyllables with medials /r,l,h/ reduced to clusters in this process: compare PMP *beli* 'buy' > Malay *běli*, Aceh. *blɔə*, PC *blɛj* with PMP *balu* 'widowed' > Malay *balu*, Aceh. *balɛə*, PC *balɔw*. Thurgood does not offer an explanation of the distribution of reduced and non-reduced forms—although the presence of unstressed schwas in the first syllable of many of the relevant forms at the PMP level suggests a phonetic rule which is yet to be formulated. The point is that Acehnese and Chamic agree exceptionlessly in terms of the etyma that do and do not show the reduction to clusters. So although this kind of change is widespread in Mainland SE Asia, including within MP (including spoken Malay, not withstanding Malay authography[5]), the distribution across a specific restricted set of etyma strongly indicates an equivalent of a "Werner's Law" for Aceh-Chamic.

6) In at least two AN etyma imploded stops developed in Chamic, with /ʔ/ reflexes in Acehnese, e.g. PMP *buhuk* 'hair', Proto-Malayic *buØ(uə)k* (< PAN *buSék*), Aceh. *ʔɔʔ*, PC *ɓuk*; PMP *nahik* 'climb'> (Proto-Malayic *naØik* ?) Malay *naik*, Aceh. *ʔeʔ*, PC *ɗïʔ*, and rather speculatively PMP *hideRaq* 'lie down' > Aceh. *ʔeh*, PC *ɗïh* (although Thurgood suggests MK origins). All three are rather problematic. Firstly, there are counter examples to the regularity of the 'hair' etymology in the reflexes of PMP *bahu* 'stench' > Malay *bau*, Aceh. *bɛə*, PC *bɔw*, PMP *bahut* 'do' > Malay *buat*, Aceh. *buət*, PC *buat*, indicating that AN medial *-h-* is exceptionally, rather than regularly, reflected as *-ʔ-* in Malayo-Chamic. Although the received view (since Lee 1966) is that PC *ɓuk* reflects a sporadic fusion of /b/ and /ʔ/, by implication it also requires the sporadic persistence of *-ʔ-* in Malayo-Chamic.

Thurgood compares PC *ɗïʔ* 'climb' to Bahnar *dək* 'go up' (citing Cabaton 1901, note that Banker et. al. 1979 gives the form as *dak*). One can also compare to Proto-Katuic (Sidwell 2005) *dïk* 'lift up, raise', although these may not be helpful—the Katuic and Bahnaric suggest a prototype *dak*, which simply does not correspond to the Chamic form. On the other hand there no problem deriving Acehnese *ʔeʔ* from PMP *nahik* in the light of examples such as PMP *niuR* > Malay *nyor*, Aceh. *bɔh ʔu*, PC *lɔʔu*. The problem is how to account for the implosive initial in Chamic, and similarly the received view is a sporadic fusion of /n/ and /ʔ/.

We do not have an obvious AN etymology for Aceh. *ʔeh*, PC *ɗïh* 'lie down', although they could reflect a radical simplification of the trisyllabic PMP *hideRaq*. I have yet to find a convincing mainland source—among MK languages Khasi *thiah* 'lie down, sleep' potentially corresponds, but the geographical distance makes it a remote prospect, while Khmer *dek*, compared by Cowan, is phonologically too different (and probably ultimately related to Khasi *thiah*).

In addition to the above three sets with Acehnese reflexes, Thurgood reconstructs 12 PC words with initial *ɓ* and 10 with initial *ɗ* that lack Acehnese reflexes—all 22 are borrowings, which must have been acquired after the separation of Acehnese. So we have three words in which Chamic implosives correspond to Acehnese /ʔ/, but we don't know

---

[5] Drawn to my attention by David Gil in 2001 during a visit to the Max Planck Institute (Leipzig).

whether there was a shift of imploded stop to glottal stop in Acehnese, or a simple loss of initial syllable from a disyllabic PAC form.

On balance there are several phonological developments that solidly belong to a phase of Aceh-Chamic unity—the formation of Primary Clusters, the diphthongisation final *-i and *-u and the lass of final *-r which followed the diphthongisations. To these phonological changes we can add the lexical innovations—borrowings—common to Acehnese and Chamic.

3.2 Lexical Innovations

In this section I discuss the data and results of two significant publications dealing with the sources of borrowings in Aceh-Chamic: Cowan (1948) and Thurgood (1999). Additionally I would have have liked to make use of Collins' (1975) PhD thesis on the sources of Acehnese vocabulary, but access to that work is restricted[6].

Cowan's 1948 paper made a fundamental contribution to discussion of the classifiation and history of Acehense half a century before Thurgood's recent synthesis appeared. Cowan discuses at length the position of Cham and Acehnese in respect of Austronesian, adducing many lexical comparisons with Malay. He groups Cham and Acehnese genetically on the basis of parallels in the phonology, morphology, lexicon and syntax, and interestingly contrasts them in respect of the use of pronouns and the "passive" voice (see Durie 1985 for a detailed analysis of Acehnese argument structure). Significantly for our present purposes, Cowan presents a list of 150 comparsions with mostly Mon and Khmer: of these I count 43 that can be confidently identified as MK loans into Acehnese, and perhaps another 60 into Aceh-Chamic, while the balance are put aside as either defective comparisons, misidentified Austronesian or other loans, imitative forms, or loans into MK languages from Chamic. A summary of Cowan's numbered examples thus excluded is at the end of Appendix 1. Of Cowan's MK loans into Aceh-Chamic, I count 17 sets not included in Thurgood's published data-set, which suggests that he did take full advantage of Cowan's contribution. This might seem a modest number at first, but in fact the total number of Thurgood's putative MK borrowings with an Acehenese reflex is modest—only some dozens—and is fact is given considerable attention in the following analysis.

Thurgood identifies some 277 Proto Chamic words of Mon-Khmer origin and another 179 of uncertain origin. One way or another we assume that the bulk of these are borrowings, although conceivably some are neologisms invented by Aceh-Chamic speakers. Dyen, in his 2001 review of Thurgood, expressing considerable scepticism about the Aceh-Chamic hypothesis. He pointed out that if Acehnese is descended from PC, it should preserve a substantial proportion of the borrowings reconstructable to PC, yet he counted only 44 Acehnese reflexes among the hundreds of PC items of MK origins. Reasoning further that those words also having Malay reflexes could well have diffused from Malay, only "twenty-eight entries, perhaps better reduced to twenty-six, then appear

---

[6] Durie (1975:3) reports Collins' conclusion that Acehenese "had contact with Old Mon, the Aslian languages of the Malay peninsula, and the languages of the Nicobar islands". In my own investigations so far I have found no particular lexical or structural features among the MK component in Acehnese that would identify an Aslian or Nicobaric source. I believe that this is consistent with the homeland of Aceh-Chamic being in Indo-China, and the relatively marginal importance of Aslian and Nicobaric in the trade networks of western Austronesia.

to constitute the basis of the hypothesis that Acehnese is a Chamic immigrant". In other words, only 10% of PC words of MK origin have Acehnese reflexes.

This is a very significant discrepancy. If Acehnese is a descendant of PC, it should reflect PC vocabulary pretty well as much as any Chamic language (subject to extraordinary social/historical factors). Furthermore, if Acehnese is the first branch of the Chamic family tree, the principal criteria for reconstructing a non-AN word to the PC level should be its attestation in at least Acehnese and one other Chamic language. Yet we have gross indications that Acehnese shares relatively few borrowings with the rest of Chamic, a fact that suggests that Acehnese separated before the bulk of borrowings into Chamic occurred.

Reviewing Dyen's count it seems that he did not consider the complete corpus of data presented by Thurgood—but ignored the words classified as of uncertain origin. I have made my own count combining both indices and the results are summarised as follows:

1. 16 borrowings also reflected in Malay
2. 7 words apparently borrowed separately into Acehnese and Chamic
3. 3 isoglosses with Moklen/Moken, origin and direction of borrowing uncertain
4. 28 AC borrowings of MK origins
5. 12 AC borrowings of unknown origins

1.)

Semantic	Aceh.	P-Chamic	Malay	Comment
'bean, pea'	*ruituɯʔ*	**rətaak*	(Iban *retak*)	Cf. Khmer *sandaek*
'bitter'	*phet*	**phit*	*pahit*	< Skt. *pitta*
'bowl, dish'	*piŋan*	**piŋan*	*pinggan*	< Persian (> Bah.)
'branch, fork'	*cabɯəŋ*	**caɓaaŋ*	*cabang*	>Aslian, Cf. Kh. *ɟəŋrmaaŋ*[7]
'broken, break'	*picah*	**picah*	*pĕcah*	> Bah.
'buffalo'	*kɯbɯə*	**kabaw*	*kĕrbaw*	> Bah.
'cotton'	*gapɯəh*	**kapaas*	*kapas*	< Skt. *karpaasa*
'cow, ox'	*lɯmɔ*	**ləmɔ*	*lĕmbu*	Cf. Khmu *lmboʔ*
'eggplant'	*truəŋ*	**trɔŋ*	*tĕrung*	> Bah.
'form, image'	*rupa*	**rupa*	*rupa*	< Skt. *rupa*
'g-grandchild'	*cʌt*	**cicẽt*	*cicit*	
'gold'	*mɯh, mɯih*	**ʔama(a)s*	*emas*	> Khmer, Bah. < ?
'lizard, gecko'	*cicaʔ*	**cicaʔ*	*cicak*	Cf. Mon *həcɛk* (imitative)
'net (casting)'	*ɟɯə*	**ɟaal*	*ɟala*	< Skt. *jāla*
'pillow'	*bantaj*	**bantal*	*bantal*	
'pineapple'	*bɔh ʔanɯh*	**manaas*	*nanas*	< Portuguese

Group 1 is an etymologically heterogenous set of borrowings that fall mostly into two main types, Indic words that probably began to be diffused by traders even before the Common Era, but particularly from the middle first millennium (as Indic scripts and other

---

[7] 'forked stick'

cultural features were widely adopted), and MK words, some of which have clear etymologies, others identified on structural grounds that are inferred to be MK. A good example is Malay *kĕrbaw* 'buffalo'—close matches are found in Bahnaric and Katuic, but the Khmer reflex is *krəbɤj*, which shows phonological differences that eliminate it as the source. The other bovid term, reflected in Khmu *lmboʔ,* Bahnar *ləmɔɔ,* Vietnamese *bò*, is drived from PMK *ʆ]bɔʔ* 'hump of ox' by Shorto (ms.) based on reflexes in Mon and Khmu. Speculatively the *kĕrbaw* word could have originated from the same root, assuming borrowing from a hypothetical MK language having lost the final glottal and added the small animal velar prefix (not uncommon changes in EMK).

Another interesting etymon is the 'gold' word. On the mainland it is restricted to languages historically in contact with Chamic, which suggests borrowing into MK, but that still leaves the question of its source in MP. An MK root **jaas* 'to shine' is reconstructable on the basis of widely distributed reflexes, and a hypothetical derivation via the *-m-* agentive infix in pre-Mon (cf. Old Mon /jimaas/) could have subsequently diffused with the very sought after trade item.

At this stage the main point I would like to make about these comparisons is that the borrowing of MK words into Malayic likely did not reflected a discrete historical process that might be localised in time or space. It is evident that the borrowings range from relatively recent Khmer, Mon and Vietic loans to very ancient times. Whatever the case Dyen is correct to set these aside from any discussion of Chamic sub-grouping.

Group 2 consists of words for which we have indications of independent borrowing of related or unrelated but similar forms:

2.)

Semantic	Aceh.	P-Chamic	Comment
'flesh, meat'	*siə, ʔasɔə*	**ʔusar*	Aceh. related to Malayic **isi*
'fly (v.)'	*phʌ/pʌ*	**pər*	PMK **par.* Anomalous aspiratred initial also found in Rade: *phiər* (Durie 1990)
'open (mouth)'	*hah*	**ʔaha*	PMK **haʔ, hah*, Ach. resembles B. & Viet.
'python'	*lhan, tlan*	**klan*	PMK **tlan* - Aceh. borrowed with apical initial; Chamic < form with velar initial
'strong, hard'	*kʌŋ*	**khaŋ*	Comp. Aceh. to Katuic **kəŋ,* Khmer *kèəŋ* (& Thai *khàŋ*) suggest **gaŋ.* Chamic < Vietnamese **khǎŋ*
'wash'	*rhah*	**raw*	Cf. Viet. *rửa* (< **raah*), Katuic/Bahnaric **ʔəraaw*
'yawn'	*sɯmɯŋɯp*	**həʔaap*	PMK **sʔaap, *sŋʔaap*, not all MK sub-groups have medial nasal

Group 2 items all show clear phonological indications that Acehnese and Chamic borrowed related forms from different MK sub-groups. This is quite understandable as lexical borrowing continued after separation, and therefore these forms are not relevant to the sub-grouping issue.

Group 3 is quite intriguing:
3.)

Semantic	Aceh.	P-Chamic	Proto-Moken/Moklen
'naked'	*lhon*	*(ma)(sa)lun*	*ɲulɔn.* No wider etymology apparent.
'urinate'	*ʔiəʔ*	*maʔiãk*	*niʔaak* >Pre-Moklen *niʔiək < PMP *[ ]iSeq
'gecko'	*paʔɛɛ*	*pak-kee*	*tɔkɛɛʔ,* imitative word?

The phonological agreements in the first two sets above are excellent, and strongly suggest ancient contact involving Aceh-Chamic and Moklen/Moken—in particular the development of the diphthong in the 'urinate' etymon indicates Moklen/Moken as the source. Larish reconstructs the Moklen/Moken homeland as the Isthmus of Kra, with their marginalised to the islands off the western coast only later. This leaves the possibility of A-C and M-M contact somewhere on the Gulf of Thailand.

Group 4 items are the most numerous, all showing indications of being borrowed from MK:

4.)

Semantic	Aceh.	P-Chamic	MK comparisons
'arm'	*sapaj*	*sapal*	Found in Asl., Kat., West-Bah.
'back'	*ruəŋ*	*rɔŋ*	Katuic *krɔŋ 'back' , Khmu kndrɔɔŋ 'back'
'bail'	*suɯət*	*sac*	PMK *saac (all but Khmu, Asl., Nic.)
'bird'	*cicem*	*cim*	PMK *cim (all but Khmer)
'carry on sldr.'	*gulam*	*gulam*	PMK *klam or *kləm (NMK & Aslian)
'chase'	*tijɯəp*	*tijaap*	Khmu ŋgjaap, Ch. > Tampuon tijạap
'cheek, jaw'	*miəŋ*	*miaŋ*	Khmu miəng 'chew', Vt. miệng, < PV *mɛɛŋʔ 'mouth'
'chin, jaw'	*kɯəŋ*	*kaaŋ*	PMK *kaaŋ (Katuic, Bah., SNic., Vietic) (+ kmaaŋ forms in Pearic, Vietic..)
'cover'	*gɔm*	*gəm*	Khmer kaem 'cover, encrust, decorate', PVietic *kəmʔ 'to bury'
'crow'	*ʔaʔaʔ*	*ʔaak*	PMK *kʔaak (all but Khasi, Nic.) – Vietic reflexes typically ʔaak, e.g. Viet. ác, but such imitative words are problematic.
'cut off'	*kɔh*	*kɔh*	PMK *kɔh (Bah.,Kat.,Nic.,Asl.)
'dry'	*tho*	*thu*	Temiar təhool, KhmuYuan thúu
'dumb'	*klɔ*	*k-am-lɔ*	Khmer kamlaw 'ignoramus'
'empty'	*sɔh*	*sɔh*	Khm., Bah., (Katuic infixed forms only)
'escape'	*lhuɯəh*	*klaas*	> Bah., other MK suggests *laas 'leave'
'forget'	*tuwʌ*	*wər*	PMK *wər 'go round' ? (all MK groups)
'hawk, kite'	*kluɯəŋ*	*klaaŋ*	PMK *klaaŋ (all MK groups)
'house'	*suɯəŋ*	*saaŋ*	Khmer saaŋ 'to build' (also >Thai/Lao)

'lick'	*liəh*	**lijah*	PMK **liət,* also Khasi *ɟliah*
'mount. range'	*cʌt*	**cət*	Khmer *caot* 'high, steep, sheer, abrupt'
'neck'	*takuə*	**takuaj*	PMK **kuuj* 'head' (Kat., Asl.)
'other, group'	*gɔp*	**gəp*	PMK **gap, gəp* 'friend, associate' (Khm., Bah., Viet.)
'peck (snake)'	*cɔh*	**cɔh*	PMK **[ʔ]cɔh* (EMK, Khmu, Asl.)
'pillar, post'	*tamɛh*	**tamɛh*	Mon *tmit* 'post supporting veranda'
'river'	*kruəŋ*	**krɔɔŋ*	PMK **ruŋ, *ruuŋ, *ruəŋ* (all but Asl., Nic.)
'stand, stop'	*dəŋ*	**dʌŋ*	Viet. *đứng,* or perhaps PMK **duŋ* 'house'
'strike, pound'	*pɔh, pɛh*	**pɔh*	PMK **pah, *puh, *puəh* (NMK, Bah., Viet.)
'wrap'	*sɔm*	**səm*	Old Khmer *sum* 'to wind, roll, wrap up'

To these we can add the Aceh-Chamic-MK comparisons from Cowan (1948) not used by Thurgood, yet which may be taken as highly indicative of MK borrowing.[8]

Aceh.	Cham	MK Comparisons
hu 'ablaze'	*hu* 'roast'	Kh. *chur* 'ignite', Bah. *huur* 'roast', Katu *huar* 'singe'
ɟa 'ancestor'	*ɟa* 'appelative'	OldMon *ɲɟaʔ,* OldKh. *ɟi* 'great-grandmother'
baʔ 'at, on'	*pak* 'at, towards'	OldMon *bak* 'up to, until'
luəŋ 'channel'	*haluŋ* 'pit, canal'	Khmer *lùŋ* 'dig hole', *ʔɔnlùəŋ* 'hole in stream-bed', Bah. *səluŋ* 'pit, ditch'
tɔm 'ever'	*tom* 'meet with, accomplish'	PMK **təm/*təəm/*tam* 'begin', e.g. Mon *tam / tɔm* 'base, beginning' (widespread in MK)
ɲum 'flavour'	*ɲəm, ɲam* 'to taste'	Praok *ɲɔm* 'to taste', Bahnar *ɲaam* 'delicious'
wɛh 'go away'	*weh* 'to dodge'	Khmer *veh* /*vèh*/ 'to slip away, escape, dodge'
gət, gɛt 'good'	*gɔt* 'just'	Khmer *gat* /*kɔt*/ 'just, exact'
chen 'affection'	*khin* 'want, like'	Viet. *xin* 'beg', Palaung. *sin* 'desire', OldMon *chān* /*chan*/ 'to pity/
khem 'laugh'	*khim* 'smile'	LitMon *k'im* /*kʔim*/ 'to smile'
buɯŋəh 'morning'	*paguh* 'morning-light'	Mon *peŋuh* 'to awaken'
khem 'laugh'	*khim* 'smile'	LitMon *k'im* /*kʔim*/ 'to smile'
weŋ 'to pedal'	*wiŋ* 'turn, whirl'	PMK **wiŋ* &c. (with many variants) 'go round'
ʔuət 'polish, rub'	*ɟak* 'rub'	Lawa *ʔuət* 'wipe', Khmu *ʔɔɔt* 'scrub body'
kuət 'scrape'(C.)	*kɟac* 'dig'	Khmer *khvaac*, Kensiw *kəwɔɟ* 'scratch up'
wɯə 'stable,pen'	*wa(r)* 'yard, stable'	Khmer *val* /*viel*/ 'plain, clearing, plaza', Mon *wa* /*wèa*/ 'open space, pasture'

---

[8] Note that Acehnese forms have been normalised to Daud & Durie (1999), Cham forms are from Cowan, MK comparisons have been corrected/augmented

dəm 'stay o.night'	*dəm* 'id.'	PMK **dəm*, e.g. Mon *dəm* /*tə̀m*/ 'to lodge'
bʌt 'stretch'	*but* 'twisted'	Khmer *bot* /*pot*/ 'to curve, fold'
cɔʔ 'take, sieze'	*cɔk* 'id'	Khmu *cɔɔk* 'catch (e.g. pig)', *cok* 'take out (e.g. entrails)', WestBahnaric **cɔk* 'take'

Examining the above sets we note no convincing pattern of borrowing from a single dominant source—Khmer and Mon are well represented but this may simply reflect the reliance on those reference material. Some etyma are well distributed across the MK family with no particular phonological clues for their source in Aceh-Chamic (such as 'crow', 'fly', 'hawk'). There are several Khmer isoglosses (e.g. 'cover', 'dumb', 'gold', 'house', 'mountain range', 'wrap') although the lack of wider MK etymology is also suggestive of borrowing into Khmer. And there are several items where the closest MK comparisons are in Northern MK languages, and it is difficult to see how they could be the source of borrowings. It is also significant that there are so very few prospective Vietic or Katuic sources for these words, given Thurgood's suggestion that:

> ...the Acehnese were the most northerly of the Chamic groups, covering an area now populated by, among others, the modern Katuic speakers. (p.42)

This idea appears to be based on the overriding assumption that the break-up of Chamic was driven by one main historical process—the Vietnamese imperial drive southward. The model assumes that as the Acehnese were the first group to break away, they must have been the first to suffer Vietnamese pressure. Logically there are other possibilities to consider, such as a southern origin of Acehnese somewhere in the vicinity of the Mekong Delta/Funan. My problem is that no particular solution appears to be supported empirically by comparative linguistic data. Thurgood bases his claims upon supposed morphological and lexical arguments. The first of these is a comparison of the *tar-, t-, ta-* prefixes in Katuic with parallels in Austronesian which Thurgood (p240-241) asserts are "too close to be accounted for by mutual inheritance", and suggests that because some lexical borrowing from Chamic into Katuic is attested, the same is likely to explain the morphological parallels. A contra-opinion is offered by Diffloth (1994) who points out that the various MK affixes with parallels in An are actually widespread in MK. He concludes that:

> Ironically, it is the relative poverty of shared vocabulary between Austroasiatic and Austronesian, combined with evident agreement in morphology, that argues for a genetic, and against a contact relationship between the two families. (Diffloth (1994:312)

Thurgood writes (p.240-241):

> Other evidence of a contact with Chamic, particularly into Acehnese, and an apparent Austronesian morphological strata (sic.) in Katu (Reid 1994), which one would presume were due to Chamic influences.
>   The obvious way to account for the Katuic strata found in Chamic is to assume that Chamic influence extended up along the coast into Katuic territory. Certainly, an examination of the appendix of forms makes it abundantly clear that there are a considerable number of MK forms, attested in the more northerly Katuic but not in the more southerly Bahnaric. Further, many of these are attested in Acehnese. Thus, the

most likely scenario is to assume that the Acehnese are the descendents of the most
northerly group of Chamic speakers.

Consistent with Diffloth above, Reid (1994) makes no claim of borrowed "Austronesian
morphological strata in Katu". In his paper Reid compares the Austroasiatic prefixes *pa-*
and *ka-*, which "can be reconstructed with a causative function" with the Austronesian
causatives **pa-* and **ka-*, exemplifying the former with examples from Katu. The
comparison is explicitly between two language families with consideration of the Austric
hypothesis in mind, with much weight given to reflex the of **pa-* in Nicobarese.

Thurgood then refers to "Katuic strata found in Chamic", including a claim that that
stratum is shared with Acehnese. No specific examples are presented for this claim, just
the assertion that it is "abundantly clear" from perusing the appendix to the book. I
strongly disagree that one could reach such a conclusion on that basis, since a careful
examination of the appendix makes it clear that there are no examples where Katuic can be
unambiguously identified as the source of an Aceh-Chamic word. Thurgood's comparisons
of Acehnese with Katuic, with my commentary, follow:

> PC **ʔɛh* 'excrement', compares with both P-Katuic and P-Vietic **ʔɛh*; Acehnese *ʔɛʔ*
> matches neither as its final suggests **ʔɛk*.

> PC **ʔaak* 'crow', Acehnese *ʔaʔaʔ*, while Katuic suggests **kaʔaak, *ʔaʔaak*, so do
> basically all MK languages, yet Acehnese fails to show the regular /ɯə/ reflex
> of /aa/, indicating a more recent imitative (re)formation.

> PC **ʔaha, *ha* 'open mouth', Acehnese *hah*, most MK language share this clearly
> sound-symbolic formation, yet the Acehnese fail to agree in the final. Thurgood
> compares to Peiros' p-Katuic **təha, *ʔəhah*, but the back vowel does not
> match.

> PC **dəŋ* 'stand; stop', Acehnese *dʌŋ*, compared to Peiros' p-Katuic **ʔtəjiŋ, *əʔjɨŋ*,
> but there is no correspondence between the forms.

> PC **kaaŋ* 'chin; jaw', Acehnese *kɯəŋ*, compared to Peiros' p-Katuic **təʔbaaŋ*, but
> there is no correspondence between the forms.

> PC **kalaaŋ* 'hawk; bird of prey', Acehnese *klɯəŋ*, compared to Peiros' p-Katuic
> **kəlhaaŋ*, but the word is found throughout MK and is even in some Malayic
> languages, e.g. Malay *helang*.

> PC **kapaas* 'cotton', Acehnese *gapɯəh*, compared to Peiros' p-Katuic **kəpaajh*, but
> the word is an Indic borrowing found throughout MK and Malayic languages,
> e.g. Malay *kapas*.

> PC **klaas* 'escape', Acehnese *lhɯəh*, compared to Thomas' p-Katuic **-klah, *-lah* but
> the distribution of the word suggests borrowing into Katuic and Bahnaric.

> PC **krɔɔŋ* 'river', Acehnese *kruəŋ*, compared to Peiros' p-Katuic **kərhuaŋ*, but other
> MK such as Vietic **krɔɔŋ* are more likely—even Thai has reflexes of this MK
> root.

> PC **lɔɔk* 'to peel', Acehnese *pluəʔ*, compared to Peiros' p-Katuic **liɛt, *luɔt* but there
> is no correspondence.

> PC **picah* 'broken; break', Acehnese *picah*, compared to Peiros' p-Katuic **pəc[ə/a]h*,
> **kəc[ə/a]h* but the phonology and distribution suggest borrowing into Katuic
> and Bahnaric.

PC *pər* 'to fly', Acehnese *phʌ*, compared to Peiros' p-Katuic *par*, *paar,* although basically any MK language could be the source for Chamic, the Aceh. aspirated initial is not explained (some Pearic languages and Khasi did shift plain stops to aspirates but there is no convincing evidence of Pearic or Khasi influence).

PC *raw* 'wash', Acehnese *rhah*, compared to Peiros' p-Katuic *ʔəriaw* but the Acehnese form does not correspond.

PC *sapal* 'arm', Acehnese *sapai*, compared to Thomas' p-Katuic *qapaal* 'shoulder'. This etymon also found in Aslian (as 'upper arm') and Pearic (as 'palm (of hand)'). The problem is that the Chamic reflex has a short main vowel, and only Aslian shows a neat semantic and phonological match.

PC *sɔh* 'only; empty; free, leasure', Acehnese *sɔh*, compared to Peiros' p-Katuic *[s/c]ənhah* but Katuic all show infixed forms, unlike Bahnaric and Khmer.

PC *trɔŋ* 'eggplant', Acehnese *truəŋ*, compared to Peiros' p-Katuic *həŋgɨŋ*, *səkɨŋ* but there is no correspondence. The word is found in Malayic, e.g. Malay *terung*, which is probably more indicative of origin.

Of these 16 comparisons, few, if any, could be put forward as evidence of a Katuic stratum in Chamic, and certainly none demonstrate a Katuic stratum in Acehnese. Importantly several (such as 'wash', 'crow', 'excrement') show differences that suggest independent borrowing. As far as I can tell from the evidence I have assembled there is nothing to indicate a geographical location for Acehnese in relation to the present distribution of Chamic languages. For this reason my default hypothesis is that Acehnese separated from Chamic at a time before Chamic had developed any significant internal diversity.

The regularity of the phonological agreements between Acehnese and Chamic in their common borrowed vocabulary strongly indicates that most, if not all, these lexical items reflect a phase of Aceh-Chamic unity. Given that there is no standout source evident among known MK languages, two possibilities present themselves: a) proto-AC had contact with a range of MK languages from which it borrowed, or b) an unknown MK language that has not otherwise survived was in contact with proto-AC and contributed these borrowings—in the latter case the MK parallels adduced above are simply related MK reflexes rather than source forms.

Below I list the Aceh-Chamic borrowings without apparent wider etymologies (with borrowing into Bahnaric via Chamic indicated):

5.)

Semantic	Aceh.	P-Chamic	Comment
'arrive'	*troh*	*truh*	(> Bah.)
'descend, sink, collapse, destroy'	*lhʌh*	*gləh*	(> Bah.)
'dry weather; drought'	*khuəŋ*	*khɔɔŋ*	
'handle (of knife)'	*gʌ*	*gər,*	(> Bah.)
'many, much'	*lə*	*luu,*	(> Bah.)
'neg. imperative'	*bɛʔ*	*bɛʔ*	(> Bah.)
'peel'	*pluəʔ*	*lɔɔk*	(> Bah.)
'pick, pluck'	*pʌt, pɛt*	*pɛt*	(> Bah.)

'snail'	*ʔubo*	**ʔabaw*	
'straw (rice)'	*ɟumpuŋ*	**puuŋ*	
'that, there'	*sideh, hideh*	**dih*	(> Bah.)
'use'	*ŋuj*	**ʔaŋuj*	

Most of the above 12 items are also present in Bahnaric languages, although the lack of reflexes in West Bahnaric (see Sidwell & Jacq 2003) and in the rest of MK clearly indicates that what Thurgood took as straightforward MK > Chamic loans were actually borrowed from Chamic into Bahnaric, originating from an unknown source. Phonologically the words look like they are from MK—half are simple monosyllables while the rest have initial clusters or are sesquisyllabic, so our default hypothesis is that they come from some MK language or languages, the identity of which is unknown.

Can we link the group 4 and 5 etyma somehow without straining possibility too far, given that they are all at least reconstructable to PAC? I believe that it is worth speculating on this. First of all, it is a fact that each MK sub-group has a set of lexicon that is not shared with any other MK sub-group, since lexical innovation is a continuous process and an important aspect of the accretion of differences that drives linguistic diversification. Logically then, if an MK speaking community were absorbed by language shift into PC, a process that we strongly suspect did happen in ancient times, one of the consequences would be the borrowing of a set of words, some of which have a wider MK etymology, and some not, although the latter would none the less have the formal structural characteristics of MK lexicon.

This statement characterises not only the 42 AC borrowings discussed above, but also the bulk of the PC lexicon of borrowed or unknown origin reconstructed by Thurgood. Allowing for some errors and reassignments we have approximately 450 words in the PC lexicon that are borrowings or otherwise innovated, of which so far only 42 (or less than 10%) have been identified in Acehnese. It thus appears that Acehnese did not participate in a major phase of the lexical development of PC, presenting us with a significant problem of historical explanation.

4. Quantification of Etymological change and distance

Now that we have some rough indication that there is a significant difference in the absolute quantity of contact-induced change experienced by Acehnese and (the rest of) Chamic, I want to move forward to quantify this in a more representative fashion. My concern is that we don't know to what extent the PC lexicon reconstructed by Thurgood is representative of the real PC lexicon, and therefore the extent to which we can fairly compare and analyse the figures discussed above.

It is in the nature of proto-languages that they are constructs that, due to the availability of sources and various accidents of history, are necessarily incomplete or even skewed in terms of their representation of the lexicon. For example, it is commonly held that some areas of the lexicon are less stable than others, such as words representing more abstract meanings over the more concrete ones, and therefore concrete meanings will be potentially over-represented in a reconstructed lexicon. Now it is clearly beyond the scope of this paper to consider complete lexicons (whatever that might mean in practice), so I set about to devise a method that would go some way towards more fairly quantifying the proportions of lexical change in Acehnese and Chamic.

In the first place we acknowledge that Acehnese and Chamic are descended directly from Proto-Malayo-Chamic or something not very much removed from that. The Malayic sub-group of AN is already the subject of a comprehensive reconstruction (Adelaar 1992), so in the absence of PMC we might reasonably use it as a base line for quantifying the amount of lexical innovation in Acehnese and Chamic. Now I understand that there are a number of assumptions here that can be challenged, but I proceed on the basis that we are looking for a broadly indicative method, rather than a very precise tool, and one whose initial results can surely be improved by subsequent more detailed analysis. Accepting this programmatic rationale we move on to the details.

I take as my starting point the Malayic basic lexicon of 200 items as reconstructed by Adelaar (1992), using the diagnostic semantic list developed for MP languages by Hudson (1967). The 200 word list contains items from a range of semantic domains and word classes, and for our purposes I take it that for any MP language which we compare on the basis of this list its genetic classification will be evident, and the degree of lexical change from PAN, PMP or any other known starting point will be readily calculated. I copied the P-Malayic items into a spreadsheet and then added the etymologically equivalent PC and Acehnese reflexes. Where lexical replacements have occurred the new words are put in place. This is different to the strictly semantic approach of lexicostatistics which is necessarily blind to etymology in the initial compilation of the lists for comparison. I did this because I want to quantify the amount of lexical borrowing as opposed to the amount of semantic change within the lexicon.

Due to the incompleteness of the PC lexicon and Acehnese sources at my disposal the total list was reduced to 183 items.[9] The resultant list is presented as an appendix to this paper. The analysis of the list begins with counting the various common etymological retentions and innovations. Note that in some cases there is more than one form given in the sources for a given gloss, these are noted in the appendix, but in the counts below I have still treated these as single items. A summary of the results follows:

- 96 items (52.5%) where all three languages (Aceh., PC, PM) show direct inheritance of AN forms or Malayo-Chamic innovations
- 51 items (27.9%) Aceh innovations (discounting Malay borrowings)—of which 26 are shared with P-Chamic and 25 are unique to Aceh.
- 73 items (39.9%) Chamic innovations, including 26 shared with Aceh, and 47 unique to Chamic.

The above figures give a sense of proportion to the great extent of borrowing in PC in particular—approximately 40% of the basic lexicon replaced by mostly borrowed vocabulary. By contrast only just over a third (26/73), of those replacements in PC are also reflected in Acehnese.

Accepting the MC hypothesis, and Blust's estimate of MC separation around 2300 BP, plus Thurgood's estimate of a late 1st millennium break-up of PC, we would look to place the separation of Acehense somewhere in a 1000 or so year window from roughly 300 BCE forward. Taking the even bolder step of assuming a more or less stable rate of lexical replacement the above figures would place the separation of Acehnese in

---

[9] I considered supplementing with available items to bring it up to 200, but decided not to lest I further skew the results by my selections.

approximately the first century CE, shortly before the first historical references to Champa appear. Citing archaeological evidence, Thurgood (p.16) places the pre-proto-Chamic settlement of the Indo-Chinese coast at sometime before 600 BCE, which on my calculations would place the separation of Acehnese in the first or second century BCE.

This is only a broadly indicative calculation. Frankly I do not wish to make a claim for a stable rate of lexical replacement—since decades of experience with glottochronology have shown that the rate of change in language in respect of borrowings is quite unstable, given the possible social factors. None-the-less the fact that Acehnese demonstrably participated in only a minority of the contact driven lexical replacement that affected the rest of the Chamic strongly indicates that it separated at a much earlier than assumed by the Thurgood model. The stratum of common borrowings suggests that Acehnese split away during the early stages of a phase of assimilation of an unknown but presumably MK speaking population into the nascent Champa.

Thus one may take Thurgood's conclusion:

> The early arriving pre-Chamic peoples most likely landed south of Danang and thus probably encountered Bahnarics. Given the major restructuring of the arriving Austronesians 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.
> [….] Probably sometime around the fall of Indrapura in the north, although it may have been as much as several centuries earlier or later, the Chamic speakers who were to become the Acehnese left the mainland on a journey that would ultimately end in northern Sumatra. (p.251)

and reformulate it as follows:

> The early arriving pre-Chamic peoples most likely landed south of Danang and encountered a Mon-Khmer speaking population of undetermined classification. Given the major restructuring of the arriving Austronesians language that took place, these pre-Chamic people must have become socially dominant, with this dominance leading many or all of the Mon-Khmer speaking people to shift to Cham.
> [….] Sometime during this early phase of language shift, perhaps before the beginning of Common Era, the Chamic speakers who were to become the Acehnese left the mainland on a journey that would ultimately end in northern Sumatra.

To what extent can we reconcile this with known history? Durie, discussing the founding of Champa in the second century CE, writes:

> From Chinese sources we know that there were several kingdoms during this period on the trade route to China around the Isthmus of Kra, the Malay peninsula, and the gulf of Thailand. One such was Funan, which was centred on the lower Mekong. Several kingdoms in the Isthmus of Kra were subject to it. It was overwhelmed by Khmers in the 6th century. We have no record of the language of Funan, but it could well have been a sister of early Chamic. During this period it would have quite likely for Funan traders to have been established in the Malay peninsula and even North Sumatra, which was in a strategic position for the trade with India. (Durie 1985:3)

So Durie suggests that Aceh may be a surviving fragment of Funan. Contra Thurgood, in that case the Acehnese were a southern branch of Aceh-Chamic that split off as Funan fell. The trouble I see with Durie's idea is that Funan fell to the Cambodians, and it is clear that the mysterious loan stratum found in Chamic and to a lessor extent Acehnese cannot be related directly to their language. I would like to suggest an alternative, in which the Funanese, or a segment of Funanese society, were speakers of an unrecognised branch of Mon-Khmer, and were absorbed into Champa as they lost their political and economic centre to Chenla/Ankor. Perhaps related events drove the Acehnese from the mainland, just as a thousand years later the Moklen/Moken were driven off the Isthmus of Kra by Thai expansion.

5. Conclusion

Thurgood's formulation of Acehnese as a "Chamic language" obscures an important distinction in the historical development of these languages. Alternatively I would suggest that we classify Acehnese as an "Aceh-Chamic" language, an offshoot of a stage intermediate between PMC and PC. The redrawn MC family tree, suggested by my analysis, is represented as follows:
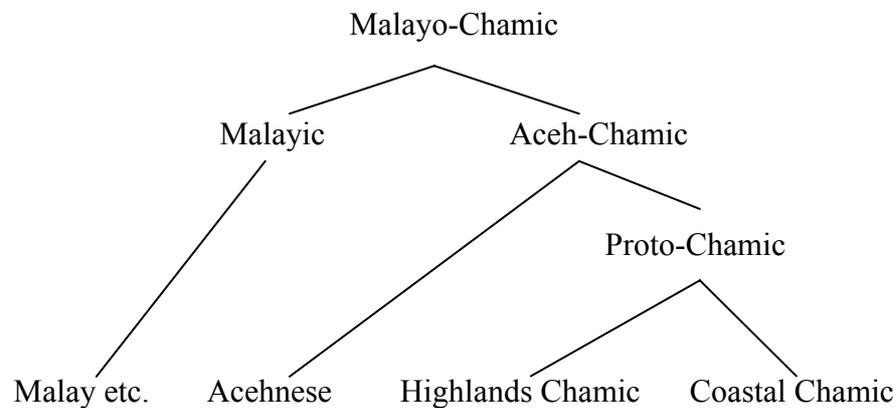
Malayo-Chamic

Malayic   Aceh-Chamic

Proto-Chamic

Malay etc.   Acehnese   Highlands Chamic   Coastal Chamic

Figure 2: *Revised Malayo-Aceh-Chamic tree*

From a programmatic perspective the redrawing of the Stammbaum begs a major overhaul of the Acehnese and Chamic comparanda and their comparative-historical analysis. The resultant phonological and lexical reconstructions should be stratified into Aceh-Chamic and Proto-Chamic levels. Naturally one would seek to include in such a project:

- any new or otherwise un(der)utilised Chamic sources
- more extensive reference to Mon-Khmer sources, especially Khmer, Vietnamese and Mon, as well as more recent Mon-Khmer comparative reconstructions
- reconstruction of Proto-Acehnese based upon dialect comparison

I expect that the latter point may prove especially important, as Acehnese, although more affected by Malay, was protected by geography from much of the MK influence that has altered the face of Chamic.

References & Sources

Adelaar, K. Alexander. 1992. *Proto Malayic: The Reconstruction of its Phonology and Parts of its Lexicon and Morphology.* Pacific Linguistics Series C-119. Canberra, Australian National University.

Benedict, Paul K. 1976. Austo-Thai and Austroasiatic. In Jenner et al. (eds.) *Austroasiatic Studies, Part 1.* Oceanic Linguistics special publications No. 13. Honolulu, Unversity of Hawaii Press. pp 1-36.

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 & Thomas Hudak (eds.). Tempe Arizona, Arizona State University. pp:25-83.

----- 1981. The reconstruction of Proto-Malayic-Javanic: an appreciation. *Bijdragen tot de Taal-, Land- en Volkenkunde* 137.4.:456-469.

----- 2000. Review of Thurgood 1999. *Oceanic Linguistics* 39: 435-445.

Burnham, Eugene C. 1976. *The Place of Haroi in the Chamic Languages.* M.A. thesis, University of Texas at Arlington.

Collins, I. Vaughn. 1975. *The Austro-Asiatic substratum in Acehnese.* PhD. dissertation, University of California, Berkeley.

----- 1969. The position of Atjehnese among Southeast Asian languages. *Mon-Khmer Studies* III:48-60.

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

Cowan, H. K. J. 1933. Het Atjèsch metrum 'sandja'' in verband met een Tjamsch gedicht. *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indië* 90:149-55.

----- 1948. Aanteekeningen betreffende de verhouding van het Atjesch tot de Mon-Khmer-talen. *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indië* 104:429-514.

----- 1974. Evidence of long vowels in early Achehnese. *Oceanic Linguistics* XIII:187-211.

----- 1981. An outline of Achehnese phonology and morphology. *Bulletin of the School of Oriental and African Studies* LXIV 3:552-49.

----- 1983. The Achehnese diphthong ɯə and its possible implications for proto-Austronesian. *Acta Orientalia* 44:153-85.

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

Diffloth, Gérard. 1994. The lexical evidence for Austric, so far. *Oceanic Linguistics,* 33.2:309-21.

Durie, Mark. 1990. Proto-Chamic and Acehnese mid vowels: towards Proto-Aceh-Chamic. *Bulletin of the School of Oriental and African Studies,* Vol.53. London, University of London, pp 99-114.

Dyen, Isidore. 2001. Review of Thurgood 1999. *Anthropological Linguistics* 43.3:390-394.

Headley, Robert K. Jr. 1976. Some Sources of Chamic Vocabulary. In Jenner et al. (eds.) *Austroasiatic Studies: Part 1,* Oceanic Linguistics Special Publications. Honolulu, Unversity of Hawaii Press. pp 453-476.

Hudson, Alfred B. 1967. *The Barito dialects of Borneo: a classification based on comparative reconstruction and lexicostatistics*. Ithaca, New York, Department of Asian Studies, Cornell University.

Larish, Michael. 1999. *The Position of Moken and Moklen within the Austronesian Language Family*. PhD thesis, 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.

Niemann, G. K. 1891. Bijdrage tot de Kennis der Verhouding van het Tjam tot de Talen van Indonesië. *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indië* 40:27-44.

Reid, Lawrence, A. 1994. The morphological evidence for Austric. *Oceanic Linguistics*, 33.2:323-44.

Shorto, Harry. ms (no date). *Mon-Khmer Comparative Etymological Dictionary.* (being prepared for piublication with Pacific Linguistics)

Shorto, H.L. 1975. *Achinese and Mainland Austronesian. Bulletin of the School of Oriental and African Studies* 38:81-102.

Shorto, H.L. 1977. Proto-Austronesian *taqən: an anomaly removed. *Bulletin of the School of Oriental and African Studies* 40:128-129.

Sidwell, Paul. 2005. *The Katuic Languages: classification, reconstruction & comparative lexicon*, Munich, Lincom.

-----. 2002. Classification of the Bahnaric Languages: a comprehensive review. *Mon-Khmer Studies*, 32:1-24.

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

Thomas, Dorothy M. 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.* Oceanic Linguistics special Publications No. 28. Honolulu, University of Hawaii Press.

Zorc, David. 1995. A glossary of Austronesian Reconstructions. In Darrell Tryen (ed.) *Comparative Austronesian Dictionary*, part 1, fascicle 2, pp.905-1197. Berlin & New York, Mouton de Gruyter.

Appendix 1: Summary of Acehnese words plausibly borrowed from MK sources, extracted from Thurgood (1999) and Cowan (1948). Note: 'PC' = Thurgood's reconstructions; 'C.' forms sourced from Cowan, 'C. No.' indicates Cowan's numbered comparison. MK comparisons cited are indicative only, and should not necessarily be interpreted and indentifying the particular donor language.

Aceh. *hu* 'ablaze'
Cham *hu* 'roast' (C.)
Khmer *chur* 'to ignite' (C.); Bah. *huur* 'roast', Katu *huar* 'singe'
C. 64

Aceh. *ɟa* 'ancestor'
Cham *ɟa* 'appelative of poor people' (C.)
OldMon *'ja /ɲaʔ/*, OldKhmer *jī /ɟii/* 'great-grandmother'
C. 66

Aceh. *sapaj* 'arm'
PC *sapal*
Reflexes in Aslian, Katuic & West-Bahnaric.

Aceh. *baʔ* 'at, on' preposition
Cham *pak* 'at, towards' (C.)
OldMon *bak* 'up to, until', *pâʔ* 'for, on, on behalf of' (C.)
C. 6

Aceh. *ruəŋ* 'back'
PC *rɔŋ*
Katuic *krɔŋ* 'back', Khmu *kndrɔɔŋ* 'back'

Aceh. *suət* 'bail'
PC *sac*
PMK *saac*, widespread in MK.

Aceh. *tɤt* 'bake in fire, burn'
Khmer *tut /dot/* 'grill, roast; kindle, set fire to'(C.)
C. 140

Aceh. *suət* 'bale'
PC *sac*
PMK *saac* 'bale out' widespread in MK
C. 128

Aceh. *rɯtɯəʔ* 'bean, pea'
PC *rətaak*
Khmer *sandaek,* Iban *retak*

Aceh. *cəgɛə* 'bear'
PC *cagɔw*
EMK *ɟkaw*, Asl. *gaaw*
C. 18

Aceh. *cicem* 'bird'
PC *cim*
PMK *cim*, reflected in all brances but Khmer, note Nicobar has redup. initial.
C. 29

Aceh. *kap* 'bite'
(PC *kɛʔ*)
PMK *kap* 'bite' indicated by widespread reflexes
C. 74

Aceh. *blɛt* 'blink'
PC ?
Khmer *blet /plet/* 'appear and disappear like a flash'(C.)
C. 11

Aceh. *pot* 'blow (wind)'
PMK *puut* 'blow' (NMK, Asl.)
C. 123

Aceh. *cabuəŋ* 'branch, fork'
PC *caɓaaŋ*
Malay *cabang* > Aslian, Cf. Kh. *ɟəŋrmaaŋ* 'forked stick'?

Aceh. *picah* 'broken, break'
PC *picah*
Cf. Malay *pĕcah.* Palatal stop indicates borrowing into Bahnaric also.

Aceh. *kɯbɯə* 'buffalo'
PC *kabaw*
Aceh. = Kh. *krəbɤj*, while Chamic = Malay *kĕrbaw*

Aceh. *gulam* 'carry on shldr'
PC *gulam*
PMK *klam* or *kləm* on the basis of NMK & Aslian reflexes.

Aceh. *drɔp* 'catch, arrest'
Cowan notes Mon *rap* /*rɔp*/ 'to catch'; PMK
**rəp, *rəp* are indicated by widespread
reflexes
C. 48

Aceh. *luəŋ* 'channel'
Cham *haluŋ* 'hole, pit, canal' (C.)
Cf. Khmer *lùŋ* 'to dig hole', *ʔɔnlùəŋ* 'hole
in stream-bed'; Bahnar *səluŋ* 'pit, ditch'
C. 107

Aceh. *tijɯəp* 'chase, run aft.'
PC *tijaap
Cf. Khmu *ŋgjaap*, Tampuon *tijạap*
borrowed from Chamic.

Aceh. *let* 'chase'
Mon *lemöt nâ* 'to drive away' (with –m-
infix?) (C.)
C. 97

Aceh. *miəŋ* 'cheek, jaw'
PC *miaŋ
Cf. Khmu *miəng* 'chew', Viet. *mi ̈ng*, < PV
*mɛɛŋʔ 'mouth'

Aceh. *kɯəŋ* 'chin, jaw'
PC *kaaŋ
PMK *kaaŋ, reflexes in Katuic, Bahnaric,
Nicobarese, Vietic, Pearic.

Aceh. *kruət* 'citrus'
PC *kruac
PMK *kruəc 'citrus'
C. 88

Aceh. *cah* 'clear undergrowth'
Borrowed > Bahnaric , C. compares Khmer
ceh 'to cut with small blows'
C. 19

Aceh. *pɯdap* 'cover, to'
PMK *dəp (widespread etymon)
C. 40

Aceh. *khop* 'cover; put face down'
PMK *ckup 'cover'; PAn *kubkub 'cover'
C. 80

Aceh. *gɔm* 'cover'
PC *gəm
Cf. Khmer *kaem* 'cover, encrust, decorate',
PV *kəmʔ 'to bury'

Aceh. *lumɔ* 'cow, ox'
PC *ləmɔ
Cf. Khmu *lmboʔ*, Viet. *bò*, Malay *lĕmbu*,
may be derived from MK **[]bɔʔ* 'hump of
ox', cf. Mon *ba'* /*pòʔ*/ id.

Aceh. *ʔaʔaʔ* 'crow'
PC *ʔaak
PMK *kʔaak (all but Khasi, Nic.) – Vietic
reflexes typically *ʔaak*, e.g. Vt. ác, but
imitative! Aceh. reflex is irregular.

Aceh. *cɛh* 'crush, pulverise'
Cham *cɛh* 'hatch' (C.) ?
Khmer *ces* 'to crush' (C.); C. also compares
Bahnar *she*, Cham *cɛh* 'hatch' the
connection to 'crush' is doubtful.
C. 22

Aceh. *kɔh* 'cut off'
PC *kɔh
PMK *kɔh (Bah.,Kat.,Nic.,Asl.)
C. 85

Aceh. *ɟluəh, gluəh* 'deer (small kind)'
Khmer *chlus* 'id.'
C. 73

Aceh. *kuəh* 'dig'
PC *kuah 'shave, scrape'
PMK *kuəs 'scrape'
C. 90

Aceh. *ɟep* 'drink'
Mon *ɟøp* /*cep*/ 'sip, taste'
C. 69

Aceh. *rɯəŋ* 'dry, dry out'
Cf. Katuic: Ta'Oi *raaŋ* 'drying rack
C. 124

Aceh. *tho* 'dry'
PC *thu
Cf. Temiar *təhool*, KhmuYuan *thúu*
C. 137

Aceh. *ʔiteʔ ʔara* 'duck-wild'
PC *ʔada*
Khmer *dā* /*tĩia*/ < PMK *daʔ*, note doublets:
Srê *ʔara* / *ʔada*, Bahnar *həraa* / *tadaa*
C. 1

Aceh. *klɔ* 'dumb'
PC *k-am-lɔ*
Cf. Khmer *kamlaw* 'ignoramus'

Aceh. *ɟɯəp ɟɯəp* 'each, every'
OldMon *jāp* /*jap*/ 'all, each, every'
C. 71

Aceh. *sɔh* 'empty'
PC *sɔh*
Khmer *suh* /*soh*/; Bah., Kat. may have
borrowed via Chamic.
C. 131

Aceh. *lhuəh* 'escape'
PC *klaas*
> Bah., other MK suggests *laas* 'leave'

Aceh. *tɔm* 'ever'
Cham *tom* 'meet with, accomplised'(C.)
PMK *təm*/*təəm*/*tam* 'begin' (all MK.);
perhaps from Mon *tam* /*tɔm*/.
C. 139

Aceh. *ʔɛʔ* 'excrement'
PC *ʔɛh*
Borrowed separately, Ch. < K/V, Ach. <
**ʔɛk*
C. 51

Aceh. *toh* 'excrete'
PC *tɔh* 'remove clothing'
Cf. Khmer *tuh* /*doh*/ 'remove clothing; to
free, release'; > Bah.
C. 138

Aceh. *ba* 'father' (C.)
(PC *ʔama* < An.)
PMK *ʔbaaʔ*, cf. Khmer *baa*
C. 2

Aceh. *dit* 'few'
PC *dVt* 'small'
PMK *kdĩt*, cf. Viet. *nít*, Khasi *khyndit*; >
Bah. (T. incorrectly states "restricted to
Highlands")
C. 45

Aceh. *gap* 'firm'
PMK *gap* 'fitting, sufficient' indicated by
widespread reflexes
C. 53

Aceh. *ɲum* 'flavour'
Cham *ɲəm, ɲam* 'to taste'(C.)
Praok *ɲɔm* 'to taste', Bahnar *ɲaam*
'delicious', Khmer *ɲaaɛm* 'exclamation used
mostly by children vaunting what they are
eating or tasting'
C. 114

Aceh. *phʌ*/*pʌ* 'fly (v.)'
PC *pər*
PMK *par*. Anomalous aspiratred initial also
found in Rade: *phiər* (Durie 1990)
C. 122

Aceh. *tuwʌ* (<*wʌ* 'stray, wander' C.)
'forget'
PC *wər*
PMK *wir* &c. 'turn' (all MK groups, with
many varients)
C. 149

Aceh. *coh coh* 'frighten animals'
Cowan notes Mon *pecuh* 'to hound on, set
on as a dog'
C. 33

Aceh. *kuət* 'gather up'
PC *kuac* 'gather, amass'
kwaac 'scrape up'

Aceh. *bit* 'genuine, real'
≠Cham *bjak* (C.)
Cowan notes Khmer *bit* /*pit*/ 'correct,
certain'
C. 10

Aceh. *wɛh* 'go away, leave'
Cham *weh* 'dodge' (C.)
Cowan notes Khmer *veh* /*vèh*/ 'to slip away,
escape, dodge'
C. 144

Aceh. *lop* 'go into, under'
Cf. Old Mon *lop* /*lop*/ 'to enter': word is
widespread in MK, but vowel varies
considerably.
C. 104

Aceh. *ɟaʔ* 'go, walk'
PMK *ɟak* 'tread, set out' indicated by
widespread reflexes
C. 67

Aceh. *muh, muɯih* 'gold'
PC *ʔama(a)s*
OldMon *jimās* 'gold' (<*jās* 'shine') > Kh.
maas 'gold'

Aceh. *gət, gɛt* 'good'
Cham *gɔt* (C.) 'just'
Khmer *gat* /kɔt/ 'just, exact'(C.)
C. 55

Aceh. *rət* 'graze (on grass etc.)'
Mon *rat* /rɔt/ 'to reap': word is widespread
in MK, but vowel varies considerably.
C. 126

Aceh. *klɯəŋ* 'hawk, kite'
PC *klaaŋ*
PMK *klaaŋ* (all MK groups)
C. 84

Aceh. *gɯ-* 'he, she'
PMK *ge[e]ʔ* '3ʳᵈ person pronoun' indicated
by widespread reflexes
C. 56

Aceh. *sɯpət* 'hit with smth.'
Mon *sapot* 'stroke or rub with hand'(C.)
C. 129

Aceh. *sɯəŋ* 'house'
PC *saaŋ*
< Khmer *saaŋ* 'to build', also >Thai/Lao)

Aceh. *goh* 'hump'
PMK *guh* 'swell', e.g. Mon *kuh* 'to swell
up', Kh, etc.
C. 60

Aceh. *chen, cen* 'in love, having strong
desire'
Cham *khin* (C.)
Palaung *sin* 'desire', Viet *xin* 'beg' < PMK
siin ? (Cowan comparisons weak)
C. 26

Aceh. *panah* 'jackfruit'
Mon *panah* 'jackfruit' (C.)
C. 116

Aceh. *khem* 'laugh'
Cham *khim* 'smile'(C.)
LitMon *ki̯m* 'smile'
C. 77

Aceh. *wiə* 'left side'
PC *ʔiə̃w*
< PMK *w[i]ʔ* 'left', with metathesis in
Chamic?
C. 147

Aceh. *ɟawiə* 'left-handed'
OldMon *jwiʔ* 'left' < PMK *w[i]ʔ*
C. 68

Aceh. *buəŋ* 'morass'
Khmer *piŋ* /bɤŋ/ 'lake, pool'; > Stieng
bhəŋ (C.), > Thai *buɯŋ*
C. 17

Aceh. *le* 'more, still more'(C.)
Riang-Lang ⁻*ləj* 'more, longer, else', Viet. *lại*
'again', Mon *lē* 'also', etc.
C. 94

Aceh. *bɯŋəh* 'morning'
Cham *paguh* 'morning light'(C.)
Mon *ŋuh* 'awake out of sleep' (C.)
C. 111

Aceh. *cʌt* 'mountain range'
PC *cət*
Cf. Khmer *caot* 'high, steep, sheer, abrupt'
C. 35

Aceh. *takuə* 'neck'
PC *takuaj*
Resembles PMK *kuuj* 'head' (Kat., Asl.),
but doubtful. C. compared to a different
etymon.
C. 135

Aceh. *kɯmuən* 'nephew'
PMK *kmun, *kmuun, *kmuən* 'nephew'
C. 92

Aceh. *coŋ* 'on top of'
Cowan notes Khmer *coŋ* 'end, tip'
C. 34

Aceh. *hah* 'open (mouth)'
PC *ʔaha*
PMK *haʔ, hah*, Ach. resembles B. & Viet.
C. 61

Aceh. *gɔp* 'other, group'
PC *gɔp*
PMK *gap, gɔp* 'friend, associate' (Khm.,
Bah., Viet.)

Aceh. *lap* 'to paint'(C.)
Khmer /srlaap/ 'to rub, anoint, smear, paint'
C. 101

Aceh. *cɔh* 'peck (as snake)'
PC *cɔh*
PMK *[ʔ]cɔh* (EMK, Khmu, Asl.)
C. 32

Aceh. *weŋ* 'pedal'(D&D), 'turn around' (C.)
Cham *wiŋ* 'turn, whirl' (C.)
PMK *wiŋ* &c. 'go round' (all MK groups,
with many varients)
C. 145

Aceh. *pət* 'pick (fruit, flower)'
PC *pɛt*
MK forms suggest *pic, but connection is
questionable.
C. 118

Aceh. *tamɛh* 'pillar, post'
PC *tamɛh*
Cf Mon *tmit* 'post supporting veranda'-
doubtful.

Aceh. *bantaj* 'pillow'
PC *bantal*
Cf. Malay *bantal*

Aceh. *cubet* 'pinch'
(PC *kapit*?)
Cowan notes Khmer *cbec* 'to pinch'
C. 38

Aceh. *bət* 'pluck, uproot'
PC *buc*
Khmer *boac* 'to pull up', Mon
bot 'unsheathe'; > Bah.,Stieng *buc*; also
Malay *cabut*
C. 16

Aceh. *ʔuət* 'polish, rub clean'
Cham *ɣak* 'rub' (C.)
Lawa *ʔuət* 'wipe', Khmu *ʔɔɔt* 'scrub body'
C. 142

Aceh. *bep* 'pout like a monkey' (C.)
Cf. Khmer *bep* /pép/ 'moue des lèvres,
contracter les lèvres, grimacer' (C.)
C. 9

Aceh. *lhan, tlan* 'python'
PC *klan*
PMK *tlan* - Aceh. borrowed with apical
initial (Kh.?); Chamic < form with velar
initial (Bah./Mon?)
C. 102

Aceh. *wɔə* 'return home'
PMK *wil* &c. 'turn' (all MK groups, with
many varients)
C. 148

Aceh. *kruəŋ* 'river'
PC *krɔɔŋ*
PMK *ruŋ, *ruuŋ, *ruəŋ*, low vowel
reflexes in Bah. & Khmu'.
C. 87

Aceh. *kuət* 'scrape/clear away' (C.)
Cham *kwac* 'dig' (C.)
PMK *kwaac* 'scratch up', e.g. Khmer
khvaac, Kensiw *kəwɔɟ*
C. 91

Aceh. *kɛh* 'scratch' (D&D 'matches')
Mon *keh* 'write with stylus' < PMK
kiəs 'scratch'
C. 75

Aceh. *ŋiəŋ* 'see, look'
Aslian: Senoi, Blanya-Sakai *neŋ* 'to see'
(C.)
C. 109

Aceh. *dɯə* 'shallow'
PC *dɛl*
Mon *da* 'shallow'(C.); PMK & Aslian
reflexes show [ɛ]
C. 42

Aceh. *be* 'size, amount'
Senoi *bē* 'very' (S&B); > Stieng
C. 7

Aceh. *caŋ* 'slash, strike, slice, chop'
Mon *caŋ* 'prick, pierce' (C.), also >
Stieng.Cf. Malay *cincang*
C. 19

Aceh. *cut* 'small', *bacut* 'a little'
Aslian: Senoi *maʔcut*, Sakai *macut*
'small'(C.)
C. 39, 4

Aceh. *chuəŋ* 'smelling of urine'
Cowan notes Khasi *jung* 'urine'
C. 28

Aceh. *luɯəŋ* 'spread out'
PC **laaŋ*
PMK **laaɲ* 'spread out'
C. 99

Aceh. *wuɯə* 'stable, pen'
Cham *wa, war* 'yard (buffalo), stable' (C.)
Khmer *val /viel/* 'plain, field, clearing,
courtyard, plaza, threshing floor'; Mon *wa*
/wèa/ 'open space, pasture'
C. 148

Aceh. *dʌŋ* 'stand, stop'
PC **dəŋ*
Viet. *đứng* (doubtful); Cowan notes Mon
demɔŋ 'remain, dwell' (with infix)
C. 47

Aceh. *dəm* 'stay overnight'
Cham *dəm* (C.)
Mon *døm /tɜ̀m/* 'to lodge'(C.); PMK
*dəm is indicated by widespread reflexes
C. 46

Aceh. *cɯɯt* 'stinging pain'
Khmer *cɔt* 'sour', Stieng *cət* 'astringent'(C.)
C. 24

Aceh. *cɯɯŋɛh* 'stink, unpleasant smell'
Khmer *chʔɛh,* Mon *həʔeh,* Stieng *ciʔih* 'to
stink'(C.)
C. 23

Aceh. *gɔp* 'stranger, other'
PMK **gəp, *gap* 'friend, to associate'; C.
notes Aslian forms with semantic match
C. 59

Aceh. *bʌt* 'stretch'
Cham *but* 'twisted' (C.)
Khmer *bot /pot/* 'to curve, fold'; also >
Stieng
C. 15

Aceh. *pɔh, pɛh* 'strike, beat'
PC **pɔh*
Khmer *pah* 'hit', *poh* 'hammer', *puh* 'hit
with stick', Mon *peh* 'kick (of horse)', *kəpɔh*
'hit with hand'
C. 117

Aceh. *pɔh, pɛh* 'strike, pound'
PC **pɔh*
PMK **pah, *puh, *puəh,* NMK, Bahnaric,
Vietic.

Aceh. *kʌŋ* 'strong, hard'
PC **khaŋ*
Katuic**kəŋ,* Khmer *kɛ̀əŋ < *gaŋ* ? Chamic
< Viet. *khăŋ ?
C. 86

Aceh. *ŋɔp* 'submerged'
Khmer *ŋup* 'incline, drop', Khasi *ŋop*
'subside' (C.)
C. 110

Aceh. *ba* 'take, carry'
PC **ba*
OldKhmer *va,* Temiar *baʔ* 'carry on back'
C. 3

Aceh. *cɔʔ* 'take, seize'
Cham *cɔk* (C.)
WestBahnaric **cɔk* 'take'; Khmu *cɔɔk*
'catch (e.g. pig)', *cok* 'take out (e.g.
entrails)' although other MK suggest **jɔ(ɔ)k,*
e.g. Khmer *jɔɔk* 'take'.
C. 31

Aceh. *criəʔ* 'tear, rip'
Khmer *criək* 'to split'(C.)
C. 36

Aceh. *sideh* 'that, there'
PC **dih*
Mon *deh* 'he or she (disrespectful)' (C.)
C. 41

Aceh. *bʌh* 'throw away'
Khmer *poh* /*bɔh*/ 'to throw'
C. 14

Aceh. *wɛt* 'turn'
PMK **wac* 'twist', e.g. Bah. *wɛc* 'twist',
Mon *wòt* 'wring out' etc.
C. 146

Aceh. *plòïh* 'unroll' (C.)
Mon *plöh* 'untwist'(C.)
C. 121

Aceh. *that* 'very'
Mon *that* /*thɔ̀t*/ 'well, healthy, strong',
Khmer *hat* 'to exert', *that* 'large, obese'
(C.)
C. 136

Aceh. *suʔuəm* 'warm'
Khmer *sʔɔm* 'to heat, warm'(C. compares a different Kh. root)
C. 130

Aceh. *rhah* 'wash'
PC **raw*
Aceh. cf. Viet. *rửa* (< **raah* ?), Chamic cf. Bah., Kat. **ʔaraaw*
C. 133

Aceh. *sɔm* 'wrap'
PC **səm*
Old Khmer *sum* 'to wind, roll, wrap up'

Aceh. *luən* 'yard'
Khmer /diilaan/, /lan/ 'flat open area, square, yard'
C. 98

Aceh. *sɯmɯŋɯp* 'yawn'
PC **həʔaap*
PMK **sʔaap, *sŋʔaap*, not all MK subgroups have medial nasal

Summary of rejected comparisons from Cowan (1948):

Phonological correspondence(s) defective: 5, 12, 13, 21, 25, 26, 30, 37, 43, 54, 57, 62, 70, 89, 93, 96, 100, 105, 106, 108, 119, 125, 127, 134, 135, 143
Semantic comparison unconvincing: 103, 113
An. or Malay: 8, 27, 65, 132
Indic: 120
Expressive/sound symbolic: 63, 82, 83, 115
No resemblant forms found beside obvious loans into Bahnaric: 49, 50, 58, , 76, 78, 79, 95, 112, 141

Appendix 2: Basic vocabulary of Acehnese, Proto-Chamic, Proto-Malayic, 183 items.

Sematic	Acehnese	P-Chamic	P-Malayic	Commentary
above/on top	*ʔatuɯəh*	**ʔataas*	**atas*	All < PAN **Caʔas*
ashamed	*malɛə*	**malɔw*	**malu*	All < Malayo-Chamic etymon
ashes	*abɛə*	**habɔw*	**habu*	All < PAN **qabúH*
at	*di*	**di*	**di*	All < PMP **di*, although the failure to diphthongise in Aceh.-Chamic is odd.
back (anat.)	*ruəŋ*	**rɔŋ*	**bΛlakaŋ*	Aceh-Chamic replaced by MK, Cf. Bahnar *rɔŋ*, Khmu *kndrɔɔŋ*. Note: Bahnaric may have back-borrowed from Chamic, the original MK form retained in West Bahnaric **krɔŋ* 'back of knife blade'
bad	*ɟuhut*	**ɟəhaat*	**ɟahət*	All < PMP **zaqát*
belly/guts	*pruət*	**pruac*	**pərut*	Metathesis in Aceh-Chamic
below	*baroh*	**ʔala*	**babah*	Aceh corresponds to Iban *baruh* and Maningkabau *baruʔh*; Chamc obscure
big	*raja, rajəʔ*	**raja*	**raja*	All < PAN **Raja*
bird	*cicem*	**cim*	**buruŋ*	Aceh-Chamic borrowed < MK **cim*
bite	*kap*	**kɛʔ*	**gigit*	Aceh < MK **kap*; Chamic form obscure
black	*ʔitam*	**hitam*	**hitəm*	All < PAN **qitém*
blood	*darah*	**darah*	**darah*	All < PAN **dáRaq*
blow	*jop*	**ʔajup*	**t/iup*	All < PAN **Siúp*
bone	*tuluɯəŋ*	**tulaŋ*	**tulaŋ*	All < PHF **CuqelaN*
branch	*dhuɯən*	**dhaan*	**dahan*	All < PMP **daqan*
breast	*tɛʔ, dɛʔ*	**tasɔw*	**susu(ʔ)*	Aceh. < Malay *tetek*; Chamic shares initial stop with Iban *tusu*
breathe	*naphãh*	**ɲawa*	**ɲawa*	Aceh. < Malay *napas* < Arabic; Malayo-Chamic < PMP **ɲáwa*
burn	*tət*	**ɓəŋ*	**bakar*	All three apparently innovated; Cf. OKhmer *tut* (*dɔt*) 'brûler'
buy	*blɔə*	**blɛj*	**bəli*	All < PAN **bĕlí*
chew	*mamʌh*	**mamah*	**mamah*	All < PMP **mamáq*
child	*ʔanɯʔ*	**ʔanaak*	**anak*	All < PAN **aNak*, widely borrowed (via Malay?) in SEAsia
choose	*pileh*	**ruah*	**pilih*	Aceh. & Malayic < PAN **píliq*, Chamic borrowed from MK, Cf. Khmer *rɤh*, Stieng *rɔɔjh,* although the Chamic vocalism is not explained
claw/nail	*gukɛə*	**kukɔw*	*kuku*^{Malay}	All < PAN **kuS+kuS*
climb	*ʔeʔ*	**dĩʔ*	**naik*	All < PMP **nahik*
cloud	*awan*	**hual*	**a(bw)an*	Aceh. borrowed Malay *awan*, Chamic obscure
cold	*siɟuək, lɯpiə*	**laʔən*	**diŋin*	Aceh. borrowed Malay *sejuk*, other Malayic < PMP **diŋ+diŋ*; Chamic

				obscure
come/arrive	*troh*	**truh*	**datən*	Aceh-Chamic etymon is shared with North+Central Bahnaric, source unknown.
cook	*tagun*	**tanak*	**tanak*	all < PHF **taNek* , assuming that Aceh. shows metathesis
count	*bilwəŋ*	**jaap*	**hituŋ*	Aceh. < PHF **bílaŋ*; Chamic < PHF **Hiáp*; Malayic < PAN **qi-(n)tuŋ*
cry/weep	*kliʔ, mɔə*	**cɔk*	**taŋis*	Malayic < PAN **Cáŋis*, Aceh. & Chamic forms obscure
cut/hack	*tɛktɛk*	**tarah*	**tətək, *taRas*	Aceh. & Malayic < PAN **tek+tek*, Chamic & Malayic < PAN **taRáq*
day/sun	*ʔurɔə*	**hurɛj*	**hari*	All < PAN **waRiH*
die	*mate*	**mataj*	**mati*	All < PAN **maCéj*
dig	*kuəh*	**kalɛj*	**kali*	Chamic & Malay < PAN **kálih*, Aceh. appears to have borrowed from MK, Cf. Bahnar *kwajh* 'dig up, scratch around for'
dirty	*kutɔ, tibəh, mɨlutəŋ*	**chəp, *grit*	**kamah/ *kumuh*	Aceh. *kutɔ* from Malay *kotor*, but other forms obscure.
dog	*ʔasɛə*	**ʔasɔw*	**asuʔ*	All < PAN **asu*, with semantic shift > 'canine' in Malay
dream	*lumpɔə*	**lumpɛj*	**m/impi/ *impi*	All < PMP **nipi*, note the Aceh-Chamic shift **n- > *l-*
drink (water)	*minom*	**minum*	**inum*	All < PMP **inúm*
dry	*kraŋ, tho*	**raŋ, *thu*	**kəriŋ*	All < MP doublet **kaRaŋ/*kaRiŋ*, plus Aceh-Chamic has innovated **thu* - origin obscure
dull/blunt	*tumpoj*	**ʔabual*	**tumpul*	Aceh. & Malayic < PAN **dump+pel* , Chamic obscure
dust	*dhoj, ʔabɛə*	**dhual/r*	**dəbu*	Aceh. + Malayic < PMP **debu*, but **dhual/r* (more probably **dhul*) is obscure
ear	*guliɲuəŋ*	**təliŋa*	**tAliŋa(ʔ)*	All < PHF **taŋíla*
earth/soil	*tanɔh*	**tanah*	**tanah*	All < PMP **tanaq* or **taneq*
eat	*makɯən*	**ɓəŋ*	**ma/kan*	Aceh. & Malayic < PAN **kán*, Chamic obscure
egg	*bɔh*	**bɔh*	**təlur*	Aceh-Chamic replaced PAN **tĕlúR* 'egg' - Thurgood suggests **bɔh* < PAN **buáq* 'fruit', although the vocalism is problematic
eye	*mata*	**mata*	**mata*	All < PAN **maCá*
fall down	*rhət*	**labuh*	**labuh*	Chamic & Malayic < PMP **ka-nabúq*, Aceh. obscure
far/distant	*ɟuʔoh*	**dɔh*	**ɟauh*	Aceh. & Malayic < PMP **Zaúq*, Chamic obscure

fat, grease	*gapah*	**ləmaʔ*	**ləmək*	Chamic + Malayic < PMP **lemak*; Aceh. obscure
father	*ʔajah, jah, ʔa bu, du, abi*	**ʔama*	**apa(ʔ)*	Aceh. forms all secondary; Chamic < PAN **ama*, Malayic < PHN **bapaʔ*
fear, afraid	*takot*	**huac*	**takut*	Aceh. + Malay(ic) < *PAN **tákut*, Chamic obscure
feather	*bulɛə*	**bulɔw*	**bulu*	All < PMP **búlu*
fire	*ʔapuj*	**ʔapuj*	**api*	All < PAN **Sapúj*
fish (n.)	*ʔuŋkot*	**ʔikaan*	**ikan*	Chamic & Malayic < PAN **Si-káʔen*; Aceh. obscure
flow	*ʔile*	**ɗuac*	**alir*	Aceh. & Malayic < PMP **a+liR*, although Aceh. may have borrowed Minangkabau *iliʔ*; Chamic obscure
flower	*buŋɔŋ*	**buŋa*	**buŋa(ʔ)*	All < PMP **búŋah*
fly (v.)	*phʌ, pʌ*	**pər*	**tʌr(ə)baŋ*	Aceh-Chamic has borrowed < MK, Cf. PMK **par*
foot/leg	*gaki*	**kakaj*	**kaki*	Aceh. has borrowed directly from Malay(ic).
forest	*ʔutwən*	**hutaan*	**hutan*	All < PMP **qutan*
four	*puət*	**paat*	**əmpat*	All < PAN **Sĕ(m)pát*
full (sated)	*punɔh, trɔə*	**trɛj*	*penuh^Malay*	Aceh. & Malay < PMP **pĕnúq*, + Aceh-Chamic innovated
give	*bri, ɟok*	**brɛj*	**bəriʔ*	Chamic & Malayic < PAN **bĕRáj*, Aceh. has borrowed Malay *beri* & an MK form, Cf. Khmer *jɔɔk* 'take'
good	*gʌt, gɛt*	**biaʔ, gɔt^Cham*	**baik*	Aceh. + Cham < Khmer *gɔt /kɔt/*
grass	*naluəŋ*	**rək*	**rumput*	All show independent innovation
green	*ʔijo*	**hiɟaw*	**hiɟaw*	All < Malayo-Chamic etymon
grow	*timoh*	**tamuh*	**t/um/buh*	All < PAN **Cú(m)buq*
hair (of head)	*ʔok*	**ɓuk*	**buø(uə)k*	All < PAN **buSék*
hand	*ɟarɔə*	**taŋaan*	**taŋan*	Cf. Malay *jari* 'finger'. Acehnese shares with Iban the semantic shift 'finger' > 'hand', using the compound *ʔanɨk ɟarɔə* 'child hand' for 'finger'. Chamic **caɗiaŋ* 'finger' borrowed from unknown source.
he/she	*ɟih*	**ɲu*	**ia*	Chamic correspondes to Minangkabau *iɲo*; Malayic < PAN **síá*; Aceh. shows a variety of forms
head	*ʔulɛə*	**ʔakɔʔ*	*kepala^Malay*	Aceh. regularly < PMP **qúluH*; Malay < Indic; Chamic < MK, Cf. Mon *kɔʔ* 'neck'
hear	*duŋʌ, luŋʌ; simaʔ* 'listen attentively'	**həməʔ*	**dəŋər*	Aceh. + Malayic < PMP **d₃ĕ+ŋéR*, although Cf. PMK **[t₁]ŋər*, e.g. Viet. *nghe* 'to hear', RiangLawa *⁻təkŋar* 'to listen'; The Aceh-Chamic *simaʔ/*həməʔ* etymon is obscure.

heavy	*ghən, brat*	**traap*	**bərat*	Aceh. * Malayic < PMP **beRʔat*; other Aceh. and Chamic obscure
hit/slap	*tampa*	**pah*	*tampar*^Malay	Aceh. < Malay; Chamic < MK, Cf. Khmer *pah* 'hit'
strike/beat	*pɔh, pɛh*	**pɔh*	**pukul*, **paluʔ*	Aceh. & Chamic < MK, Cf. Khmer *pah* 'hit', *poh* 'hammer', *puh* 'hit with stick', Mon *peh* 'kick (of horse)', *kəpɔh* 'hit with hand'
hold	*ruigam, mat*	**ʔɟaaʔ*, **ʔapan*	**pəgaŋ*	Aceh. < PAN **gem +gem*, Chamic & Malayic obscure although MK forms such as OldMon *bgan* 'to yoke, take hold of' are suggestive
horn	*luŋkɛə*	**tuki*	*tanduk*^Malay	Ache-Chamic has borrowed from MK, the etymon is found in Bahnaric & Katuic, Cf. Bahnar *ʔəkɛɛ*
house	*suɯəŋ*	**saaŋ*	**rumah*	Malayic < PAN **Rumaq*, Aceh. & Chamic borrowed , Cf. Thai/Lao *saaŋ* 'granary, warehouse'
I	*kɛə*	**kɔw*	**aku*	Aceh-Chamic < PAN **ku*, Malayic < PAN **akú*
inside	*dalam*	**dalam*	**(d-)aləm*	All < PAN **d₂á+lem*
knee	*tuʔot, tuʔot*	**tuʔut*	**tuʔ(uə)t*	All < PHF **túSud*
know (things)	*thɛə*	**thɔw*	**tahu*	All < PMP **taqú*
lake	*danɔ*	**danaw*	**danaw*	All < PAN **dánaw*
laugh	*khem*	**klaw*	**tawaʔ*	Malayic < PAN **Cáwa*, Aceh. & Chamic forms obscure.
leaf	*ʔon*	**sula*	**daun*	Aceh. & Malayic < PMP **d₂ahun*, Chamic < MK, Cf. PMK **slaʔ*
left side	*wiə*	**ʔiə̃w*	**kA-iri/*kibaʔ*	Malayic < PAN **ka-wiRí*; Aceh-Chamic < MK Cf. Khmu *trweʔ*, Jenai *wĩʔ*, Mon *c'wei* (with metathesis in Chamic and > Bahnaric).
lightning	*kilat*	**kataal*	**kilat*	Aceh. & Malayic < PHF **kilát*; Chamic is obscure, but could be derived by metathesis
live	*ʔudep*	**hudip*	**hudip*	All contine PAN **qúd₂ip*
liver	*ʔate*	**hataj*	**hati*	All < PAN **qaCéj*
louse	*gutɛə*	**kutɔw*	**kutu*	All < PAN **kúCuH*
man/male	*lakɔə*	**ʔakɛj*	**laki*	All < PMP **láki*
many	*lə*	**lu*	*banyak*^Malay	Aceh-Chamic obscure
meat/flesh	*ʔasɔə*	**ʔasɛj*	**isiʔ*	All contine PAN **Sesi* (Malayic also innovated **dagiŋ*)
moon	*buluɯən*	**bulaan*	**bulan*	All < PAN **bulaN*
mosquito	*ɟamɔʔ, ɲamɔʔ*	**ɲamuk*	**ɲamuk*	All < PMP **ɲamúk*
mother	*maʔ, ma*	**mɛʔ*	**(ə)ma(ʔ)*	Aceh. corresponds to Malayic, Chamic resemble numerous MK forms

				suggesting PMK *mee?
mountain	*gunɔŋ, cɔt/cʌt*	*cɔt	*gunung*^Malay	Cf. Khmer *cót* 'escapé'; Aceh. *gunɔŋ* < Malay
mouth	*babah*	*babah	*mulut	Aceh-Chamic < PMP *baqbaq
name	*nan*	*ʔanan	(Malay *nama* < Skt.)	Aceh-Chamic etymon obscure, borrowed into Bahnaric, Cf. Bahnar *ʔənan*
narrow	*ʔubit/ʔubɯt*	*ganiat	*səmpit	Aceh. and Malayic may reflect independent varients of PMP *kapit*; Chamic obscure
near	*tɔə, rap*	*ɟɛʔ	*dəkək	All show independent developments
neck	*takuə*	*takuaj	*lihər	Aceh-Chamic resembles PMK *kuuj* 'head'
needle	*ɟarom*	*ɟarum	*ɟarum	All < PAN *ZáRum
new	*baro*	*bahrɔw	*baharu?	All < PAN *baq(e)RuH
night	*malam*	*malam	*ma-lə(hø)m	All < Malayo-Chamic etymon
nose	*ʔidoŋ*	*ʔiduŋ	*hiduŋ	All < PAN *i+júŋ
not	*hʔan, tan*	*ɓuh...ʔɔh	*-daʔ	All show independent developments
old (person)	*tuha*	*klap	*tuha(?)	Aceh. & Malayic < PAN *tuqáS; Chamic obscur
one	*sa*	*sa	*əsa?	All < PAN *sa
open/uncover	*pɯhah*	*pəh	*buka?	Chamic < MK, Cf. Bahnar *pɔh*, Palaung *pɯh*, Aceh. Cf. Viet. ha?; Malayic < PMP *buká?
other	*bukʌn*	*bukən	*bukən	All < Malayo-Chamic etymon
person/human	*ʔurɯəŋ*	*uraaŋ	*uraŋ	All < Malayo-Chamic etymon
rain	*ʔuɟɯən*	*huɟaan	*huɟan	All < PAN *quZáN
rat	*tikoh*	*tikus	*tikus	This Malayo-Chamic etymon resembles MK words for 'porcupine', e.g. PWaic *ŋkos*, PSemai *kuus*; also borrowed into Moken as *koh* 'porcupine'
red	*mirah*	*mahirah	*(ma-)irah	All < PMP *ma+iRaq
right side	*ʔunun*	*hanuã?	*k/anan	Aceh. corresponds to Malayic. Chamic is obscure, but is perhaps an infixed reflex of the same etymon as Minangkabau *suo?* 'right side'
road/path	*ɟalan*	*ɟalaan	*jalan	Aceh. < Malay(ic) (otherwise *ɟaliən* expected)
root	*ʔukhɯə*	*ʔughaar	*akar	Aceh-Chamic < PMP *wakaR (note influence of *w* on minor-syllable vocalism), Malayic < PMP *akaR
rope/string	*talɔə*	*talɛj	*tali	All < PAN *Calís
rotten	*broʔ*	*bruʔ	*busuk	Aceh-Chamic < PAN *buRúk, Malayic < PMP *busuk
salt	*sira*	*sira	*sira, garam*^Malay	All < PAN *qasiRa, plus some replacement with *garam* in Malay and others

sand	*Ɂanɔə*	**cuah*	**pasir*	Aceh. < *PAN *qĕnaɟ; Chamic & Malayic independently innovated
say/speak	*muɯtuto*	**lac*	**tutur*	Aceh. corresponds to Malayic; Chamic etymon obscure
scratch (itch)	*krut*	**kabac*	**garut, *garuk*	Aceh. & Malayic < PMP *ka+Ruɗ; Highlands Chamic borrowed from Bahnaric, Cf. Bahnar *kəbaɟɁ,* infixed PMK *kaac
sea/ocean	*laot*	**tasiɁ*	**tasik*	Malayo-Chamic < PMP *tasik, Aceh. borrowed Malay *laut*
see	*kalʌn, ŋieŋ, Ɂɯ*	**ɓuh*	**lihat*	Aceh. forms obscure; Chamic > Bahnar *ɓoh,* Cf. also OldMon /təmɓah/ 'to appear'
sew	*cɔp*	**ɟahit*	**ɟahit*	Chamic & Malayic < PMP *záɋit, Aceh. obscure
sharp	*taɟam*	**haluaɁ*	**taɟəm*	Aceh. & Malayic < PMP *tazím, Chamic obscure
shoot (arrow)	*panah 'arrow'*	**panah*	**panah*	All < PAN *panaq
shoulder	*baho*	**bara*	*(PAN *qabáRaH)*	Aceh. < Malay *bahu*
sick	*saket*	**sakit*	**sakit*	All < PMP *sakít
sit	*duəɁ*	**dɔɔk*	**duduk*	All < PMP *d₂uk+d₂uk, note: Aceh. resembles Minangkabau *duduəɁ,* Chamic vowel quality is not explained
skin	*kulet*	**kulit*	**kulit*	All < PAN *kúliC
sky	*laŋet*	**laŋit*	**laŋit*	All < PAN *láŋit
sleep/lie down	*Ɂeh*	**ɗíh*	**tidur*	Aceh-Chamic < PMP *hideRáɋ 'lie down'; Malayic < PAN *tid₂ur 'to sleep'
small	*Ɂubɯut, Ɂubit, cut*	**dVt*	**kəcil, *kətik*	Aceh. & Chamic forms obscure
smoke	*Ɂasap*	**asap*	**asəp*	All < Malayo-Chamic the etymon
snake	*Ɂulɯə*	**Ɂular*	**ulər*	All < PAN *ulaR
sniff, smell	*com*	**cum*	*ciumMalay*	Malayo-Chamic etymon of obscure origin, also borrowed into North & Central Bahnaric
spider	*rambiduɯən*	**waɟ*	**lawaɁ, *laba(Ɂ)*	Aceh. appears to correspond, at least partially, to Iban *əmpəlawaɁ,* Highlands Chamic has borrowed a word meaning 'turn' (> 'spin (web)' Cf. Bahnar *waaɟ* 'roll up, turn'
spit	*ludah, rudah*	**kacua, *kacuh*	**ludah*	Aceh. borrowed < Malay; Malayic < PMP *luZáɋ; Chamic < MK, Cf. Khmu *kɟuh,* Bahnar *ksɔh*
split (v.t.)	*plah*	**blah*	**bəlah*	All < PAN *bĕ+láɋ
squeeze	*ɟuɯpat/ɟupat, prah*	**kapit, *cupa/et*	**pərəs, *pərah*	Aceh. and Malayic < PMP *peRáɋ, while Aceh-Chamic has borrowed a prefixed from of PMK *pat
stab	*tɔp*	**kləp*	**tikəm,*	Aceh. and Chamic have independently

			*tusuk	borrowed from MK while Malayic < AN etyma
stand/stay	*dʌŋ*	**dəŋ*	**diri*	Malayic < PMP **d₂iRi*; Aceh-Chamic resembles Viet. *đứng*, 'be standing, to set' but initial voicing is problematic, an alternative comparison is PMK **duŋ* 'house'
stand up	*bɯdəh*	**taguuʔ*	*bangun*^Malay	?
star	*bintaŋ*	**bituʔ*	**bintaŋ*	Chamic < PAN **bi-(n)túqen*, while Aceh. has borrowed the Malayic varient with final velar nasal
steal	*puplɯŋ, cuə*	**klɛʔ*	**maliŋ*	Aceh. *puplɯŋ* relates to Malayic, but *cuə* is obscure, as is Chamic **klɛʔ*
stick (wood)	*kajɛə* 'wood'	**kajɔw* 'tree, wood'	**kajuʔ*	All < PAN **kájuH*
stone	*batɛə*	**batɔw*	**batu*	All < PAN **batú*
suck, sip	*hirop, pĩəp*	**sarip, *mam*	**hiRup*^PMP, **hi(ŋ)səp*	Aceh. *pĩəp* plausibly < Malayic **hi(ŋ)səp*, Chamic **mam* is clearly a nursery word
swell (abscess)	*barah*	**barah*	*barah*^Malay	All < PMP **baReq*
swim	*laŋuə*	**luaj*	**(mb)A-rənaŋ*	Aceh. < PHF **laŋúj*, Chamic is replaced by MK
tail	*ʔiku*	**ʔiku*	**ikur*	All < PAN **íkuR*
that (far)	*ɲan, nan*	**ʔanan*	**(i)na(n), *(a)na(ʔ)*	All < PAN **i-náʔ*
thick	*tɯbaj*	**kapaal*	**təbəl*	Chamic < PMP **kapal*; Aceh. & Malayic appear to reflect MK loan, Cf. PMK **[t]бəl*
think	*pike*	**saniŋ*	——	Aceh. < Malay *pikir* < Arabic; Chamic is obscure
this (near)	*nɔə*	**ʔiniʔ, *inɛj*	**(i)ni(ʔ)*	All < PAN **i-ní*
three	*lhɛə*	**klɔw*	**təlu*	All < PAN **tĕlú*
thunder	*gɯlantɯə*	**grəm*	**guntur*	Aceh. corresponds to Malayic, plus *-l-* infix which MK languages use to indicate repeated action; Chamic < MK, Cf. PMK **grəm[ʔ]*
tie/fasten	*ʔikat*	**ʔikat*	**ʔikət*	All < PMP **hi+ket*
tongue	*dilah, lidah*	**dilah*	**dilah*	All < PHF **d₃ílaq* 'lick', Aceh. also shares metathesised reflex with Malay
tooth	*gigɔə*	**gigɛj*	**gigi*	All < the Malayo-Chamic etymon
true	*bəna*	**biaʔ*	**bənər*	Aceh. & Malayic < PMP **bener*, while Chamic has merged with **biaʔ* 'good'
turn over	*baleʔ*	**blək*	**biluk*	Aceh-Chamic < PAN **balík* 'turn around'
two	*duwa*	**dua*	**dua(ʔ)*	All < PAN **d₃uSá*

vomit	*muntah*	**patah*	**m/u(n)tah*	Aceh. < Malay; Aceh-Chamic < PAN **utaq+m*
walk/go	*gaki*	**labaat,* **naw*	**((mb)Ar)ɟalaŋ*	Aceh. borrowed Malay *kaki*
warm	*suʔuəm*	--	**panas*	Aceh. < MK, Cf. Khmer *sʔɔm* 'warm'; Malayic < PMP **panas*
water	*ʔiə*	**ʔiar*	**air*	All < PMP **wáhiR*
we (excl.)	*kamɔə*	**kamɛj*	**kami*	All < PAN **kamí*
wet	*basah*	**basah*	**basah*	All < PMP **basáq*
what?	*puɯə, puə*	**hagɛt*	**apa*	Aceh. & Malayic < PMP **apa*, Chamic obscure
white	*puteh*	**putih*	**putih*	All < PAN **putíq*
who?	*sɔə*	**sɛj*	**sai, *si-apa*	All < PMP **i-sai*
wind	*ʔaŋɛn*	**ʔaŋin*	**ʔaŋin*	All < PMP **háŋin*
wing	*sajɯəp*	**sajaap*	**sajap*	All < PHN **sajap*
woman/female	*binɔə*	**kumɛj*	**bini*	Aceh. & Malayic < PMP **ba-b(in)áHi*, Chamic obscure
work, do	*buət*	**buat,* **bruãʔ*	**buat*	All < PAN **buhat*, Chamic **bruãʔ* borrowed into some Katuic & Bahnaric langs., but origin obscure, possibly secondary from **buat*
worm	*ʔulat*	**hulat*	**hulət*	All < PAN **qúleɟ*
yawn	*suɯuɯŋuɯp*	**həʔaap*	**uap*	Aceh-Chamic < MK, Cf. Khmer *sŋaap,* Bahnar *kəʔaap*; Malayic < PAN **Suab*
year	*thon*	**thun*	**tahun*	All < PMP **taqún*
yellow	*kunɛŋ, kuɲɛt* 'tumeric'	**kuɲit*	**kunit,* *kuning^Malay*	Malayic forms indicate **kuniŋ* yet Adelaar reconstructs **kunit* from PMP **kuniɟ*. Both are found in Aceh.
you (pl.)	*kah*	**hã*	**kamu(ʔ)*	Malayic < PAN **kamú*, Chamic < MK(?), Aceh. obscure
you (sg.)	*gata, kah*	**ih*	**kau*	Malayic < PAN **i-kaSú*, Chamic/Aceh.?